Étude des dynamiques du réservoir hydrologique continental de surface par réflectométrie GNSS : de l'in-situ au spatial

Pierre Zeiger

THÈSE

En vue de l'obtention du

DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

Délivré par : *l'Université Toulouse 3 Paul Sabatier (UT3 Paul Sabatier)*

Présentée et soutenue le *02/12/2022* **par :**

Pierre ZEIGER

Étude des dynamiques du réservoir hydrologique continental de surface par réflectométrie GNSS : de l'in-situ au spatial

---

JURY

Adriano CAMPS	Professeur d'université, UPC	Rapporteur
Marie Paule BONNET	Directrice de recherche, IRD	Rapporteure
Manuel MARTÍN-NEIRA	Senior Engineer, ESA	Examinateur
Catherine PRIGENT	Directrice de recherche, CNRS	Examinatrice
Nicolas BAGHDADI	Directeur de recherche, INRAE	Examinateur
Mehrez ZRIBI	Directeur de recherche, CNRS	Président du jury
Frédéric FRAPPART	Directeur de recherche, INRAE	Directeur de thèse
José DARROZES	Maître de conférences, UT3	Co-directeur de thèse

---

École doctorale et spécialité :
SDU2E : Océan, Atmosphère, Climat
Unité de Recherche :
LEGOS (UMR 5566)
Directeur(s) de Thèse :
Frédéric FRAPPART et *José DARROZES*
Rapporteurs :
Adriano CAMPS et *Marie Paule BONNET*

Résumé —

Les variations spatio-temporelles du cycle de l'eau restent aujourd'hui sujettes à de nombreuses incertitudes. Pourtant, la ressource en eau est d'une importance capitale pour la vie sur Terre et les populations humaines, sur les plans sanitaire, de l'industrie et de l'agriculture. Les écosystèmes naturels dépendent aussi de la disponibilité de l'eau, et les interactions entre les cycles énergétique, hydrologique et biogéochimiques ont des répercussions directes et importantes sur le climat. Dans le cadre de l'adaptation au changement climatique, de nombreuses politiques publiques et actions environnementales ont besoin d'estimations fiables des stocks d'eau et de leur évolution dans le temps. Cela nécessite la connaissance des niveaux d'eau et de l'étendue des zones inondées, qui est actuellement très mal contrainte. En particulier, dans les grands bassins tropicaux de l'Amazone et du Congo, la plupart des techniques de télédétection sont limitées par les denses forêts tropicales. Les seules cartes globales avec des longues séries temporelles ont été dérivées de capteurs micro-ondes passifs (par exemple GIEMS), et leur résolution spatiale de 0,25° est trop faible pour de nombreuses applications hydrologiques. Cette équation en apparence insoluble peut bénéficier de l'apport de la réflectométrie GNSS, ou GNSS-R. Il s'agit d'une technique de télédétection consistant à capter les signaux GNSS en bande L d'ordinaire utilisés pour le positionnement, pénétrant mieux la végétation que les mesures micro-ondes à plus hautes fréquences, après qu'ils aient été réfléchis par la surface terrestre. Les signaux réfléchis sont notamment sensibles à la présence d'objets d'eau. Les applications hydrologiques incluent à la fois le suivi de niveaux d'eau à l'échelle locale, et la détection de grandes zones inondées dans les forêts tropicales. Ce dernier point a été rendu possible par le lancement en 2016 d'une mission GNSS-R spatiale collectant de nombreuses observations sur la bande intertropicale, nommée CYGNSS. Les différentes applications hydrologiques et échelles spatiales ont été au cœur de mon travail de thèse. Tout d'abord, j'ai étudié la problématique locale du suivi des niveaux d'eau sur les rivières par GNSS-R in-situ. Les méthodes existantes étaient développées pour des cas d'étude océaniques, et limitées sur les rivières du fait de mesures plus bruitées, d'un nombre de satellites GNSS visibles plus faible, ainsi que de variations plus rapides des niveaux d'eau. Tout cela engendrait un bruit important que j'ai pu filtrer grâce à l'ajout de méthodes itératives aux techniques existantes. Le résultat permet de calculer les niveaux d'eau des rivières pour des variations de hauteurs d'eau allant jusqu'à 1 mm/s. Cela ouvre des perspectives intéressantes pour le suivi des crues à l'aide d'une instrumentation à bas coût. Le second axe de mon travail a consisté à étudier la dynamique des inondations avec CYGNSS, sur toute la bande intertropicale. Une première étude a permis de constater que les variables dérivées des observations CYGNSS peuvent détecter les eaux de surface, malgré l'influence d'autres paramètres géophysiques: humidité des sols, biomasse, rugosité de la surface réfléchissante (eau ou sol), entre autres. Une classification dynamique de ces variables a permis d'extraire une cartographie à 0,1° de résolution spatiale des zones inondées. Une seconde étude visait ensuite à extraire, à partir des variables CYGNSS, la fraction d'eau contenue dans chaque pixel de manière hebdomadaire. Pour cela, j'ai implémenté une régression linéaire dont les paramètres dépendent de la densité de la biomasse qui atténue le signal GNSS. Les résultats montrent des dynamiques spatiales et temporelles très similaires à plusieurs cartes régionales

et à GIEMS, tout en pointant certains biais de ce dernier. Les fractions d'eau CYGNSS permettent une détection presque exhaustive des grandes zones inondées dans la bande intertropicale, et l'étude de leur dynamique saisonnière.

Mots clés : GNSS-R, CYGNSS, hydrologie, inondations, hauteurs d'eau, in-situ

Abstract —

Spatial and temporal variations of the water cycle are still subject to large uncertainties. Yet, freshwater is essential for life on Earth and human populations, for health, industry and agriculture. Natural ecosystems also depend on freshwater availability. The interactions between the energy, water and biogeochemical cycles directly impact the climate. Within the frame of climate change mitigation, public and environmental policies need reliable estimates of water stocks and their dynamics. The knowledge of water levels and flooded extent are required, and the latter is currently poorly constrained. It is particularly the case for the large Amazon and Congo basins, where most remote sensing techniques are limited by dense tropical forests. The only global maps of surface water extent (*e.g.* GIEMS) were derived from passive microwave sensors, and their spatial resolution of 0.25° is too low for many hydrological applications. For this, the contribution of GNSS reflectometry (GNSS-R) can be very useful. GNSS-R is a remote sensing technique that uses L-band GNSS signals after their reflection on the Earth's surface. These signals dedicated to positioning penetrate well the vegetation layers. The reflected signals are particularly sensitive to the presence of water bodies. Hydrological applications include both the monitoring of water levels at local scale, and the detection of large flooded areas in tropical forests. The latter was made possible by the launch of a spaceborne GNSS-R mission named CYGNSS in 2016, which collects numerous observations over the intertropical band. The different hydrological applications and spatial scales have been the focus of my thesis work. First, I studied water level on rivers monitored by in-situ GNSS-R. Existing methods were developed for oceanic case studies, and they are limited over rivers due to noisier measurements, a smaller number of GNSS satellites in visibility, as well as faster water level variations. All this generated a lot of noise that I was able to filter out by adding iterative methods to the existing techniques. The result allows to compute river water levels with variations of the reflecting surface up to 1 mm/s. The perspectives are interesting for monitoring flash floods using low cost instrumentation. The second axis of my work consisted in studying flood dynamics over the whole intertropical band with CYGNSS. A first study showed that the CYGNSS-derived parameters detect surface water, despite the influence of other geophysical parameters: soil moisture, biomass, soil or water roughness, among others. A dynamic classification of these variables was performed to extract a 0.1° spatial resolution static map of flooded areas. A second study aimed at extracting weekly water fractions in each pixel from CYGNSS parameters. For this, I implemented a linear regression whose parameters depend on the biomass density (which attenuates GNSS signals). The results show spatial and temporal dynamics very similar to regional reference maps and GIEMS, while pointing out some biases of the latter. CYGNSS water fractions allow an almost exhaustive detection of large flooded areas in the intertropical band, and the study of their seasonal dynamics.

Keywords: GNSS-R, CYGNSS, hydrology, floods, water heights, in-situ

Acknowledgments

Beaucoup de doctorants disent qu'il est difficile de terminer une thèse et d'être satisfait de son travail. C'est une réalité que j'ai également vécue, mais pour moi la plus grande difficulté est de laisser ces trois années riches et épanouissantes derrière moi. Les relations sociales créées à Toulouse, dans le cadre du travail et en dehors, ont rythmé le cours de ma thèse. Je remercie immensément toutes les personnes que j'ai rencontrées, avec quelques mentions particulières à venir.

Dans la sphère du travail, j'ai eu la chance d'avoir de super co-directeurs de thèse, un tandem aux nombreuses qualités scientifiques et humaines. A chaque appel avec Frédéric, j'ai toujours reçu des encouragements et des éloges, surtout lorsque je doutais le plus. C'est quelque chose dont je me souviendrai toujours, et cette approche pédagogue et positive m'a largement donné envie de m'investir dans cette thèse. Quant à José, l'importance de sa présence à quelques dizaines de mètres de mon bureau a été énorme, pour accueillir toutes mes questions et mes incertitudes au cours de ces trois années. Je vous ai toujours senti à l'écoute, et prêts à me relancer sur la bonne piste grâce à quelques conseils et éléments scientifiques qui me manquaient. J'ai apprécié cette relation de travail autant que j'apprécie toute relation humaine en dehors, car j'abordais chaque discussion avec vous plein de bonne humeur.

Un grand merci également à tous les collègues du labo, mes co-auteurs, et à toutes les personnes avec qui j'ai échangé. Je pense en particulier à Catherine qui m'a apporté un point de vue critique et bienveillant, et avec qui les discussions sont toujours passionnantes. Merci à Luc pour ses nombreux conseils et pour nos échanges et projets, merci à Carlos, à Jean-Pierre et Eric qui ont formé mon comité de thèse. Un grand merci également à Adriano, Marie-Paule et Manuel qui ont accepté d'évaluer mon travail, ainsi qu'à Nicolas et Mehrez qui ont composé mon jury de thèse avec nombre des personnes précitées.

Un bon environnement de travail est nécessaire, mais ne permet pas forcément de garder le moral chaque jour. Pour cela, j'ai pu compter sur le soutien et l'amitié de personnes exceptionnelles. Les amis de longue date, la famille proche, mes parents, grands-parents, frères et sœurs, oncles et tantes, cousins et cousines, tous représentent mon équilibre de sorte que plutôt qu'une longue énumération, je préfère dire un grand MERCI à tout le monde. J'ai beaucoup de reconnaissance pour Christine et Rémi, l'ancrage toulousain, accueillants et généreux, et d'une compagnie toujours agréable. J'espère ne pas avoir été une présence trop pesante, surtout au cours des derniers mois de thèse ! Après quatre ans ici j'ai toujours un plaisir intact à passer chez vous, et ça n'a rien à voir avec mon courrier. Je pense beaucoup à l'anonyme Laurane avec qui j'ai presque commencé cette thèse, grâce à qui j'ai redécouvert un peu Paris et beaucoup les Alpes du Nord. Ton soutien et tes encouragements ont été précieux pour continuer mon travail, et je ne suis jamais aussi efficace que lorsque j'ai la tête pleine de rêves.

Je pense fort aussi à Gabriela et Lisa, mes deux co-bureau pour l'éternité, partenaires de

rigolade, de craquage et de vie quotidienne pendant trois belles années. Amies surtout, pour partager des bons repas et des apéros, des week-ends en montagne et des sortie toulousaines. Un immense merci à Romain, grâce à qui je me suis senti bien dans ce laboratoire que je n'ai jamais voulu quitter, sauf lorsque lui en était déjà parti. Toutes ces séances d'escalade et ces jeux de société trop faciles, ces bières artisanales et ces nuits en tente, cette proximité et cette amitié naturelles. Sans compter quelques sorties ski magnifiques ! Merci à Adé à qui je me suis souvent identifié, qui a été un soutien de poids et une amie proche, quelqu'un sans qui les journées au LEGOS n'auraient pas eu la même saveur, et sans qui la vie toulousaine aurait été plus fade. Hâte de se retrouver en conférence quand tu seras au stand des exposants, du côté du café ! Merci à Manon et Flo (dur de vous citer l'un sans l'autre), souriants et drôles, accueillants et prévenants, souvent occupés et toujours partants ! Que de bons moments vécus avec vous en montagne et pendant nos soirées jeux, même si la fermeture du Santal a été un gros coup dur.

Merci aux « anciens » qui ont su m'accueillir lors de mon arrivée et m'accompagner ces années, en particulier Alice entre Arnaud B. et François Verdier, Lise et sa visite toulousaine, Simon mon partenaire de ténor, et Marion avec ses histoires rocambolesques, son empathie et ses conseils en matière de teinture de cheveux. Merci à la team latin LEGOS pour les bouillons de culture et les bons repas, et toute cette joie ! Je suis particulièrement heureux d'avoir appris à parler andalous avec Julia, et d'avoir eu ces moments d'échange avec vous. A ceux qui sont arrivés ensuite comme Antonin et Morgane et qui ont apporté une bourrasque de fraîcheur, des sourires chaque midi et certains soir. Je dois aussi une fière chandelle au collègue tennisman et déménageur Arne, qui prend chaque expérience que la vie lui propose avec le sourire. Je pense également à Alexandra qui a toujours su me motiver à aller grimper, même lorsqu'elle n'en avait pas envie elle-même. Merci enfin à Juliette, Adrien, Elisa et Simon, Amélie, Gaëlle, Camille et tous les autres plus récemment arrivés, avec qui j'ai eu le temps de partager de chouettes moments. Je n'ai pas parlé de tout le monde mais les doctorants du LEGOS, dans leur collectivité, ont été une petite famille pour moi au cours de ces années.

Je souhaite également remercier chaleureusement mon ange gardien, Brigitte, ainsi que l'ensemble du GESSEC pour leur travail quotidien et leurs attentions régulières. De Martine et Nadine, parties trop tôt pour moi, à Agathe, Catherine, Bertrand et Nicolas, des présences rassurantes et aidantes. La direction du LEGOS aussi et notamment Frédéric, pour les échanges enrichissants et l'atmosphère créée au laboratoire. Le LEGOS et l'OMP ont été pour moi presque un foyer, surtout lors d'une fin de thèse plus agréable que prévue.

Finalement, j'espère que cette parenthèse de vie et de travail ne se refermera pas, mais pourra au contraire déboucher sur des relations longues et durables !

Contents

List of Figures

List of Tables

Acronyms

AGB	*Above Ground Biomass*
ALOS	*Advanced Land Observing Satellite*
CDMA	*Code Division Multiple Access*
cGNSS-R	*Conventional GNSS-R*
CNES	*Centre National d'Etudes Spatiales*
CYGNSS	*CYclone GNSS*
DDM	*Delay Doppler Map*
DEM	*Digital Elevation Model*
DGNSS	*Differential GNSS*
ESA	*European Space Agency*
FDMA	*Frequency Division Multiple Access*
FFZ	*First Fresnel Zone*
GIEMS	*Global Inundation Extent from Multiple Satellites*
GLONASS	*Globalnaya Navigatsionnaya Sputnikovaya Sistema*
GNSS	*Global Navigation Satellite System*
GNSS-R	*Global Navigation Satellite System Reflectometry*
GPS	*Global Positioning System*
GRACE	*Gravity Recovery and Climate Experiment*
iGNSS-R	*interferometric Global Navigation Satellite System Reflectometry*
InSAR	*Synthetic Aperture Radar Interferometry*
IPCC	*Intergovernmental Panel on Climate Change*
IPT	*Interference Pattern Technique*
JERS	*Japanese Earth Resources Satellite*
LEO	*Low Earth Orbit*
LHCP	*Left Hand Circular Polarization*

LMB *Lower Mekong Basin*

LSP *Lomb-Scargle Periodogram*

MODIS *Moderate Resolution Imaging Spectro-radiometer*

NASA *National Aeronautics and Space Administration*

NDVI *Normalized Difference Vegetation Index*

NISAR *NASA-ISRO Synthetic Aperture Radar*

PALSAR *Phased-Array type L-band Synthetic Aperture Radar*

PPP *Precise Point Positioning*

RHCP *Right Hand Circular Polarization*

RMSE *Root Mean Square Error*

SAR *Synthetic Aperture Radar*

SM *Soil Moisture*

SMAP *Soil Moisture Active Passive*

SMOS *Soil Moisture and Ocean Salinity*

SNR *Signal-to-Noise Ratio*

SSH *Sea Surface Height*

SSM/I *Special Sensor Microwave/Imager*

SSMIS *Special Sensor Microwave Imager / Sounder*

SWAMPS *Surface WAter Microwave Product Series*

SWE *Surface Water Extent in this manuscript (it is also commonly used to mean Snow Water Equivalent in other works)*

SWH *Significant Wave Height*

SWOT *Surface Water and Ocean Topography*

SWS *Surface Water Storage*

TDS-1 *TechDemoSat-1*

TWS *Terrestrial Water Storage*

WF *Water Fraction*

Introduction générale

La ressource en eau est essentielle à la vie sur Terre ainsi qu'au développement des écosystèmes naturels et des activités humaines. Elle d'une importance capitale pour la consommation directe des ménages, les usages industriels et l'agriculture (Abbott et al., 2019). Les stocks d'eaux de surface et les aquifères conditionnent la croissance de la végétation et la distribution des espèces de faune et de flore sur les surfaces terrestres (Junk et al., 2006). Surtout, l'influence du cycle de l'eau sur le climat est énorme (Chahine, 1992). Les intéractions constantes entre les grands cycles énergétique, hydrologique et biogéochimiques sont essentielles pour mieux comprendre le changement climatique. En particulier, les zones humides jouent un rôle essentiel dans les émissions mondiales de gaz à effet de serre

La végétation et les sols captent du dioxyde de carbone (CO_2) de l'atmosphère, notamment via la photosynthèse, ce que l'on appelle le puits de carbone continental. Cette pompe à carbone a jusqu'à présent suivi l'augmentation des émissions et concentrations de CO_2, permettant de réguler en partie leur hausse exponentielle (Friedlingstein et al., 2020; Walker et al., 2021). En revanche, son efficacité pourrait être réduite dans le futur sous l'effet d'un stress hydrique plus important (Humphrey et al., 2018). De plus, les zones humides et inondées sont la première source d'émissions de méthane (CH_4), qui est un gaz à effet de serre beaucoup plus puissant que le CO_2 (Saunois et al., 2020). L'assèchement des tourbières, les feux de forêt, la mortalité plus importante de la végétation et le dégel du pergélisol sont autant de conséquences potentielles du réchauffement climatique, qui augmenteraient les émissions naturelles de CH_4 et de CO_2. Ces mécanismes cont couramment appelés des *boucles de rétroaction positives*, et diminuent de fait le budget carbone de l'humanité fixé par exemple par les Accords de Paris en 2015 (Hoegh-Guldberg et al., 2018).

Les projections climatiques indiquent une augmentation future des précipitations, et surtout des extrêmes, à l'échelle mondiale (Douville et al., 2021). Cependant, de nombreuses régions devraient au contraire subir une diminution des précipitations qui, couplées à la hausse des températures, diminuera les ressources en eau disponibles (Polade et al., 2015; Douville et al., 2021). Ces régions incluent le bassin méditérranéen, l'Amérique Centrale, plus globalement la plupart des tropiques, et même le bassin de l'Amazone (Almazroui et al., 2020a,b, 2021). Elles rejoignent des régions dont l'accès à l'eau se voit menacé par la fonte des glaciers, comme les régions andines de Bolivie et du Pérou. Une meilleure connaissance du cycle de l'eau, de ses composantes régionales et de ses interactions avec les cycles biogéochimiques est donc essentielle pour s'adapter face au risque d'aridification.

Le changement climatique devrait également causer une augmentation de la fréquence et de l'intensité des crues extrêmes (Hirabayashi et al., 2013, 2021). Ces évènements font des dégats humains et matériels très importants, surtout dans les pays du Sud dont le niveau de vie est plus faible, et qui sont plus exposés aux aléas climatiques. Ces catastrophes deviennent également de plus en plus fréquentes dans les pays développés. Une adaptation pour créer des sociétés plus résilientes passe notamment par la connaissance plus approfondie des risques

climatiques, et par une meilleure prédiction et surveillance des évènements extrêmes (Kreibich et al., 2022). Dans le domaine de l'hydrologie, cela passe par un meilleur suivi des variations du cycle de l'eau, à des échelles spatiales et temporelles de plus en plus fines.

L'étude de l'étendue des inondations, des niveaux et volumes d'eau, des débits des rivières et autres variables hydrologiques s'est longtemps basée sur des réseaux d'observations in-situ. Cependant de nombreux désavantages sont à noter, parmi lesquels un très mauvais échantillonnage spatial, particulièrement hors des zones développées. De plus, le nombre de stations de mesure actives a fortement diminué depuis plusieurs décennies (Shiklomanov et al., 2002). La surveillance du cycle de l'eau repose donc essentiellement sur l'utilisation de la télédétection. Les images et mesures prises par les satellites d'observation de la Terre permettent de détecter et de quantifier la présence d'eau. On peut généralement opposer deux types de signaux utilisés par ces satellites, selon leur fréquence et donc leur utilité pour étudier les surfaces terrestres.

Tout d'abord, il y a les signaux dans les domaines du visible et de l'infrarouge (proche et moyen). Plusieurs satellites permettent d'obtenir des images optiques avec une résolution allant de quelques mètres à quelques centaines de mètres, notamment Landsat et MODIS (*Moderate Resolution Imaging Spectroradiometer*). Leur avantage principal, outre la résolution spatiale assez élevée, est la disponibilité de longues séries temporelles pour étudier à la fois les tendances et la variabilité du cycle de l'eau (Woodcock et al., 2008; Wulder et al., 2016). Les satellites Landsat opèrent ainsi depuis 1975, AVHRR (*Advanced Very High Resolution Radiometer*) depuis 1981, MODIS depuis 2000, et Sentinel-2 depuis 2015. Leur inconvénient est la longueur d'onde du signal étudié, de l'ordre de quelques micromètres maximum, qui est bloqué par le couvert nuageux ainsi que la végétation, et ne fonctionne pas de nuit. Plusieurs études récentes ont fourni des détections des eaux de surface à l'échelle globale en se basant sur MODIS (Carroll et al., 2009), ou Landsat (Verpoorter et al., 2014; Yamazaki et al., 2015; Feng et al., 2016). Le produit de référence fournit des estimations mensuelles depuis 1984, en se basant sur les archives Landsat (Pekel et al., 2016). Récemment, la nouvelle mission européenne Sentinel-2 a montré son potentiel pour la cartographie des eaux de surface (Martinis et al., 2022). Cependant, la limitation intrinsèque de ces produits reste l'obstruction des signaux par la végétation qui ne leur permet pas de détecter de nombreuses zone humides, en particulier dans les grands bassins équatoriaux du Congo et de l'Amazone.

Pour cela, les signaux micro-ondes (ou ondes radar) sont parfaitement adaptés. Ils permettent de traverser le couvert nuageux, fonctionnent de nuit, et pénètrent en partie le couvert végétal. Plusieurs types de données sont disponibles : les mesures passives, captant le signal émis par la surface de la Terre avec une résolution spatiale très basse (de 25 à 50 km), et les mesures actives, dont l'instrument émet un signal qui sera ensuite réfléchi par la surface de la Terre puis capté par le récepteur. Les systèmes actifs permettent d'avoir une résolution spatiale très élevée, de l'ordre de quelques mètres pour les Radars à Synthèse d'Ouverture (RSO, ou SAR pour l'acronyme anglais). Ils ont donc été très utilisés pour étudier les surfaces continentales. En hydrologie, les SAR permettent d'obtenir de l'information sur les inondations indépendamment des conditions climatiques et presque en temps réel (Pierdicca et al., 2013; Westerhoff et al., 2013; Martinis et al., 2015a; Twele et al., 2016; Chini et al., 2017). La

pénétration des couverts végétaux les plus denses est cependant mauvaise pour la plupart des capteurs, sauf ceux fonctionnant en bande L, assez rares. Ce manque de données ainsi que la variabilité de la rétrodiffusion SAR font que ces données ont été utilisées pour cartographier les inondations à l'échelle locale ou régionale, et très peu en global. Quelques études utilisent la complémentarité entre le SAR et les images optiques Martinis et al. (2022), mais restent limitées dans les zones de végétation dense. De plus, le SAR n'a jamais permis de produire de longues séries temporelles pour dépasser la variabilité interannuelle du cycle hydrologique.

Finalement, les radiomètres dans les micro-ondes représentent les systèmes les mieux à même de cartographier les variations spatiotemporelles des inondations. Ils ont permis d'estimer des produits globaux, comme le Global Inundation Extent from Multi-Satellite (GIEMS, Prigent et al. (2007, 2020)) fournissant des estimations mensuelles de fractions d'eau depuis 1992. Sa résolution spatiale est très basse (0.25°), car limitée par celle des capteurs passifs utilisés dans le produit. Il y a donc un vrai manque, à l'heure actuelle, de connaissance sur la dynamique des inondations dans les grands bassins tropicaux, notamment le Congo et l'Amazone. L'enjeu est de proposer des estimation d'étendue des eaux de surface avec une résolution spatiale plus élevée que GIEMS, permettant d'étudier les variations interannuelles du cycle hydrologique, et fournissant de l'information sur les forêts tropicales inondées.

Dans le cadre de cette thèse, mon approche a consisté à développer le potentiel pour l'hydrologie d'une nouvelle technique de télédétection, la Réflectométrie GNSS (GNSS-R). Le principe du GNSS-R repose sur l'utilisation des signaux émis par les satellites GNSS (*Global Navigation Satellite System*), qui servent au positionnement des utilisateurs sur Terre. Ces signaux sont également réfléchis par la surface de la Terre, et peuvent être acquis par un récepteur pour étudier les propriétés géophysiques de la surface réflechissante (Martin-Neira et al., 1993; Zavorotny et al., 2014). De plus en plus de satellites GNSS sont disponibles, grâce au développement des constellations chinoise BeiDou, européenne Galileo, et russe GLONASS, qui s'ajoutent à ceux de la constellation américaine initiale GPS (*Global Positioning System*). Le GNSS-R possède de multiples applications. Sur les océans, il est possible d'estimer la hauteur d'eau (Martin-Neira et al., 1993; Larson et al., 2013a; Roussel et al., 2015) et la hauteur significative des vagues, permettant de calculer la vitesse du vent (Foti et al., 2015; Clarizia and Ruf, 2016). Sur les continents, les applications majeures sont l'humidité des sols (Larson et al., 2008a; Rodriguez-Alvarez et al., 2009; Chew et al., 2016), la mesure des épaisseurs de neige (Larson et al., 2009b; Rodriguez-Alvarez et al., 2012), ou encore l'étude de la biomasse (Carreno-Luengo et al., 2020) et la hauteur de la végétation (Munoz-Martin et al., 2022). Le cas de l'hydrologie est traité à part ci-dessous.

L'une des caractéristiques intéressantes du GNSS-R pour l'hydrologie, c'est la possibilité de réaliser des acquisitions in-situ à l'aide d'une antenna géodésique simple (Kavak et al., 1998; Larson et al., 2008a), ou alors des applications à l'échelle régionale et globale avec des récepteurs embarqués respectivement sur un avion (Cardellach et al., 2011) ou un satellite (Ruf et al., 2016). A l'échelle locale, les signaux GNSS réfléchis peuvent être utilisés pour calculer les niveaux d'eau. Ceci a particulièrement été étudié en milieu côtier sur les cycles de marée, avec des techniques plus (Roussel et al., 2015) ou moins (Larson et al., 2013a) sophistiquées. Les applications sur les eaux continentales ont été beaucoup plus réduites, car

plus complexes. C'est l'objet de la première partie de ma thèse, avec un cas d'étude sur les marées asymmétriques et le mascaret dans la Garonne. Les résultats montrent la possibilité d'étudier des changements soudains et de grande amplitude du niveau d'eau par GNSS-R, et ont été publiés dans le journal *Remote Sensing* (voir le Chapitre 3).

A l'échelle globale, la visibilité du GNSS-R a bénéficié du lancement d'une mission spatiale par la NASA fin 2016. Cyclone GNSS (CYGNSS) représente une constellation de 8 micro-satellites dédiés, comme son nom l'indique, à l'étude des cyclones tropicaux (Ruf et al., 2016). Cependant, cette mission fournit une grande quantité de données en milieu continental, dans la bande inter-tropicale ($\pm 38°$ de latitude). La combinaison de 8 satellites permet également de réaliser un grand nombre de mesures. Il est donc logique que ces données aient été utilisées pour estimer l'humidité des sols (Chew et al., 2018; Al-Khaldi et al., 2019; Clarizia et al., 2019; Chew and Small, 2020a), la végétation (Carreno-Luengo et al., 2020), et bien sur pour chercher à détecter la présence d'eau.

Comme je l'ai expliqué précédemment, l'estimation de l'étendue des inondations dans les grands bassins tropicaux est actuellement très mal contrainte. Il se trouve que CYGNSS possède la capacité de détecter les eaux de surface (Chew et al., 2018; Wan et al., 2019; Gerlein-Safdi et al., 2021; Al-Khaldi et al., 2021a), grâce aux propriété de cohérence et à la réflexion quasi-spéculaire du signal sur les surfaces d'eau lisses (Voronovich and Zavorotny, 2018; Loria et al., 2020). La résolution spatiale de ces observations permet de détecter des objets d'eau de l'ordre de quelques centaines de mètres (Camps, 2019; Camps and Munoz-Martin, 2020), voire inférieurs à 100 m en utilisant les (rares) données complexes disponibles (Li et al., 2021, 2022). Cependant, la rugosité des sols et des surface d'eau, la végétation, ou encore l'humidité des sols contribuent à des variations du signal GNSS réfléchi. Les études régionales permettent de s'affranchir des différences entre les régimes climatiques et les paramètres géophysiques au niveau global, où les travaux sont bien moins nombreux.

Le coeur de mon travail de thèse représente le développement d'une chaîne de traitement des données CYGNSS, permettant la détection des surfaces inondées sur toute la couverture spatiale de la mission. Ce travail a fait l'objet d'une publication récente dans *Remote Sensing of Environment* (Chapitre 4). Il y est montré que malgré l'influence de tous les parametres géophysiques, et notamment l'atténuation des signaux réfléchis par le couvert végétal, CYGNSS peut détecter des inondations sur tous les types de climats rencontrés dans la zone inter-tropicale. Le compromis trouvé entre la résolution spatiale ($0.1°$) et la résolution temporelle (7 jours) de nos estimations est très intéressant en comparaison avec les produits existants tels GIEMS. Ce travail se poursuit toujours, avec pour objectif de fournir une estimation des fractions d'eau dans chaque pixel de $0.1°$, et donc une estimation de l'étendue et des dynamiques saisonnières des inondations. Cela donnera bientôt lieu a une troisième publication, dont les principaux résultats sont présentés dans le Chapitre 5. A terme, ces estimations pourront être étendues sur toute la période d'acquisitions de CYGNSS, courant depuis 2017. Elles pourraient avoir un intérêt majeur pour la communauté scientifique, et notamment permettre:

- La comparaison avec GIEMS, qui ne dispose actuellement d'aucune base de validation globale, afin de montrer les améliorations à fournir sur les deux produits.

- La mise à disposition d'estimation plus précises de fractions d'eau depuis 2017, pour étudier les mécanismes hydrologiques des bassins tropicaux, et notamment leur saisonnalité et leur variabilité interannuelle (quand une période de temps suffisamment longue sera disponible).
- La constitution d'un jeu de données complémentaire des produits optiques et de GIEMS, utilisable comme base de validation pour les futures missions spatiales dédiées à l'hydrologie.

Le structure de ce manuscrit est organisée de la manière suivante. Le Chapitre 1 présente un état de l'art sur le cycle de l'eau et les cycles biogéochimiques, pour fournir au lecteur les clés de la compréhension de ce travail. Sont ensuite présentées les différentes techniques de télédétection permettant d'étudier les eaux de surface. Le Chapitre 2 introduit les notions de base sur le GNSS, la réflectométrie GNSS, ainsi que les techniques de mesure, les instruments et les applications de cette dernière. Le Chapitre 3 présente les résultats obtenus sur la détermination des niveaux d'eau lors de marées asymmétriques dans la Garonne, au moyen d'une station GNSS géodésique installée sur la berge, qui ont été publiés dans *Remote Sensing*. Le Chapitre 4 présente l'étude des observations CYGNSS appliquées à l'hydrologie, pour détecter les zones affectées par des inondations, qui a fait l'objet d'une publication dans *Remote Sensing of Environment*. Le Chapitre 5 prolonge ces résultats en montrant la possibilité d'estimer, sur une base spatio-temporelle de 0,1° et 7 jours, des fractions d'eau dans chaque pixel grâce à CYGNSS. Cela fera l'objet d'une troisième publication prochainement. Enfin, la conclusion et les perspectives de mon travail de thèse clôtureront ce manuscrit.

General introduction

Freshwater is essential to life on Earth, to the development of natural ecosystems and human activities. It is used for household consumption, industry and agriculture (Abbott et al., 2019). Surface water and groundwater determine the vegetation growth and the distribution of fauna and flora species on land (Junk et al., 2006). Furthermore, the water cycle has a huge influence on climate (Chahine, 1992), and constantly interacts with the energy and biogeochemical cycles. In particular, wetlands play a key role in global greenhouse gas emissions.

Vegetation and soils capture carbon dioxide (CO_2) from the atmosphere, particularly through photosynthesis. This mechanism is called the land carbon sink and partially mitigates the exponential growth of CO_2 emissions and concentrations. Its effectiveness could be reduced in the future due to water stress (Humphrey et al., 2018). In addition, wetlands and flooded areas are the primary source of methane (CH_4) emissions, which is a much more powerful greenhouse gas than CO_2 (Saunois et al., 2020). Drying peatland, fires, increased vegetation mortality and thawing permafrost are all potential consequences of climate change, that would increase in return natural CH_4 and CO_2 emissions. These mechanisms are commonly referred to as *positive feedbacks*. Their combined effect decreases the remaining carbon budget set by the Paris Agreement in 2015 for example.

Climate projections indicate a future increase in global precipitation, and especially in extremes (Douville et al., 2021). However, many regions are expected to experience drying due to decreasing regional precipitation coupled with increasing temperatures (Polade et al., 2015; Douville et al., 2021). These regions include the Mediterranean Basin, Central America, most of the subtropics, and even the Amazon Basin (Almazroui et al., 2020a,b, 2021). In other regions such as the Altiplano, the access to freshwater is threatened by melting glaciers. A better monitoring of the water cycle, its regional components and its interactions with biogeochemical cycles is therefore essential to mitigate the expected aridification.

Climate change is also expected to increase the frequency and intensity of extreme floods (Hirabayashi et al., 2013, 2021). They cause very important human and material damages, especially in the Southern countries which are more exposed to climatic hazards. These disasters now become more and more frequent in Northern developed countries. Adaptation to create more resilient societies requires a better prediction and monitoring of extreme events (Kreibich et al., 2022). The variations of the water cycle should be monitored accurately at fine spatial and temporal scales.

The study of inundation extent, water levels and volumes, river flows and other hydrological variables has long been based on ground station networks. There are many disadvantages including a very poor spatial sampling, especially in remote regions. In addition, the number of active stations has strongly decreased for decades (Shiklomanov et al., 2002). The monitoring of the water cycle therefore relies on the use of remote sensing. The images and measurements taken by Earth observation satellites make it possible to detect and quantify the presence of water. These satellites use two major types of signals to monitor land

and ocean surfaces, which are further distinguished depending on their frequency and their usefulness for land remote sensing.

There are first signals in the visible and infrared domains. Several satellites produce optical images of the Earth surface with a resolution ranging from a few meters to a few hundred meters, including Landsat and MODIS (*Moderate Resolution Imaging Spectroradiometer*). In addition to the high spatial resolution, their main advantage is the availability of long time series to study both the climatic trends and the variability of the water cycle (Woodcock et al., 2008; Wulder et al., 2016). Landsat satellites have been operating since 1975, AVHRR (*Advanced Very High Resolution Radiometer*) since 1981, MODIS since 2000, and Sentinel-2 since 2015. Their drawback is the wavelength of the signal of a few micrometers at the most. They do not penetrate neither the cloud cover nor the vegetation, and do not operate by night. Several studies estimated global surface water distribution and extent based on MODIS (Carroll et al., 2009) or Landsat (Verpoorter et al., 2014; Yamazaki et al., 2015; Feng et al., 2016). The reference product provides monthly estimates at 30 m spatial resolution since 1984, based on the Landsat archive (Pekel et al., 2016). However, these estimations are inherently limited over tropical wetlands as the vegetation obscures floods for visible and infrared signals.

For this, microwave signals are well suited. They penetrate cloud cover, operate by night, and partially penetrate the vegetation. The passive and active measurements are distinguished. Passive observations captures the signal emitted by the Earth's surface with a very low spatial resolution (25 to 50 km), while active sensors emit a signal that is further reflected by the Earth's surface and captured by the receiver. Active systems have a very high spatial resolution, down to one meter for some Synthetic Aperture Radars (SAR). For hydrology, SAR can detect floods under all weather conditions and in near real time (Pierdicca et al., 2013; Westerhoff et al., 2013; Martinis et al., 2015a; Twele et al., 2016; Chini et al., 2017). However, the penetration of the densest canopies is poor for most sensors, except for the few L-band SAR. This lack of data as well as the variability of the backscatter made SAR images mostly used to map inundations at the local or regional scale. Few studies used the complementary SAR and optical images (Martinis et al., 2022), and they remain limited in densely forested areas. Furthermore, SAR has never been able to produce long time series of surface water extent to overcome the interannual variability of the hydrological cycle.

Finally, microwave radiometers represent the best systems for mapping spatio-temporal flood dynamics. They have been used in global estimations such as the Global Inundation Extent from Multi-Satellite (GIEMS, Prigent et al. (2007, 2020)) providing monthly water fraction estimates since 1992. Its spatial resolution is very low (0.25°), as it is limited by the passive sensors used in the product. The dynamics of floods are thus poorly monitored for now in large tropical basins, including the Congo and the Amazon. The challenge is to propose surface water extent estimates with a higher spatial resolution than GIEMS, to study interannual variations of the hydrological cycle in flooded tropical forests.

During my PhD, I studied the potential of GNSS-Reflectometry (GNSS-R) for hydrology. GNSS-R is a new remote sensing technique based on the use of GNSS (*Global Navigation Satellite System*) signals, primarily dedicated to positioning on Earth. These signals are

also reflected by the Earth's surface, and can then be acquired by a receiver to study the geophysical properties of the reflecting surface (Martin-Neira et al., 1993; Zavorotny et al., 2014). An increasing number of GNSS satellites are available, due to the development of the Chinese BeiDou, European Galileo, and Russian GLONASS constellations, in addition to the original US GPS (*Global Positioning System*). GNSS-R has multiple applications. Over the ocean, it is possible to estimate the water level (Martin-Neira et al., 1993; Larson et al., 2013a; Roussel et al., 2015) and the significant wave height, related to wind speed (Foti et al., 2015; Clarizia and Ruf, 2016). Over land, the major applications are soil moisture (Larson et al., 2008a; Rodriguez-Alvarez et al., 2009; Chew et al., 2016), snow depth (Larson et al., 2009b; Rodriguez-Alvarez et al., 2012), biomass (Carreno-Luengo et al., 2020) and vegetation height (Munoz-Martin et al., 2022). Other hydrological applications are presented below.

GNSS-R is possible at local scale with in-situ acquisitions using a geodetic antenna (Larson et al., 2008b), or support regional and global studies with receivers onboard aircrafts (Cardellach et al., 2011) or satellites (Ruf et al., 2016), respectively. Both spatial scales are particularly interesting for hydrology. At the local scale, the reflected GNSS signals can be used to calculate water levels. It was particularly studied in coastal environments, with more (Roussel et al., 2015) or less (Larson et al., 2013a) accurate techniques. Applications over inland water include lake level (Holden and Larson, 2021), but the monitoring of rivers was very limited up to now due to the more complex data acquisition and processing techniques. During the first part of my PhD, I studied asymmetric tides and tidal bores in the Garonne River using GNSS-R in-situ measurements. The results demonstrated the possibility to retrieve sudden and large amplitude changes in the water levels, and were published in *Remote Sensing* (see Chapter 3).

At global scale, GNSS-R benefited from the launch of a space mission by NASA in December 2016. The Cyclone GNSS (CYGNSS) is a constellation of 8 micro-satellites dedicated to study tropical cyclones (Ruf et al., 2016). This mission still provides a large amount of data over land and in the inter-tropical band ($\pm 38°$ latitude), due to the combination of 8 satellites. These observations were logically used to estimate soil moisture (Chew et al., 2018; Al-Khaldi et al., 2019; Clarizia et al., 2019; Chew and Small, 2020a), vegetation (Carreno-Luengo et al., 2020), and of course to detect the presence of water. This specific application is further introduced in the next paragraph as it constitutes the main state-of-the-art of my thesis work.

As it was already stated, the estimation of inundation extent in large tropical basins is currently poorly constrained. CYGNSS observations can detect surface water (Chew et al., 2018; Wan et al., 2019; Gerlein-Safdi et al., 2021; Al-Khaldi et al., 2021a), due to the quasi-specular reflection of GNSS signals on smooth water surfaces (Voronovich and Zavorotny, 2018; Loria et al., 2020). The spatial resolution of these observations permits to detect water bodies of a few hundred meters (Camps, 2019; Camps and Munoz-Martin, 2020), and below 100 m using the sparse complex data available (Li et al., 2021, 2022). However, the coherence and the power of reflected GNSS signals show variations depending on geophysical parameters of the reflecting surface (soil and water surface roughness, vegetation, and soil moisture). For this reason, few global studies exist on surface water detection using CYGNSS.

The core of my PhD work was the development of a CYGNSS data processing chain, to detect flooded areas in the entire spatial coverage of the mission. This work has been recently published in *Remote Sensing of Environment* (Chapter 4). It demonstrates that CYGNSS can detect floods over all the inter-tropical band despite the influence of geophysical parameters, and in particular the attenuation of the reflected signals by the vegetation. The compromise between the spatial (0.1°) and the temporal (7 days) resolutions of our estimations is very interesting compared to existing products such as GIEMS. This work is still in progress, with the objective of estimating the fraction of water contained in each 0.1° pixel with a weekly time sampling. This will soon lead to a third publication, with the main results presented in Chapter 5. Eventually, these estimates will be extended over the entire CYGNSS acquisition period, since 2017. They could be of great interest for the hydrological community for several reasons:

- Intercomparison with GIEMS in order to correct its biases and finally improve both products.
- Provision of more accurate water fraction estimations since 2017 to study tropical wetland hydrology, in particular its spatial distribution, seasonality and interannual variability (with a sufficiently long time series).
- Constitution of a dataset complementary to optical products and GIEMS, to compare with the future space missions dedicated to hydrology (*i.e.* SWOT, NISAR and HydroGNSS).

This manuscript is organized as follows. Chapter 1 presents a state-of-the-art on the water biogeochemical cycles, to provide the keys to understand this work. The different remote sensing techniques used to study surface water are then presented and discussed. Chapter 2 introduces the concepts of GNSS, GNSS reflectometry, as well as the measurement techniques, instruments and applications of the latter. Chapter 3 presents the results obtained on the determination of water levels during asymmetric tides in the Garonne River using a geodetic GNSS station, which were published in *Remote Sensing*. Chapter 4 presents the analysis of CYGNSS observations for detecting inundated areas, which was published in *Remote Sensing of Environment*. Chapter 5 extends these results by presenting a 0.1°, 7-day spatio-temporal estimation of the fraction of water based on CYGNSS. This will feed a third publication soon. Finally, I will present the general conclusions and the perspectives of my work.

Chapter 1

Monitoring of surface water dynamics

Contents

1.1 The water cycle

The water or hydrological cycle is essential to life and is also closely linked to the climate system and its changes (Chahine, 1992). It encompasses all the water stored in the oceans, atmosphere, land and ice sheets, as well as their interactions. This section introduces the basis of land hydrology, to assess its requirements and pave the way to further analysis. It describes in particular the interactions between freshwater on land, and water vapour in the atmosphere. For the purpose of this manuscript, the ocean and cryosphere components of the water cycle are left aside and only their principal interactions with land hydrology are mentioned.

1.1.1 Components of the water cycle

Freshwater is one of the most essential natural resources. It represents less than 2% of all the water on Earth, and 96% of freshwater reserves are stored in ice sheets, glaciers and snow. The remaining 4% represent a total volume of 835,000 km^3, which are accessible for essential needs of human societies and ecosystems functioning (Abbott et al., 2019). The surface reservoirs (lakes, rivers and wetlands) store only a quarter of this volume, the rest being contained in groundwater. Freshwater supports human development, water consumption, cropland irrigation, industrial use, among other anthropogenic processes. The use of freshwater resources by humans comports direct surface water withdrawals (called blue water), soil moisture use by livestock, agriculture and forestry (green water), and water polluted by industrial processes (grey water). Together, these three categories represent an anthropogenic water consumption of ∼24,000 $km^3\ yr^{-1}$, being half of yearly river discharge or double of yearly groundwater recharge (Abbott et al., 2019).

Freshwater is also essential to the functioning of many natural ecosystems, in particular wetlands, which are impacted in several ways by the deficit in moisture or standing water: destruction of natural habitats and loss of biodiversity (Junk et al., 2006), feedbacks to regional and global biogeochemical cycles (Jung et al., 2017; Humphrey et al., 2018), and emergence and propagation of infectious diseases (Kouadio et al., 2012; Suk et al., 2020), among others. An overview of the water cycle is presented in Figure 1.1, with both the major water pools on Earth (a) and the major fluxes between all components of the water cycle (b).

The water evaporated over the oceanic and land surfaces supply atmospheric moisture, which is redistributed as precipitation on the Earth surface. In Figure 1.1, a strong atmospheric moisture transport from the oceans to land is highlighted, contributing in a large fraction (∼ 40%) of land precipitation. This moisture redistribution compensates the water fluxes from land to ocean, primarily supplied by river discharge. Another important source of moisture for land precipitation is terrestrial evapotranspiration (ET). It is the sum of three components: the evaporation from soil and open water (E), the transpiration of the vegetation (T), and the vaporization of rainfall intercepted by the vegetation (I):

$$ET = E + T + I \tag{1.1}$$

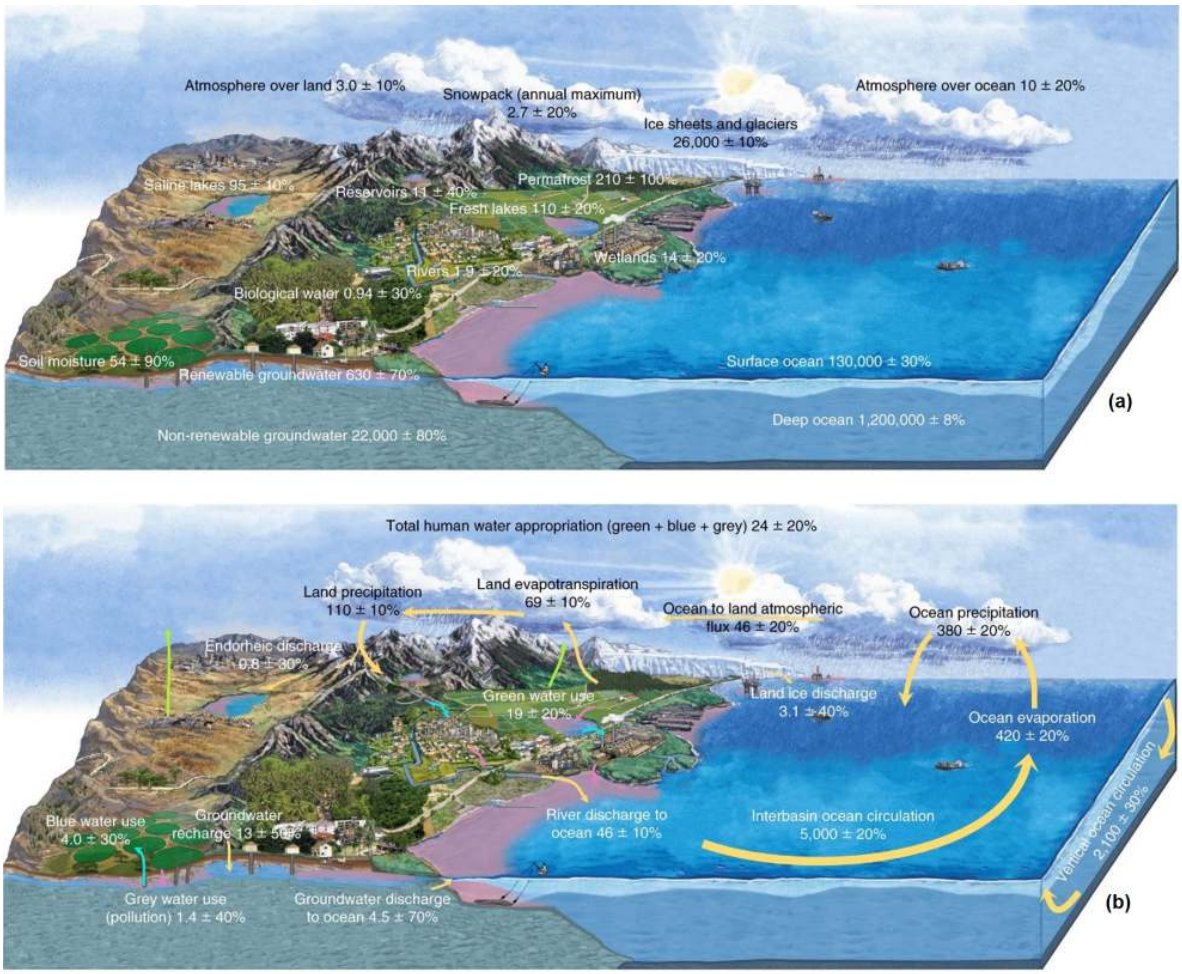

Figure 1.1: Overview of the global hydrological cycle, from Abbott et al. (2019). (a) major water pools on Earth (expressed in 10^3 km^3 yr^-1), and (b) major water fluxes between the different components of the water cycle. The uncertainties are given in percentage of the water pool or flux considered.

Evapotranspiration therefore represents the total transfer of moisture from land to the atmosphere, through the interfaces of soils and vegetation. Precipitation is the response to atmospheric moisture, and supply soil moisture, groundwater recharge and river streamflow. They drive the seasonality and amount of permanent or temporary surface water storage in lakes and floodplains. The variations in precipitation therefore induce a large range of responses from the water cycle over land, with water scarcity as an extreme response to severe droughts or water deficits (Schewe et al., 2014).

Precipitation at the global and regional scales are constrained by the Earth energy balance, and moisture fluxes in the atmosphere and at the surface. At global scale, thermodynamic effects drive the mean annual precipitation, through the impacts of radiative forcing by greenhouse gases (GHGs) and anthropogenic aerosols on the Earth energy balance. This is described by the Clausius-Clapeyron relationship applied to the saturation water vapor

pressure (e_s), which is written as follow:

$$\frac{\mathrm{d}\, e_s}{\mathrm{d}\, T} = \frac{L_v\, e_s}{R\, T^2} \tag{1.2}$$

where T is the temperature in kelvins (K), L_v is the specific latent heat of evaporation of water in J·mol^{-1}, and R is the gas constant in J·mol^{-1}·K^{-1}. According to this relation, the atmospheric moisture can increase by ∼7% per 1°C of warming. The radiative forcing and the resulting climate warming has several impacts according to the Clausius-Clapeyron relation. First, there is an increase of atmospheric evaporative demand, leading to water stress for soils and vegetation in case of limited soil moisture. Then, mean annual precipitation is likely to increase at the rate of 1-3% per 1°C according to the IPCC AR6 report (Douville et al., 2021). This include fast atmospheric adjustments to radiative forcing by GHGs and aerosols, and a slower but more robust response to surface temperature warming and vegetation only estimated to 2-3% $^{\circ}C^{-1}$ (Andrews et al., 2010; Cao et al., 2012). As the mean increase of precipitation is largely lower than 7% $^{\circ}C^{-1}$, the excess of atmospheric water vapour is likely to increase the intensity of extreme rainfall events close to the rate predicted by the Clausius-Clapeyron relation (Douville et al., 2021).

At the regional scale, the global increase in precipitation amounts due to thermodynamic effect is most of the times strengthened and sometimes tempered by atmospheric moisture transport. It was shown that at a spatial scale of more than 3000-4000 km, the precipitation trends follow the global mean, whereas for lower spatial scales, they digress from it (Dagan et al., 2019; Dagan and Stier, 2020). These regional patterns mostly depend on water supplied by dynamic moisture transport in the atmosphere (Dagan et al., 2019). The regional moisture transport is one of the drivers of the intensification of wet events under global warming. On the contrary, the atmospheric evaporative demand is stronger in a warmer climate due to the increase in vapour pressure deficit (VPD) (Scheff and Frierson, 2014; Vicente-Serrano et al., 2018), and is one of the drivers of land aridification along with the precipitation (Scheff and Frierson, 2015). The intensification of wet and dry events in a warming climate will likely exacerbate regional droughts and heavy rainfalls, therefore increasing both the contrasting risks of aridification and floods at regional scale (Hirabayashi et al., 2013, 2021). The next section deeper analyzes the observed changes in the water cycle in the Anthropocene, and the future expected trends.

1.1.2 Projected water cycle changes

The IPCC AR5 and AR6 reports, modeling results from the phases 5 and 6 of Coupled Model Intercomparison Project (CMIP5 and CMIP6, respectively), as well as numerous studies, show that an intensification of the water cycle is expected as response to climate change. The total precipitation (P), evaporation (E), runoff, and water availability (P-E) are expected to increase regardless of the Shared Socioeconomic Pathway (SSP) considered. The CMIP6 simulations predict an average increase of precipitation for 2081-2100 relative to 1995-2014, by 2.4% in the SSP1-1.9 which is the low-emission scenario, and 8.3% in the SSP5-8.5 which is the high-emission scenario (Douville et al., 2021). The projected increase in atmospheric moisture

is close to the rate predicted by the Clausius-Clapeyron relation (7% $°C^{-1}$ of warming).

Regionally, CMIP6 models predict an increase of precipitation in the tropics and high latitudes, and lower precipitation in the subtropical drylands (see Figure 1.2). According to global simulations and regional studies (Almazroui et al., 2020a,b, 2021; Douville et al., 2021), annual mean rainfall is likely to increase in Southeast and East Asia, boreal high latitudes, most of North America, Central and East Africa, and Southeast South America. The risk of floods in these areas is likely to increase according to CMIP5 and CMIP6 models (Hirabayashi et al., 2013, 2021). On the contrary, annual mean rainfall is expected to decrease in South, North and West Africa, Amazon and Central America, Southwest Australia and the Mediterranean region, with a high risk of aridification in semi-arid regions. In most of the drying areas, all hydrological variables will also exhibit a seasonal cycle of larger amplitude. Extreme precipitations are expected to strengthen worldwide at a rate close to 7% $°C^{-1}$ of warming discussed in Section 1.1.1. This is constant regardless of the emission scenarios (Pendergrass et al., 2015), as shown in the right panel of Figure 1.3. However, the frequency of rainfall is likely to experience heterogeneous regional changes, with more (respectively fewer) dry days in regions where the mean precipitation is expected to decrease (respectively

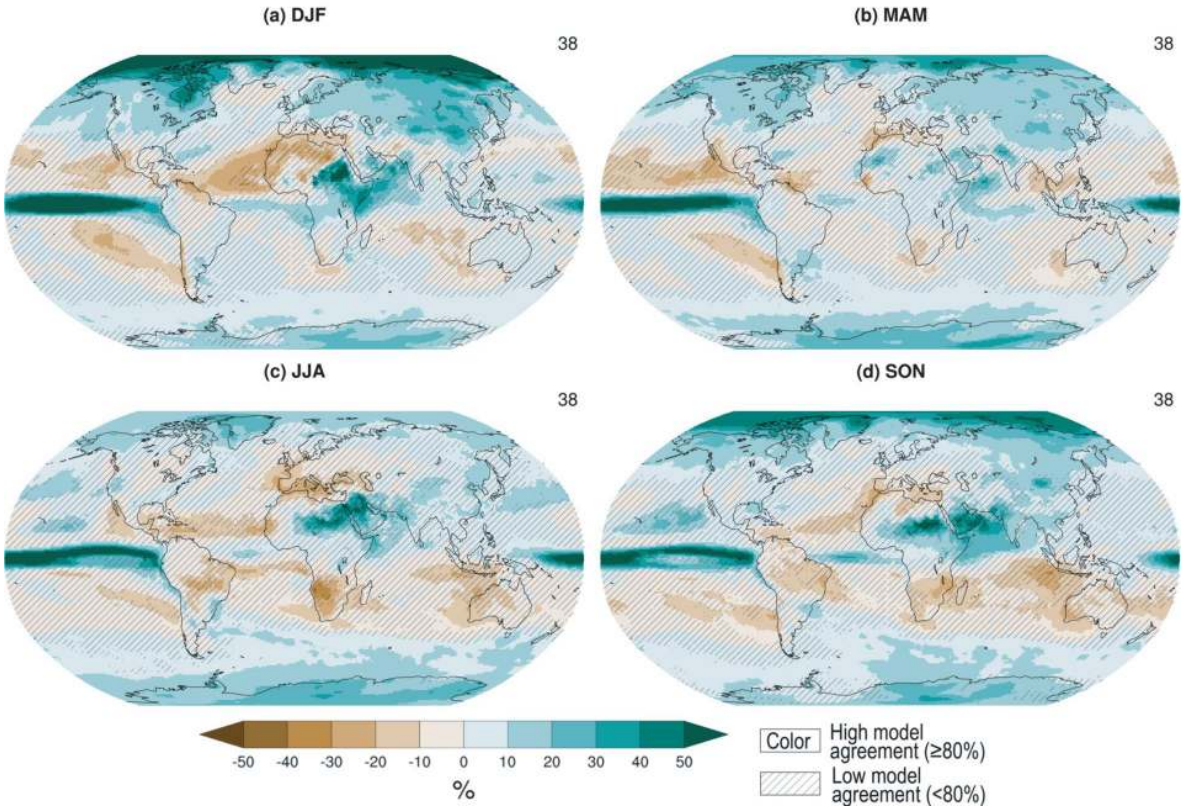

Figure 1.2: Projection of relative changes in mean seasonal precipitation, for 2081-2100 relative to 1995-2014, from Douville et al. (2021). The precipitation prediction is the average of the CMIP6 models predictions forced in the SSP2-4.5 scenario. The overlay used to assess uncertainties indicates whether ≥80% of the predictions agree on the sign of change.

increase, see left panel in Figure 1.3). A larger number of dry days is expected to dominate the negative trend and the interannual variability of annual precipitation totals over most of the subtropics shown in Figure 1.2 (Polade et al., 2015).

In response to the intensification of short-duration extreme rainfall, flash-floods are likely

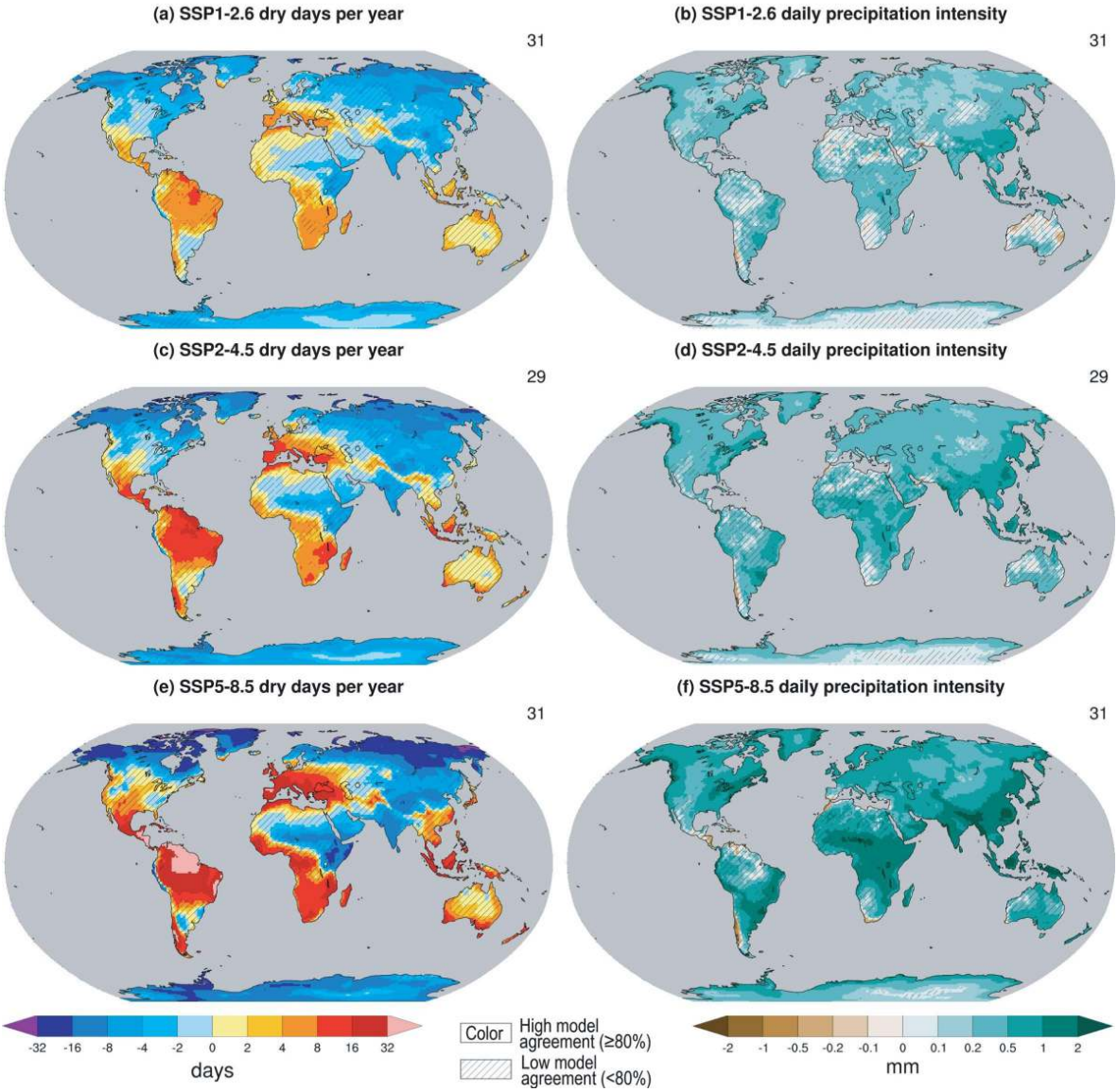

Figure 1.3: Projection of relative changes in daily precipitation statistics, for 2081-2100 relative to 1995-2014, from Douville et al. (2021). **(a,c,e)** Changes in the number of dry days per year, defined as days with less than 1 mm of rain. **(b,d,f)** Changes in daily precipitation intensity, defined as the average total rainfall on wet days. The projections are averaged from the predictions of CMIP6 models, forced with three SSP scenarios: **(a,b)** SSP1-2.6, **(c,d)** SSP2-4.5, and **(e,f)** SSP5-8.5. The overlay used to assess uncertainties indicates whether ≥80% of the predictions agree on the sign of change.

to become more frequent and severe (Hirabayashi et al., 2013; Chan et al., 2016a; Sandvik et al., 2018; Hirabayashi et al., 2021). Observations are required to validate these model predictions which are not well constrained at sub-hourly time scale (Chan et al., 2016a). The persistence of heavier rainfall during the wet season and the increase in total precipitation are expected to have contrasting impacts, from local to regional scales. More intense rainfall increase the proportion of water dedicated to runoff and surface storage (blue water), at the expense of a lower infiltration in soils and water use by the vegetation (green water), impacting the vegetation productivity and the agriculture (Eekhout et al., 2018). When the soils are saturated, larger runoff and flooding risks occur (Yin et al., 2018). The landslide risk is also expected to grow due to changing precipitation patterns (Gariano and Guzzetti, 2016). Changes are also expected in the groundwater recharge rate, as it highly relies on nonlinear relationships with the intensity of rainfall (Taylor et al., 2013a). Higher recharge rate were observed for extreme precipitation (Taylor et al., 2013a,b; Cuthbert et al., 2019), which are then essential to offset groundwater depletion in the context of a global groundwater crisis (Famiglietti, 2014). Indeed, the groundwater use exceeds their recharge in many regions of the world, mainly due to the extraction of non-renewable water for farming in productive dryland agricultural areas (Döll et al., 2014; Bierkens and Wada, 2019). Extreme precipitation could thus help in ensuring water security within the frame of climate change. Nevertheless, they will especially strengthen natural hazards worldwide and affect the efficiency of risk management policies (Kreibich et al., 2022).

The cryospheric component of the water cycle is also extremely sensitive to climate change. As warming is higher in high mountains, the snow cover (Marty et al., 2017) and glacier mass (Zemp et al., 2019) losses have been important since the 1950s, and even accelerated over the last two decades (Hugonnet et al., 2021). Both snow cover and glacier mass losses are consistently attributable to climate change, and will very likely continue in the twenty-first century regardless of any emission scenario (Fox-Kemper et al., 2021). This considerably impacts the regional water cycles in high mountains and at high latitudes. Streamflows increase as large glaciers melt, and then decrease when the water stock has lowered due to glacier mass loss and retreat (Bolch et al., 2010; Huss and Hock, 2018). Climate change also causes a lower fraction of precipitation falling as snow (except over the colder regions), which shifts and reduces streamflow in snowmelt-dependant catchments (Berghuijs et al., 2014). Permafrost thaw was already observed in the northern and high mountain regions, and is expected to strengthen in a warming climate, although the uncertainties are high. Thawing permafrost is likely to modify the regional hydrological processes (Walvoord and Kurylyk, 2016), and to release important CO_2 and CH_4 stocks through linear and non-linear processes (Turetsky et al., 2020). As a consequence to scarcer and more extreme precipitation, and to glacier and permafrost retreats, the river streamflows are expected to increase in winter and decrease in summer by 2100, with a global decreasing trend on annual mean (Douville et al., 2021). An earlier snowmelt due to global warming will also shift in time the peak runoff in snow-dominated river basins (Kang et al., 2016; Dudley et al., 2017).

The soil moisture (SM) projections show a general stronger decline in surface SM (top 10 cm) than in the total SM column (Berg et al., 2017). This is consistent with an increase in atmospheric evaporative demand, which dries the top-surface layer first. According to CMIP6

ensemble simulation results, regions that will experience drying of both surface and total SM column are mainly located in the subtropical regions (Central America, Mediterranean region, southern Africa, Southwest Australia), and the Amazonian basin (Cook et al., 2020; Douville et al., 2021). This corresponds roughly to areas with decreasing precipitation totals by 2100 in Figure 1.2. Other areas including northern high latitudes will experience a drying of the surface layer only. Overall, more than 40% of land areas are likely to experience enhanced drying by 2100, even under low emission scenarios (Cook et al., 2020). Drying is a consequence of lower precipitation, but also provoke a lower evapotranspiration and atmospheric moisture, contributing as a positive feedback to land aridification (Berg et al., 2016). Drought duration will significantly increase over drying subtropical areas (Ukkola et al., 2020), with a magnitude emissions-pathway dependent. In these regions, the likelihood of the most extreme droughts rises by 100% under SSP1-2.6 (low-emission scenario), up to 300% under the high-emission scenario (SSP5-8.5) (Cook et al., 2020). An increase in drought intensity is also projected in the tropics, despite the positive trend on annual mean precipitation except over the Amazon. This is due to the intensification of the water cycle, with an increasing number of dry days following more intense rainfall events (Figure 1.3).

The impact of land-use on the entire water cycle needs also to be considered. Deforestation at large scale causes lower precipitation (Boysen et al., 2020) and higher average runoff (Guzha et al., 2018; Levy et al., 2018), reducing the amount of water stored in the soils, the vegetation or at the surface. A direct water extraction in rivers reduced streamflow and caused the transformation of 139,000 km^2 of inland water areas to land from 1985 to 2005 (Donchyts et al., 2016). It is partially compensated by 95,000 km^2 of new artificial reservoirs created for irrigation or hydropower exploitation. The use of non-renewable water resources and climate change are expected to reduce groundwater recharge (Wada and Bierkens, 2014), especially in regions with intense irrigation such as northern India, (Asoka et al., 2017). It is altering the aquifers that store the largest freshwater resources on Earth, as for example the North African aquifer (Ramillien et al., 2014).

Due to the increased anthropogenic pressure and climate change, the magnitude and seasonality of surface water storage, runoff, and other hydrological-related variables will experience significant changes. As human populations grow and migrate, the access to freshwater in the world could be reduced for many people, exacerbating water scarcity and food insecurity (Schewe et al., 2014; Hoegh-Guldberg et al., 2018). The industrial, energy and agriculture sectors are all high consumers of water resources. There is therefore an emerging competition between food and energy security, called the "food-energy-water nexus" (D'Odorico et al., 2018). In this context, water scarcity already impacts up to 4-billion people (two-thirds of the current global population) at least one month a year, and half a billion people all year round (Mekonnen and Hoekstra, 2016). It is essential to predict, in order to mitigate, the consequences of climate change on the water cycle. Some uncertainties reduce the confidence in long-term predictions of CMIP6 and other Earth System Models (ESMs). It is the case for the interactions between water cycle and vegetation, which impact the evapotranspiration and precipitation patterns.

1.1.3 Interactions of the vegetation with the water cycle in a changing climate

In the previous section, the divergent regional predictions were highlighted, from drying in the subtropics and the Amazon, to wetting in the high latitudes and most of the tropics. Overall, an increased aridification is expected due to changes in precipitation, streamflow, and evapotranspiration patterns. The ambivalent influence of vegetation layers was not fully described up to now. It is especially essential to better understand the interactions between vegetation and the water cycle components, within the frame of climate change, in order to better constrain the Earth System and hydrological models, and to improve the confidence in their predictions. The vegetation is a key obstacle to aridification, as it favours water uptake by the soils, and exchanges moisture with the atmosphere through plant respiration, acting as a positive feedback to land precipitation.

The increase in atmospheric GHGs concentrations and climate change influence the interactions between water cycle and vegetation on several ways. At the leaf level, higher atmospheric CO_2 leads to enhanced photosynthesis, so to a higher Water-Use Efficiency (WUE) (De Kauwe et al., 2013; Hatfield and Dold, 2019; Lian et al., 2021). WUE is defined as the ratio between the plant, crop or biomass unit carbon uptake (through photosynthesis), and its water losses through transpiration or evapotranspiration when considering the response from the soils (Hatfield and Dold, 2019). Higher WUE due to rising atmospheric CO_2 concentrations have feedbacks on both the water cycle, and the biogeochemical cycles (see Section 1.2.1). It could contribute to save water in arid regions affected by precipitation losses. However, this effect may be offset by other processes.

Another physiological effect of rising CO_2 is the growth of plants and the expansion of leaves, leading to an increase in plant transpiration (De Kauwe et al., 2013; Ukkola et al., 2016). While a better WUE saves water, plant growth and leaf expansion increase the water use in vegetation layers, therefore decreasing runoff and soil moisture. The lengthening of the growing season is a direct effect of global warming, which increases the yearly evapotranspiration of vegetated land surfaces. There is low confidence in how these opposite processes balance each-other, and in what is the net hydrological response to higher atmospheric CO_2, although some studies suggest a reduced availability of water. In semi-arid regions, an increase in plant water use and reduced streamflows were observed, favoring aridification (Ukkola et al., 2016). In Europe, an increase in WUE from 14% (broadleaf forests) to 22% (coniferous forests) was found associated with a 5% increase in forest transpiration, due to the combined effects of larger greenness, evaporative demand and lengthened growing season (Frank et al., 2015).

Overall, evapotranspiration over land surfaces has consistently been increasing since the 1980s (Zhang et al., 2015, 2016; Zeng et al., 2018). This positive trend has been mainly attributed to anthropogenic forcing (Dong and Dai, 2017) and comprises interannual and decadal variability due to climatic events such as El-Niño/ Southern Oscillation (ENSO) (Miralles et al., 2014). At the same time, an increase in the Leaf Area Index (LAI) was found in most of the world (Zhang et al., 2016; Zhu et al., 2016). The positive trend in greening

was mainly attributed to CO_2 fertilization effect (enhanced photosynthesis with higher CO_2 concentrations) in the tropics, and climate change in the high latitudes (Zhu et al., 2016). The increasing vegetation greenness and transpiration is considered as one of the major factors of increasing land evapotranspiration, with the rise of atmospheric moisture demand (Zhang et al., 2015; Wei et al., 2017; Zeng et al., 2018).

The rising atmospheric CO_2 concentrations and climate change strongly impact the water cycle, through an increase of global atmospheric moisture demand, precipitation and evapotranspiration. Contrasting trends at the regional scale increase the risks of aridification and flooding, impact the magnitude and seasonality of river streamflows, the rate of groundwater recharge, agriculture and industrial activities. However, changes in the water cycle also interact with the biogeochemical cycles at regional and global scales, through positive or negative feedbacks to climate change. These interactions are an essential object of study in the frame of climate change mitigation.

1.2 The biogeochemical cycles

The IPCC AR6 report (Masson-Delmotte et al., 2021) has stated unequivocally that human activity is responsible for the rise of well-mixed atmospheric greenhouse gases (GHGs) concentrations since the pre-industrial era, and the resulting climate change. The well-mixed GHGs - carbon dioxide (CO_2), methane (CH_4) and nitrous oxide (N_2O) - are relatively evenly distributed at global scale, because their molecules have a quite long steady-state lifetime, at least several years. The residence times in the atmosphere are on average 9.1 years for CH_4 (Section 6.3 in Szopa et al. (2021)), 116 years for N_2O (Prather et al., 2015) and from one to hundred thousands of years for CO_2 which has multiple lifetimes, depending on the timescales of several removal processes (Ciais et al., 2013). Their lifetimes influence the global accumulation and distribution of these GHGs, and the radiative forcing associated. The Global Carbon Project, established in 2001, publishes reports to assess the budgets of $C0_2$ (Friedlingstein et al., 2020, 2022), CH_4 (Saunois et al., 2020) and N_2O (Tian et al., 2020). The trends of natural and anthropogenic sources, sinks and feedbacks related to GHGs accumulation in the atmosphere are described, and deeply linked to evolutions in the water cycle and wetland ecosystems.

The cycles of the two GHGs responsible for the major radiative forcing, namely carbon dioxide and methane, are described in the two following subsections. The current state of the art is further used to define the CO_2 and CH_4 budgets, sources and sinks (Friedlingstein et al., 2020; Saunois et al., 2020; Canadell et al., 2021). In particular, for the purpose of this manuscript, two interesting processes are described. For the CO_2 budget, the importance of the land sink is highlighted, as well as the dependence of climate-carbon feedback mechanisms on the water cycle evolution. For the CH_4 budget, the large estimated emissions from wetlands and inland waters with their uncertainties are shown, to highlight the need for a better understanding of tropical wetland dynamics.

1.2.1 Carbon dioxide budget and feedbacks

The global CO_2 budget (Friedlingstein et al., 2020, 2022) provides yearly estimations of the five major sources and sinks of CO_2, namely the fossil CO_2 emissions (E_{FOS}), emissions from land-use change (E_{LUC}), the growth rate of atmospheric CO_2 concentrations (G_{ATM}), and the ocean and land sinks (S_{OCEAN} and S_{LAND}, respectively). All the terms being independently estimated, either from observations or modeling processes, the positive and negative CO_2 fluxes do not balance perfectly. The budget imbalance (BIM) corresponds to the difference between estimated sources and sinks, and the final CO_2 budget is estimated as:

$$E_{FOS} + E_{LUC} = S_{OCEAN} + S_{LAND} + G_{ATM} + BIM \tag{1.3}$$

The yearly and cumulative evolution of the carbon budget, with its six terms, is represented in Figure 1.4. The difference between the red, dashed lines and the total negative fluxes represents the budget imbalance. The atmospheric growth rate in blue is still increasing, up to the value of 5.1 ± 0.02 GtC·yr^{-1} on average for the decade 2010-2019 (Friedlingstein et al., 2020). When combined, the carbon uptake from land ($S_{LAND} = 3.4 \pm 0.9$ GtC·yr^{-1}, ~31% of total emissions) and ocean ($S_{OCEAN} = 2.5 \pm 0.6$ GtC·yr^{-1}, ~23% of total emissions) sinks barely exceed a half of the combined emissions from anthropogenic activity ($E_{FOS} + E_{LUC} = 10.9 \pm 0.9$ GtC·yr^{-1}). The remaining part is stored in the atmosphere (G_{ATM}), with a small budget imbalance. All the terms in Equation (1.3) have significantly increased since the pre-

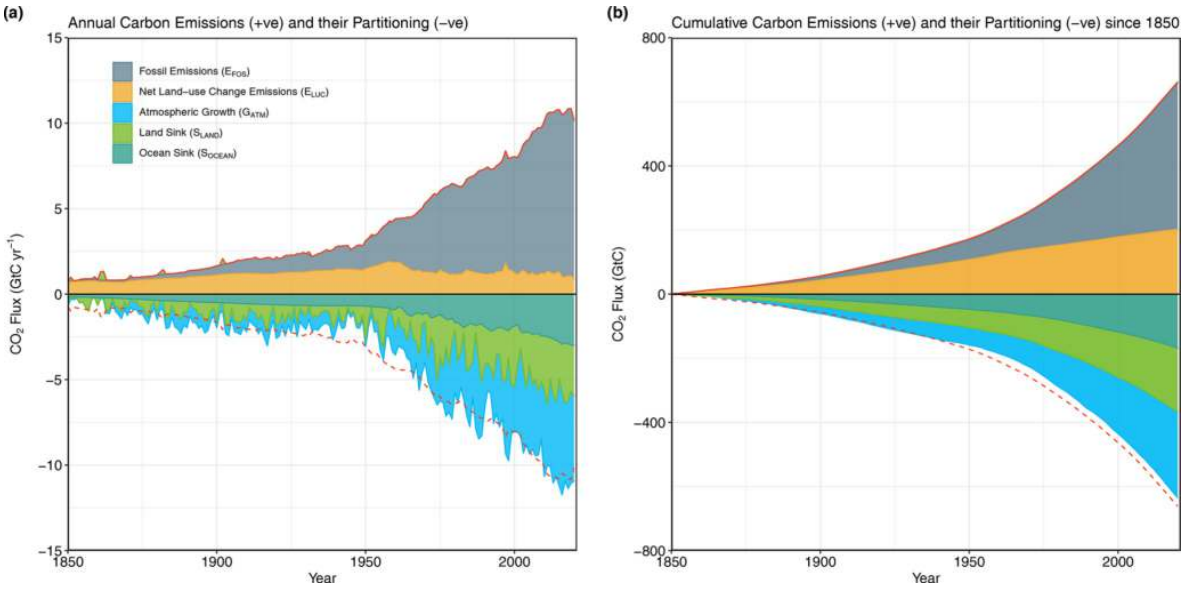

Figure 1.4: Historical records since 1850 of the carbon sources and sinks described in Equation (1.3), from Friedlingstein et al. (2022). **(a)** annual CO_2 fluxes, **(b)** cumulative CO_2 fluxes. The red dashed lines are symmetric to the red lines, which represent the total of positive fluxes (CO_2 sources). The difference between the dashed lines and the total negative fluxes (sinks and atmospheric growth) represents the budget imbalance (BIM), which balances Equation (1.3).

industrial era except E_{LUC}. Fossil energies are now responsible for 86% of the anthropogenic CO_2 emissions and are still growing at the rate of 3.0% in the 2000s, and 1.2% in the 2010s. E_{LUC} includes both gross emissions and gross removals (as carbon uptake by the growing new vegetation as an example) from land-use change, and contributes to the remaining fraction of anthropogenic CO_2 emissions.

The airborne fraction represents the fraction of anthropogenic CO_2 emissions which is stored in the atmosphere every year. It has been quite stable since 1960 around an average value of 44% (Friedlingstein et al., 2020), with a strong interannual variability due to climatic events such as ENSO (see Figure 1.4a). Every year CO_2 accumulates in the atmosphere, which is obvious looking at Figure 1.4b. The rate of accumulation depends on the contributions of land and ocean sinks, and the mechanisms of climate-carbon feedbacks. These feedbacks decrease (positive feedback) or increase (negative feedback) the fraction of emitted CO_2 which is captured by carbon sinks, and therefore directly affect the carbon budget. The rise of atmospheric CO_2 concentration acts as a negative feedback, while the rise of surface temperatures acts as a positive feedback (Williams et al., 2019). Overall, the stability of airborne fraction over the last decades implies a growth of both land and ocean CO_2 sinks, to compensate larger emissions (Friedlingstein et al., 2020). Two major contributions of increasing ocean and land uptakes are the buffering capacity of the ocean, and the CO_2 fertilization effect which enhances photosynthesis.

The relationship between cumulative CO_2 emissions and the level of warming can be studied through the Transient Climate Response to cumulative CO_2 Emissions (TCRE) (MacDougall, 2016). TCRE shows a near-constant warming response to cumulative CO_2 emissions on decadal to centennial time scales (Goodwin et al., 2015; MacDougall and Friedlingstein, 2015; Williams et al., 2016; MacDougall et al., 2017). This property is due to the equilibrium of the radiative forcing per unit mass of CO_2, and the carbon and heat uptake by the ocean (MacDougall and Friedlingstein, 2015; Williams et al., 2016). At interannual time scale, the strong variations of the land carbon sink (see Figure 1.4a) and feedbacks also cause variations in TCRE, with low effect on longer climate responses (MacDougall and Friedlingstein, 2015; Goodwin et al., 2015). At decadal time scales, the climate warming response to cumulative CO_2 emissions remains linear due to the ocean carbon uptake. TCRE has been very useful to estimate the remaining carbon budget to exceed a certain warming limit - such as the objective of maximum 1.5 °C of global warming in the Paris agreement (Hoegh-Guldberg et al., 2018; Jones and Friedlingstein, 2020), although its uncertainties remain high (MacDougall et al., 2017).

The land CO_2 sink is deeply linked to carbon uptake from the vegetation through photosynthesis. Photosynthesis has increased over the past decades, which is mainly attributed to a fertilization effect due to the rise of atmospheric CO_2 concentrations (Walker et al., 2021). The increase of atmospheric CO_2 therefore clearly represents a negative climate-carbon feedback as shown in Figure 1.5a. On the contrary, the sign and magnitude of the feedback from a warming climate alone are divergent across model outputs, spatially heterogeneous (Figure 1.5b), and represent a large cause of uncertainty. The high interannual variability of land sink is mainly responsible for the fluctuations of atmospheric CO_2 growth rate (see

Figure 1.4a and Friedlingstein et al. (2020)). This variability is dominated by the response of climate-sensitive tropical and semi-arid ecosystems to climate variations (Cox et al., 2013; Piao et al., 2020).

Two key parameters were intensively studied to explain climate-driven variations: the temperature (Cox et al., 2013; Fang et al., 2017; Jung et al., 2017) and the water availability

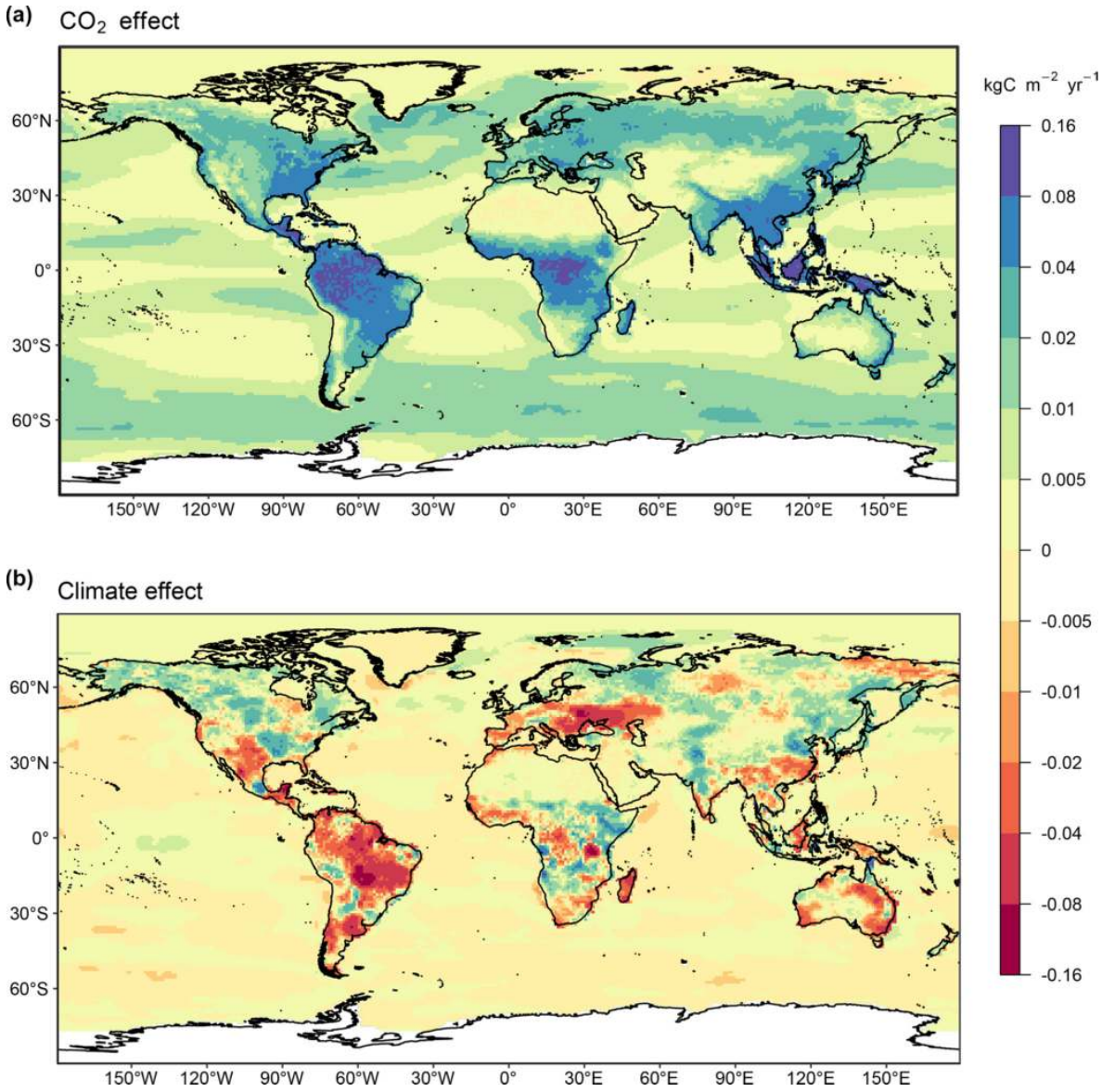

Figure 1.5: Attribution of the CO_2 fluxes between land, ocean and atmosphere to increasing atmospheric CO_2 concentrations and climate effect, from Friedlingstein et al. (2022). The effect of CO_2 only exceeds the contrasting effect of climate only, in particular due to the large contribution of tropical and boreal forests to the land sink. The climate effect shows high climate-carbon positive feedback (in red) in regions where negative precipitation trends and drying are found (see Section 1.1.2 and Figure 1.3).

(Fang et al., 2017; Jung et al., 2017; Humphrey et al., 2018). The responses to heat and moisture have a strong interannual variability, due to shifts in climate indexes such as El Niño Southern Oscillation (ENSO), the Indian Ocean Dipole (IOD) and the Southern Annular Mode (SAM). For the case of ENSO, El Niño events are both warmer and drier, and La Niña cooler and moister. Overall, the contributions of heat and moisture are hard to dissociate but both are admitted. Jung et al. (2017) used modeling approaches to conclude that the temporal variability of land - atmosphere CO_2 fluxes is mostly driven by the fluctuations in temperature, while the availability of water accounts for spatial anomalies. On the contrary, Humphrey et al. (2018) showed a relationship between the observed changes in terrestrial water storage (TWS) derived from the GRACE space mission and CO_2 growth rate, higher on drier years. They also concluded that this relationship is independent from the observed temperatures. The droughts reduce water availability and certainly weaken CO_2 uptake by the vegetation through lower photosynthesis. They also increase forest mortality and wildfires, all contributing to a weakening of the land CO_2 sink (Phillips et al., 2009; da Silva et al., 2018; Bennett et al., 2021). During strong El-Niño events, the land sink shows a large negative anomaly, and the airborne fraction consequently increases. Although a decrease of ∼20% on global burnt areas was observed over the last two decades, CO_2 emissions from fires will likely increase in the future in response to global warming (Canadell et al., 2021).

As it was previously discussed in Section 1.1.3, higher atmospheric CO_2 and climate change increase the plant water-use efficiency, leaf expansion and plant growth, and lengthen the growing season of the vegetation (De Kauwe et al., 2013; Ukkola et al., 2016; Lian et al., 2021). Although it is not totally understood and comprises large uncertainties, this may increase the water used by the vegetation in the future (Frank et al., 2015; Ukkola et al., 2016). With larger droughts and a reduced availability of water, there is a risk of a lower efficiency of the land CO_2 sink at the end of the twenty-first century due to an increasing water stress affecting the vegetation productivity (Green et al., 2019). Such an effect would provoke a rise of the airborne fraction, and higher CO_2 accumulation in the atmosphere. This could affect the linearity of TCRE, reduce the remaining carbon budget to be emitted before exceeding the 1.5 °C or 2 °global warming objective, and finally lower the efficiency of climate change mitigation processes being conducted.

Finally, permafrost (as peatlands) will likely exert positive climate-carbon feedbacks, as it has already been converted from carbon sink to net source of CO_2 (Canadell et al., 2021). It also represents a large source of CH_4 (see Section 1.2.2). The magnitude of the phenomena is highly uncertain, even within the path of one emission scenario. In particular, there is low confidence on the linearity of permafrost carbon feedbacks due to warming. Abrupt permafrost thaw could provoke collapsing ground, erosion and landslides, and release as much CO_2 within these processes as projected for gradual thaw emissions (Turetsky et al., 2020). Overall, the linearity of TCRE and remaining carbon budgets will be affected by permafrost CO_2 feedbacks (MacDougall et al., 2015).

To conclude, the land carbon sink shows variations and impacts the CO_2 budget at interannual time scale, and has been stable at longer time scales. Terrestrial climate-carbon feedbacks will affect either positively or negatively the future climate. The land CO_2 sink

was shown to be strongly dependent on the vegetation productivity, which is driven by water availability and temperatures. It could be weakened in case of prolonged droughts leading to water stress, and warming climate. Such parameters also directly affect the methane emissions from wetlands, which are explored on the following subsection.

1.2.2 Methane budget

Methane (CH_4) is a much more powerful GHG than CO2, and is responsible for the second highest radiative forcing in the atmosphere (Etminan et al., 2016). The CH4 emissions come from both human activity - mainly fossil energies, agriculture and waste, and biomass burning - and natural sources like wetlands, inland waters, geological sources, oceans, animals and wildfires (Saunois et al., 2020). The major sinks are chemical: 90% of the atmospheric CH_4 is removed in the troposphere while reacting with hydroxyl radical (OH), and 5% is removed in the stratosphere through several reactions. The resulting 5% sink is from land uptake by microbial activity (Saunois et al., 2020). The CH_4 emissions have nearly doubled since the pre-industrial era and this rise is mostly driven by anthropogenic emissions, resulting in a continuously positive growing rate of atmospheric CH_4. Still, large uncertainties exist in the partition of the emissions, mainly from natural sources, due to discrepancies between the *top-down* estimates based on atmospheric inversion, and the *bottom-up* estimates based on the upscaling of measurements and inventories. For now, bottom-up estimates are always larger than top-down estimates, the latter being constrained by consistent measurements of atmospheric CH_4 concentrations. There is particularly a consequent lack of quantification and partition between the various natural sources of methane (Canadell et al., 2021).

Figure 1.6 represents the global methane budget for 2008-2017, as published in the IPCC AR6 report (Canadell et al., 2021) close to the budget from Saunois et al. (2020). The values in Figure 1.6 are the minimum-maximum ranges in bottom-up model estimates, to decompose the several natural and anthropogenic sources. Both top-down estimates and bottom-up average estimates are given in the IPCC AR6 report (Table 5.2 in Canadell et al. (2021)). The anthropogenic part of total CH_4 emissions vary between 49% (bottom-up) and 62% (top-down), the former likely overestimating the natural CH_4 sources (Saunois et al., 2020). They are dominated by agriculture and waste management, of which livestock production form a large part due to enteric fermentation and manure. It represented on average 109·Tg yr^{-1} in the 2008-2017 decade (∼30% of anthropogenic emissions), in a rising trend inferred to the augmentation of total animal number. Landfill and waste (18% of anthropogenic emissions), rice cultivation and biomass burning / biofuels (5% each) also contribute to the growing atmospheric CH_4. Finally, the fossil fuels represent ∼32% of all anthropogenic CH_4 emissions and are on an rising trend.

Wetland ecosystems represent the single largest source of methane, from 149 Tg·yr^{-1} to 180 Tg·yr^{-1} and from 20% to ∼30% of total CH_4 emissions, depending on bottom-up or top-down estimates, respectively (Saunois et al., 2020; Canadell et al., 2021). Furthermore, bottom-up approaches count separately the freshwater emissions from lakes and rivers (159 Tg·yr^{-1}), while top-down atmospheric inversions does not differentiate freshwater and wetland

emissions. This causes a large discrepancy between the total natural CH_4 emissions estimated with both methods: 371 $Tg{\cdot}yr^{-1}$ against 215 $Tg{\cdot}yr^{-1}$, respectively. While bottom-up estimates of freshwater and wetland emissions represent 42% of total CH_4 sources, it is likely that a double-counting occurs which partially explains the gap with top-down approaches (Canadell et al., 2021). The inland water are furthermore the largest source of uncertainty in the methane budget (Saunois et al., 2020).

The changes in inundated areas, air temperature and microbial activity cause large interannual variations of wetlands CH_4 emissions (Bridgham et al., 2013). Trees in floodplains play a key role in transporting methane from the soils to the atmosphere in tropical rainforests (Pangala et al., 2017). This increases bottom-up estimates of methane emissions and widens the gap with top-down estimates, although a good agreement was found for the Amazon CH_4 budget in Pangala et al. (2017). It highlights the need for a better understanding and modeling of the CH4 fluxes. In particular, it is essential to better constrain the CH_4 emissions in the tropics, which are negatively impacted by a poor definition and mapping of wetland dynamics (Bridgham et al., 2013; Pangala et al., 2017; Covey and Megonigal, 2019). This is a key point to help closing the methane budget which is dominated by tropical emissions of

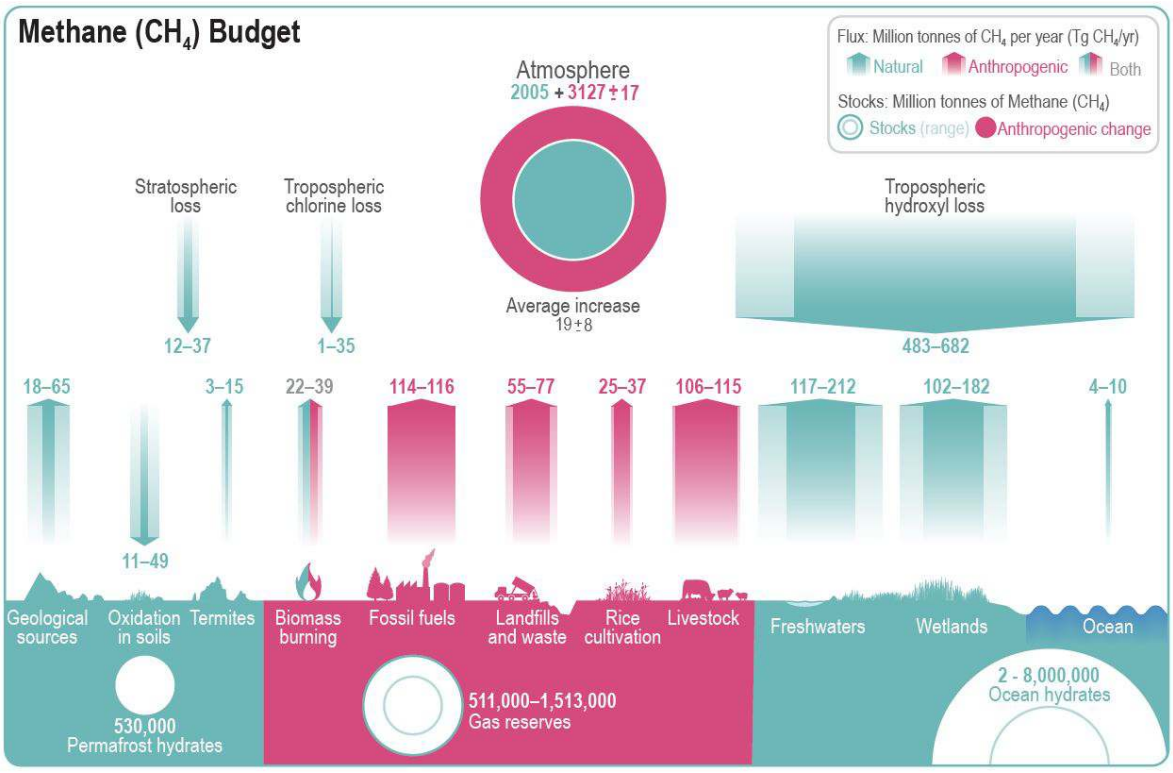

Figure 1.6: Global methane budget representing the average fluxes per year over 2008-2017, from Canadell et al. (2021). The fluxes are representing bottom-up estimates with the separation of many anthropogenic (in red) and natural (in green) CH_4 sources. Each flux is given as the range between minimum and maximum model estimates. The data are from the global methane budget in Saunois et al. (2020) and updated with several sources.

up to 65% (Saunois et al., 2020).

To conclude, the methane budget presents high uncertainties related to the partitions of natural and anthropogenic sources. First, no agreement is found between bottom-up and top-down estimates at global scale, which highlights the need for better constraining the CH_4 dynamics and fluxes through improved observations and modeling processes. Then, large interannual variations of the natural emissions are found in response to climatic events and microbial activity, in a flurry of processes that require further investigations to be fully understood. Finally, it is essential to disentangle the relative contributions of freshwater in lakes and rivers, and wetlands to the global CH_4 budget from bottom-up estimates. Comprehensive modeling of the inland-water emission processes and their spatial extent is required. Also, a better mapping of wetland dynamics, in particular to monitor the temporal variations of freshwater storage in floodplains, should reduce uncertainties in the upscaling of methane fluxes based on observations (Pangala et al., 2017; Saunois et al., 2020). A finer representation of whether positive trends in atmospheric CH_4 growth is due to human activity of natural processes, is required to implement adapted mitigation strategies in response to climate change.

1.3 Remote sensing of the surface water dynamics

It was shown in Section 1.1 and Section 1.2 that freshwater terrestrial water storage (TWS) and its dynamics are essential to understand the regional and global changes in water and biogeochemical cycles. The different terrestrial water components are the surface water storage (SWS), water stored in soil layers, ice and snow packs, and groundwater. Since 2002, the Gravity Recovery and Climate Experiment mission (GRACE, Tapley et al. (2004)) and its successor GRACE Follow-On were able to monitor changes in TWS (ΔTWS), from basin to global spatial scales (Tapley et al., 2019). ΔTWS in a basin is defined using the terrestrial water balance equation (Peixoto and Oort, 1992):

$$P - ET - G = Q + \Delta TWS \tag{1.4}$$

where P and ET are the precipitation and evapotranspiration in the area, respectively, G represents the groundwater fluxes, and Q is the total discharge across the basin boundaries. Although ΔTWS is computed from GRACE, the partition of precipitated water into the different TWS components suffers from high uncertainties. Satellite observations, in particular active and passive microwave remote sensing, were used to derive snow water equivalent, soil moisture (SM) and SWS, while groundwater (GDW) is inferred using TWS from GRACE and all other estimations. However, snow water equivalent is derived from passive microwave instruments with a coarse spatial resolution, and is not correctly estimated on mountains (Foster et al., 2005). The surface SM is also derived from passive microwave sensors at low spatial resolution (Chan et al., 2016b, 2018; Wigneron et al., 2021), but Root Zone Soil Moisture (RZSM) and the total SM column are very complicated to estimate. Finally, the spatial and temporal variations of SWS are poorly determined, although the upcoming Surface Water

and Ocean Topography (SWOT) mission should provide reliable estimations (Biancamaria et al., 2016). The uncertainties propagate into global hydrological and biogeochemical models, and cause a large dispersion in the estimates of land CO_2 sink and natural CH_4 emissions from wetlands (see Section 1.2).

The use of in-situ gauges is an efficient way to monitor the water levels, flows and temporal variations of river discharges. However, they are not able to monitor the spatial extent and water storage in most of the floodplains due to inherent limitations. They cover a very few portion of the major rivers and wetlands, are mainly located in industrialized regions, and the number of data available has consistently declined over the past decades (Shiklomanov et al., 2002; Alsdorf and Lettenmaier, 2003). So, surface water extent (SWE) and SWS monitoring could strongly benefit from global and consistent estimates derived from remote sensing. Although SWE and SWS were derived from several remote sensing instruments and techniques, improvements are still required to provide a consistent knowledge of the evolution and partitioning of freshwater stocks. In this section, a critical review of current surface water products with their limitations is provided. The recent advances and future developments are also highlighted. Within the frame of this manuscript, the poor monitoring of tropical wetlands is emphasized to pave the way to new insights provided by GNSS-Reflectometry.

Section 1.3.1, Section 1.3.2 and Section 1.3.3.2 introduce respectively the early global SWE determinations, the estimations based on visible and infrared imagery and those based on both passive and active microwave datasets. These sensors, their characteristics, the main advantages and the limitations of all products are highlighted. Then, Section 1.3.4 presents and discusses the methods for retrieving SWS based on SWE. Finally, Section 1.4 and Section 1.5 discuss the current limitations and future advances of hydrological sciences.

1.3.1 Early Surface Water Extent estimations

The determination of the total extent, size and geographical distributions of lakes and other inland water bodies is a key issue for a better understanding of the global water cycle. It has long been conduced using inventories at the local, regional, national or continental scales. With the rising concerns about global water and biogeochemical cycles, the monitoring of SWE in land surfaces is crucial. The availability of a global, high-resolution water mask is also very important for many land applications where the water bodies should be masked out: fire disturbances, surface reflectance and temperatures, soil moisture, cloudiness, etc. The early global maps of water bodies, including wetlands, were generated through the combination of several vector datasets at different spatial scales, such as the Global Lakes and Wetlands Database (GLWD, Lehner and Döll (2004)). They suffered from a low spatial resolution, and concerns with the vectorization or the information represented.

Global estimates of the lake sizes and total surface water area first relied on statistical extrapolations to include small reservoirs (Downing et al., 2006). These methods concluded that small lakes and pounds were of extreme importance and contributed to about half of the total inland water area (Downing et al., 2006; Downing, 2010). Further studies also found

that the very small reservoirs emit around twice more CO_2 and five times more CH_4 than other inland water bodies (Holgerson and Raymond, 2016). A precise quantification of the global size distribution and the total surface of inland water bodies is therefore essential, to better constrain freshwater stocks in both hydrological and biogeochemical models.

A first high-resolution water mask was created using data from the Shuttle Radar Topography Mission (SRTM) launched by the National Aeronautics and Space Administration (NASA). During 11 days in February 2000, the SRTM performed Interferometric Synthetic Aperture Radar (InSAR, see Section 1.3.4.3) measurements of the Earth's topography at C and X bands, with 1 arcsec spatial resolution (∼30 m at the equator). It supplied a near global high-resolution Digital Elevation Model (DEM) at 90 m (30 m now), covering land surfaces from 56°S to 60°N. The SRTM Water Body Data (SWBD) was further derived at 30 m spatial resolution, but does not cover the high latitudes. Moreover, some issues were highlighted as the connectivity of channels was not guaranteed. Still, it was used as input for further water body masks based on remote sensing data (Carroll et al., 2009).

In the following sections, two distinct types of water masks and SWE estimations are presented. First, the products based on visible and infrared imagery have high-to-moderate spatial resolutions, but are limited over tropical wetlands and other vegetated areas. They are mainly based on two sensors: the Moderate Resolution Imaging Spectroradiometer (MODIS) onboard Terra and Aqua, and the Landsat program, all being NASA satellites. Secondly, the microwave-based surface water products, which are either active or passive. Some of these sensors are able to monitor flooded vegetation, but they have distinct characteristics and limitations which are described in Section 1.3.3.1.

1.3.2 SWE mapping from visible and infrared imagery

A first global, moderate resolution (250 m) and reliable water mask was produced by Carroll et al. (2009), using the NASA SWBD product and MODIS reflectances at 250 m as inputs. In this product, MODIS complements SWBD to fill the gaps in the spatial coverage and correct the wrong channel connectivities. It was produced as the remote sensing community needed a precise, moderate resolution (250 m) water mask for a flurry of applications over land surfaces. However, with the increasing computational and storage capacities and the need for a global, high-resolution surface water dataset, its spatial resolution was too low. In particular, it was not able to connect small river streams with lakes and ponds, and shoreline pixels had insufficient level of detail.

The detection of inundated areas in wetlands using MODIS was assessed in Chen et al. (2013) against Landsat images. Both the daily MODIS observations and the 8-day composite product showed good results in detecting water. The seasonal and interannual changes in wetland extent were also observed from MODIS at 500 m spatial resolution, over the Sudd wetlands in Sudan (Di Vittorio and Georgakakos, 2018). Similar results were obtained over other river basins, as the Mekong and the Mackenzie (Sakamoto et al., 2007; Normandin et al., 2018). Although they are consistent at the regional scale, a global retrieval is unlikely

due to limitations of the MODIS sensors over densely vegetated areas. Moreover, the spatial resolution at 250 or 500 m is also too low for many applications.

The European Space Agency's (ESA) Sentinel-2 satellites are also in orbit since 2015 (Sentinel-2A) and 2017 (Sentinel-2B). They provide high-resolution data (10 or 20 m for most of the spectral bands) useful for mapping floods. Sentinel-2 was already used in several regional studies (Caballero et al., 2019; Goffi et al., 2020), and it is incorporated in Copernicus flood monitoring services. The combination or radar and optical Sentinel-1 and Sentinel-2 datasets is particularly interesting for flood mapping, both at the regional (Huang and Jin, 2020; Bai et al., 2021; Tarpanelli et al., 2022) and global (Martinis et al., 2022) scales. More information on Sentinel-1 and combined products can be found in Section 1.3.3.2. However, Sentinel-2, having been launched recently, is not able to provide insights into interannual and decadal variability of the water cycle.

The high-resolution and long-term requirement can be solved using Landsat. The opening of Landsat archives since 2008 (Woodcock et al., 2008), and their further consolidation at United State Geological Survey (USGS) (Wulder et al., 2016), have make available millions of images from the 1970s up to now in a common database. The spatial resolution is 30 m since Landsat 4, launched in 1982, and 15 m for the panchromatic band in Landsat 7-9 starting from 1999. They were used in several studies to map either water masks or seasonal and interannual changes in SWE at the global scale (Verpoorter et al., 2014; Yamazaki et al., 2015; Pekel et al., 2016; Feng et al., 2016). The first global dataset was provided by Verpoorter et al. (2014), where the abundance and size distribution of lakes is described using a set of Landsat 7 Enhanced Thematic Mapper Plus (ETM+) from 2000. They found ∼117 millions of lakes greater than 0.002 km^2 covering about 3.7% of the land non-glaciated areas (5.0×10^6 km^2), which represent a lower number of lakes but a larger surface covered than early estimates based on statistical extrapolation (Downing et al., 2006).

Further efforts were made to map all components of surface water. Feng et al. (2016) also considered a snapshot of year 2000 to produce a Landsat-based inland water mask at ∼30 m spatial resolution. They found $\sim 3.65 \times 10^6$ km^2 of inland water area at global scale, dominated by wetlands and lakes in boreal high latitudes. However, these static estimates of water vs land miss the temporal changes of SWE, with large seasonal and interannual variations of inland water area. Yamazaki et al. (2015) used four sets of Landsat images from 1990, 2000, 2005 and 2010, to map surface water with its evolution at 3" spatial resolution (∼90 m). This allowed it to define a probability of water and to produce a mask of both permanent and temporal water, covering ∼3.25 and $\sim 0.49 \times 10^6$ km^2, respectively.

The reference inland water dataset was provided by Pekel et al. (2016), with an example shown in Figure 1.7. Using 3 millions of Landsat images, the Global Surface Water (GSW) product from the Joint Research Center (JRC) permits a mapping of surface water occurrence, recurrence and seasonality at 30 m spatial resolution since 1984. The long time series and the precision of this dataset allowed to map changes in both permanent and seasonal water at high-resolution, and interannual to decadal time scales. This is essential for climate applications to overcome the interannual variability of the water cycle. It provides a very useful tool to observe changes in regional and global hydrology over the past four decades, and evaluate the

contribution of anthropogenic pressure and climate change to these evolutions.

The inland water mapping using optical and infrared imagery suffers from several limitations. Misclassifications occur due to the confusion of water with shadow or snow and ice, but great efforts have been made to overcome this thresholding problem in the retrieval algorithms. More problematic, the signals are not able to penetrate either cloud cover or vegetation layers. This inherent limitation of the sensors highly affects the surface water detection in tropical wetlands, where both clouds and dense canopies are regularly found. As a consequence, the global SWE estimations using optical sensors are dominated by northern boreal lakes and wetlands, and with a clear underestimation of inundation extent in the tropics (see Figure 1.7 compared to the estimate based on passive microwave sensors in Figure 1.13).

In compliance with this limitation, the validation of the GSW product suffers from a high rate of omissions affecting seasonal water. According to the Extended Data Table 1 in Pekel et al. (2016), it omits ∼25% of the seasonal water pixels, but only 1-2% of the permanent water pixels. It has therefore proven to be extremely useful to map permanent water globally, and seasonal water regionally only when the vegetation is sparse or below water. The mapping of tropical wetlands, where flooded forests are commonly found, therefore highly relies on passive or active microwave sensors.

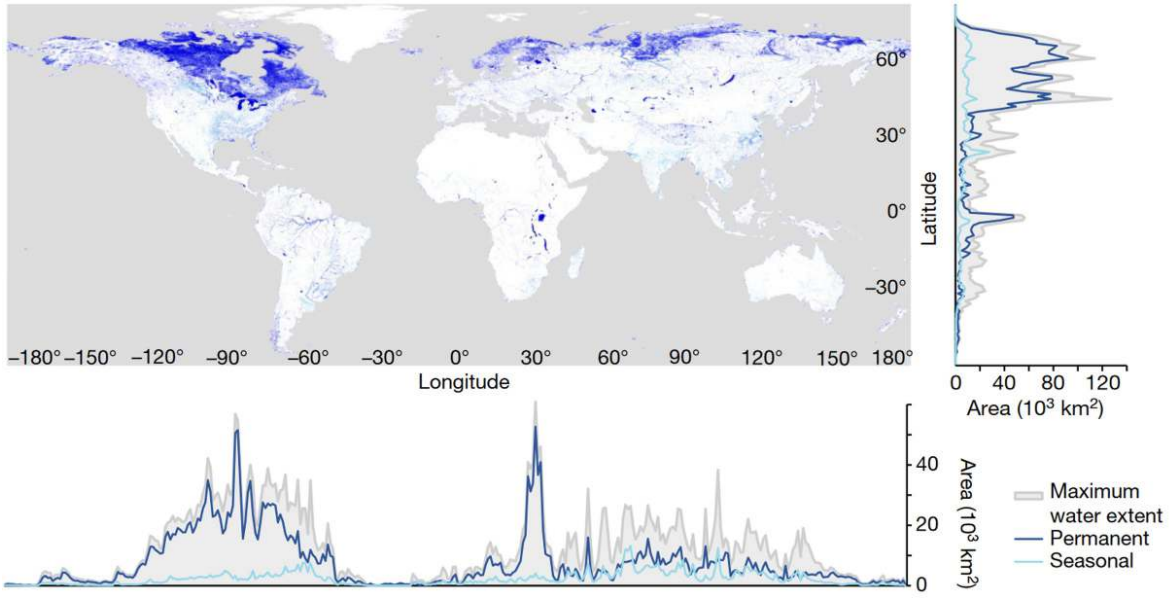

Figure 1.7: Global distribution of permanent and seasonal water in the GSW dataset from October 2014 to October 2015, from Pekel et al. (2016). The latitudinal and longitudinal profiles corresponding to the map are shown with the seasonal, permanent and maximum water extent. A dominant contribution from the northern boreal regions is highlighted.

1.3.3 SWE mapping from passive and active microwave remote sensing

1.3.3.1 Microwave sensors

Microwave signals are operating in almost all weather conditions during both day and night, and are able to partially penetrate the vegetated layers. These are the main advantages that make them suitable to remotely sense information from land surfaces, especially over the tropics where both cloud cover and vegetation obstruct the visible and infrared wavelengths. Microwave sensors are crucial for monitoring tropical wetland hydrology, where MODIS and Landsat-based products perform the worse. Several wavelengths are used by microwave sensors to observe the Earth surface. As illustrated in Figure 1.8, a lower frequency (higher wavelength) permits a better penetration of the signal in vegetation layers, especially trees. The P and L-band sensors are therefore optimal for monitoring water under forests and crops canopies.

Two different types of sensors are found: passive systems that measure the brightness temperature related to the emissivity of soil and vegetation layers at a given wavelength, and active systems carrying onboard a transmitter emitting the radar signal, and a receiver

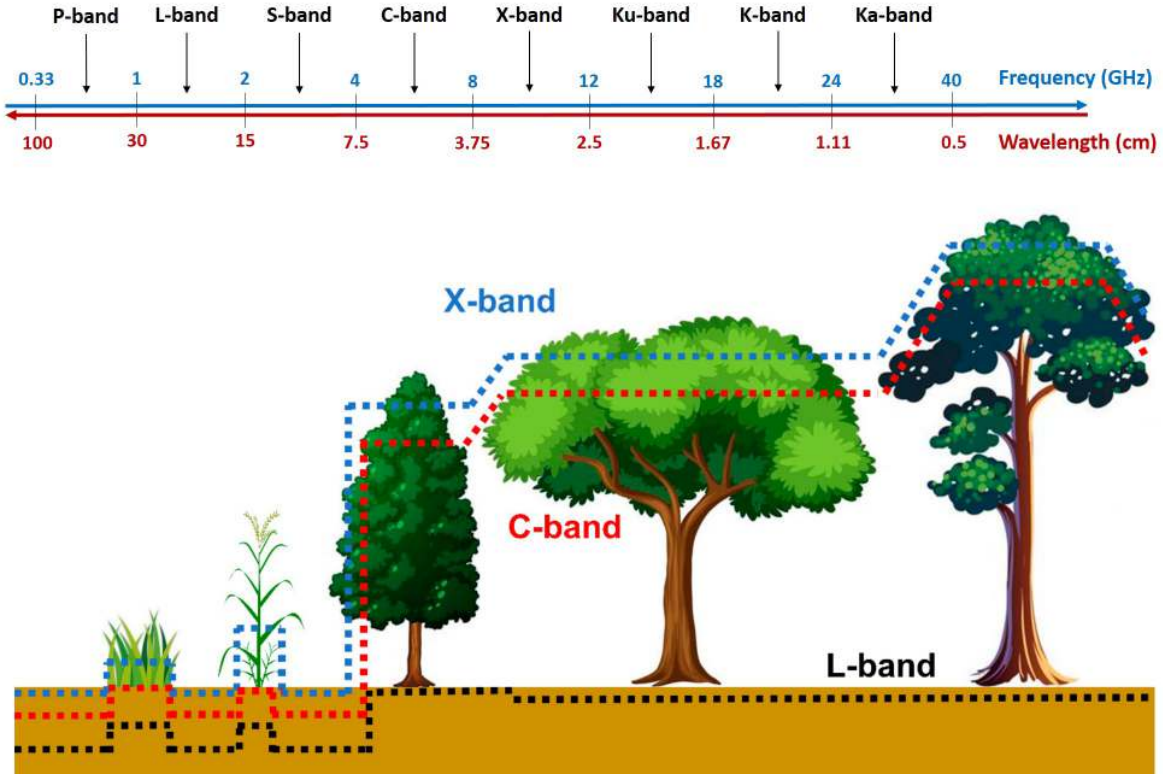

Figure 1.8: **Top**: spectrum of the microwave frequency bands used in remote sensing sensors and their penetration in vegetation layers. **Down**: a higher frequency (lower wavelength) reduces the penetration of microwave signals, as illustrated with X, C and L bands. Adapted from Frappart et al. (2020).

collecting the reflections of this signal on the Earth's surface. Radar altimeters are one type of active sensor, with nadir-pointing antennas emitting one beam and recording the waveform of the reflected signal. Other type of active sensors are Synthetic Aperture Radars (SAR), which emit successive beams while the satellite is moving. They use side-looking antennas with large swath widths and the motion of the spacecraft to reconstruct 2D images from the reflected signals. The motion of the satellite platform makes it possible to obtain high-resolution images ($<$ 1m) from relatively small physical antennas, a process named *synthetic aperture*. Overall, many types of passive radiometers, SAR and radar altimeters were launched and used to study surface water dynamics. They are respectively listed in Table 1.1, Table 1.2 and Table 1.3.

Microwave radiometers

For passive microwave sensors listed in Table 1.1, the lower is the frequency of the signal, the lower is the spatial resolution of observed brightness temperatures. The radiometers have been widely used since the 1970s, with the launch of the Scanning Multichannel Microwave Radiometer (SMMR) in 1978, and its successors which are the Special Sensor Microwave/Imager (SSM/I) and Special Sensor Microwave Imager/Sounder (SSMIS). Together, they create a long time series of brightness temperatures at several frequencies and polarizations, which is very useful for the mapping of long-term inundation dynamics. The presence of freshwater indeed causes a decrease in observed brightness temperatures, more pronounced for the horizontal than for the vertical polarization (Sippel et al., 1994; Prigent et al., 2001).

Other microwave sensors were dedicated mainly or partially to hydrology researches. The Advanced Microwave Scanning Radiometer - Earth Observation System (AMSR-EOS) and its successor AMSR2 cover a wide range of frequencies and, therefore, spatial resolutions. Similarly, the Tropical Rainfall Measuring Mission (TRMM) and its successor Global Precipitation Measurement (GPM) use several instruments including a microwave radiometer operating from 10.7 GHz to 183 GHz, to monitor precipitation amounts. The Soil Moisture

Table 1.1: Summary of the main microwave radiometers used for remote sensing of the water cycle in the last decades, with their principal features. The data are organized by date of launch.

Platform/ Sensor	Band/ Frequency	Resolution	Revisit	Acquisition period
SSM/I	Ku-Ka (19.35-85.5 GHz)	13-69 km	<daily	1987-
TRMM/ TMI	X-Ka (10.7-85.5 GHz)	6-72 km	<daily	1997-2015
AMSR-E	X-Ka (6.9-89 GHz)	5-74 km	1-2 days	2002-2016
SSMIS	Ku-Ka (19.35-183 GHz)	13-73 km	<daily	2003-
SMOS	L (1.4 GHz)	35-50 km	2-3 days	2009-
SAC-D/ Aquarius	L (1.4 GHz)	∼100 km	7 days	2011-2015
AMSR2	X-Ka (6.9-89 GHz)	5-10 km	2 days	2012-
GPM/ GMI	X-Ka (10.7-183 GHz)	4-32 km	<daily	2014-
SMAP	L (1.4 GHz)	40 km	2-3 days	2015-

and Ocean Salinity (SMOS, Kerr et al. (2001)) and Soil Moisture Active Passive (SMAP, Entekhabi et al. (2010)) missions are notably dedicated to studying soil moisture (SM) dynamics. They provide L-band brightness temperatures at 35-50 km spatial resolution and 2-3 days temporal resolution. They were also used to produce consistent estimations of the Vegetation Optical Depth (VOD) at L-band (L-VOD; Konings et al. (2017); Wigneron et al. (2021)) through radiative transfers.

The main limitations of passive microwave measurements are their very coarse spatial resolution, as the emissivities of the Earth components have a low power and must be aggregated over large areas in order to be monitored by spaceborne instruments. Most of the microwave radiometers have a spatial resolution lower than 25 km, except at the highest frequencies. For SMOS and SMAP, this behavior is particularly critical as the 35-50 km pixels cover a large and often heterogeneous ground terrain, over which a unique of SM and L-VOD is performed. The brightness temperatures are also contaminated by water bodies inside the pixel. For surface water monitoring, only the largest floodplains and rivers (as the Amazon) can be partially monitored using passive microwave sensors.

Side-looking Synthetic Aperture Radars

SAR systems are particularly interesting for hydrological researches. As listed in Table 1.2, they have a spatial resolution ranging from <1 m to a hundred of meters, depending on the acquisition mode. Most of the sensors are able to shift between several modes, to produce either images with very high resolution focused on a local area, or images with lower resolution and a larger spatial coverage. All SAR instruments were widely used to detect flood signatures, as a low backscatter is obtained over smooth open water surfaces, and a high backscatter over flooded vegetation where the tree trunks produce a so-called *double-bounce* effect (see Figure 1.9 and Richards et al. (1987)).

Table 1.2: Summary of the main SAR missions used for surface water remote sensing in the last decades, with their principal features. The data are organized by frequency band and date of launch.

Platform/ Sensor	Frequency band	Resolution	Repeat cycle	Acquisition period
JERS-1	L (1.275 GHz)	18 m	44 days	1992-1998
ALOS/ PALSAR	L (1.27 GHz)	10 m	46 days	2006-2011
ALOS-2/ PALSAR-2	L (1.27 GHz)	3-10 m	14 days	2014-
NISAR	L (1.25 GHz) + S	3-10 m	12 days	2024 (planned)
RADARSAT-1	C (5.3 GHz)	10-100 m	24 days	1995-2013
ENVISAT/ ASAR	C (5.3 GHz)	30-1000 m	35 days	2002-2012
RADARSAT-2	C (5.3 GHz)	3-100 m	24 days	2007-
Sentinel-1A/1B	C (5.405 GHz)	5-20 m	12 days each	2014-
RISAT-1	C (5.35 GHz)	3-50 m	25 days	2012-2016
TerraSAR-X	X (9.6 GHz)	1-16 m	11 days	2007-
TanDEM-X	X (9.6 GHz)	1-16 m	11 days	2010-
COSMO SkyMed 1-4	X (9.6 GHz)	1-100 m	16 days each	2007-

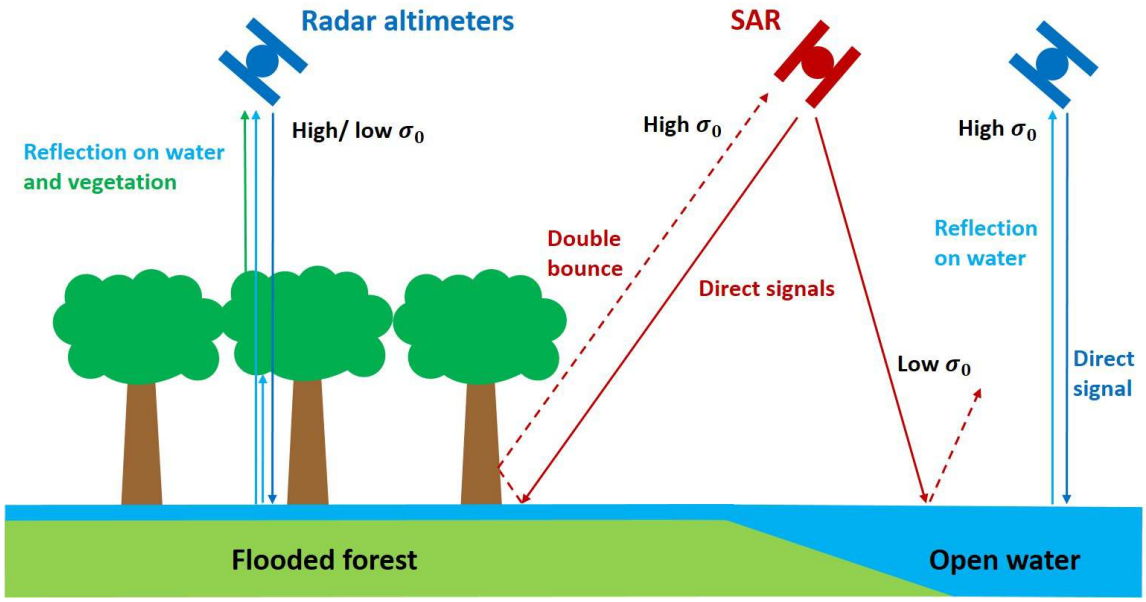

Figure 1.9: Mechanism of backscattering from nadir-looking radar altimeters and side-looking SAR, over open water and flooded forests. Both produce a specular reflection over water surfaces, but the signals can be attenuated, obstructed or reflected by the vegetation. SAR produce low backscattering over open water and high backscattering over the inundated vegetation, due to the side-looking antennas and double bounce effect.

Most of the SARs were limited for hydrological applications due first to a low temporal resolution, with revisit times ranging from two weeks to over a month for the oldest ones (Table 1.2). This is insufficient to sense the rapid changes in SWE following heavy precipitation. It was solved by designing constellations of satellites like the twins TerraSAR-X and TanDEM-X, the four COSMO-SkyMed, or the twins Sentinel-1A and Sentinel-1B. The last ones provide C-band SAR observations with a 6-day combined revisit time and global coverage. Secondly, SAR do not provide long time series of hydrological variables, because they were only active over short periods of time. So, the study of the interannual and decadal variability of the water cycle is based on optical imagery and passive microwave observations. Thirdly, it may be complex to identify water with the SAR backscatter at global scale, as the signatures of geophysical parameters vary across different environments, incidence angles and polarization modes (Musa et al., 2015). And lastly, few data are available at L-band, although it penetrates better the dense forests. The Advanced Land Observing Satellite 2 (ALOS-2) is currently the only platform carrying onboard a L-band SAR, namely Phased Array L-band SAR 2 (PALSAR-2).

Nadir-looking radar altimeters

The radar altimeters emit a signal in the direction of nadir (mostly at Ku band, see Table 1.3), which is strongly reflected by water surfaces. A high backscatter is therefore

observed over surface water (Fatras et al., 2015; Frappart et al., 2021a; Abdalla et al., 2021), while the backscatter of side-looking SAR is low, except over flooded vegetation due to the double bounce effect on natural corner reflectors (see Figure 1.9 and Richards et al. (1987)). Radar altimeters only acquire one waveform for each beam pulse, making these observations sampled along the ground tracks of the platform. The distribution of the observations on the Earth surface depends on the design of the sensor. The less distance there is between two ground tracks, the longer is the revisit time of the satellite. Depending on the missions, the compromise between spatial and temporal samplings results in altimetry tracks spaced from tens to hundreds of kilometers at the equator, with revisit times ranging from 10 to 35 days (Table 1.3).

Due to their poor spatial coverage, the radar altimeters are not operational for monitoring SWE, although they are capable of detecting both open water and floodplains below the canopy (Frappart et al., 2021b). On land, they are mostly used for monitoring time series of water levels over rivers, lakes and recently floodplains (Crétaux et al., 2011; Cretaux et al., 2017). Combined to SWE, these water levels are then used to retrieve the volume of freshwater stocks (or surface water storage) in river basins. They represent therefore a very useful complementary information for deriving SWS using multi-mission information (see Section 1.3.4.1).

1.3.3.2 Microwave-based surface water detection

Mapping SWE with SAR

Table 1.3: Summary of the main radar altimeters used for remote sensing of the water cycle in the last decades, with their principal features. The data are organized by date of launch.

Platform/ Sensor	Frequency band	Ground track separation	Time sampling	Acquisition period
ERS-1/ RA	Ku (13.8 GHz)	80 km	35 days	1991-2000
TOPEX/Poseidon, SSALT	Ku (13.65 GHz)	315 km	9.9156 days	1992-2006
ERS-2/ RA	Ku (13.8 GHz)	80 km	35 days	1995-2011
Jason-1/ Poseidon-2	Ku+C (13.6/ 5.3 GHz)	315 km	9.9156 days	2001-2013
ENVISAT/ RA-2	Ku+S (13.6/ 3.2 GHz)	80 km	35 days	2002-2012
Jason-2/ Poseidon-2	Ku+C (13.6/ 5.3 GHz)	315 km	9.9156 days	2008-2019
HY-2A+B+C+D	Ku+C (13.6/ 5.3 GHz)	-	10-14 days	2011-
SARAL/ AltiKa	Ka (35.75 GHz)	80 km	35 days	2013-2016
Jason-3/ Poseidon-2	Ku+C (13.6/ 5.3 GHz)	315 km	9.9156 days	2016-
Sentinel-3A+3B/ SRAL	Ku+C (13.6/ 5.4 GHz)	104 km	27 days	2016-
Jason-CS/ Sentinel-6	Ku+C (13.6/ 5.3 GHz)	315 km	9.9156 days	2020-
SWOT	Ka (inSAR), Ku, C (nadir)	Two 50 km swaths 10 km separation	21 days	2022 (planned)

The SAR missions listed in Table 1.2 support a wide range of high-resolution operational flood detection methods, in particular using automated image segmentation and classification techniques. First, the twin satellites TerraSAR-X and TanDEM-X were used to produce maps of surface water. During the generation of the high-resolution (12 m) and high-precision (∼2 m) Digital Elevation Model WorldDEM, data from the TanDEM-X mission (encompassing both satellites) were used to produce a reference water mask (Martinis et al., 2015b). More interesting, the field of near real time (NRT) flood monitoring has been widely investigated to support mitigation policies and decision-making. It has shown good capabilities using TerraSAR-X data resulting in operational flood detection services available online (Martinis et al., 2009, 2015a).

The Canadian C-band SAR RADARSAT and RADARSAT-2 were also used to monitor large inundations (Bolanos et al., 2016; Nakmuenwai et al., 2017). Additionally, some interesting results for the mapping of rapid flood events were obtained using the X-band COSMO SkyMed (COnstellation of small Satellites for the Mediterranean basin Observation) satellites (Pulvirenti et al., 2011; Refice et al., 2014). The constellation flights four satellites allowing a short revisit time over specific scenes (up to 12 hours) and a very high-resolution (1 m in spotlight mode), which makes it suitable for NRT flood detection (Pierdicca et al., 2013). The Advanced SAR (ASAR) onboard ENVISAT was also used to produce a mask of permanent water bodies using multi-year data from 2005 to 2012 (Santoro et al., 2015), and in a NRT automatic extraction of water surfaces (Westerhoff et al., 2013).

A first drawback of these SAR systems is revealed when considering the literature. Although most of the SAR perfectly fit the needs of NRT inundation mapping, they are unable to provide an operational mapping of global or large regional areas over long timescales, as data are not available with sufficient temporal resolution. Some sensors also need to be activated or programmed to acquire several scenes on a specific target.The first continuous, global and high-resolution capability for mapping floods with SAR data and high time sampling is provided by the European Space Agency's (ESA) Sentinel-1 program. The two C-band SARs onboard Sentinel-1A & 1B satellites permit a complete overview of land surfaces within 6 days, and data are freely available. The characteristics of Sentinel-1 data perfectly fit the NRT inundation detection (Twele et al., 2016). They were used in numerous regional studies for mapping the spatial and temporal variations of floods (Bioresita et al., 2018; Uddin et al., 2019; Singha et al., 2020; DeVries et al., 2020). More interestingly, global SAR-based flood products are now possible. The Copernicus Emergency Management Service (CEMS) recently implemented in its Global Flood Awareness System (GloFAS) a NRT Global Flood Monitoring (GFM) product, that is directly updated through the analysis of new incoming Sentinel-1 images (Salamon et al., 2021). Three different and independent algorithms are used (Twele et al., 2016; Chini et al., 2017; Bauer-Marschallinger et al., 2022), to produce three independent water masks on the same scene. The decision to classify a pixel as land or water is performed at the majority of the ensemble algorithms (at least two agreeing). This combines the advantages of each algorithm for the detection of water, while adding robustness to avoid omission/commission errors in complex terrains where individual estimates can be wrong (arid soils, snow and ice, shadows).

Although C-band and X-band SAR data have shown significant performances in mapping SWE, and despite a growing capability for global analysis, some problematic features remain. A major issue is the attenuation of the signals by vegetation layers affecting particularly the possibility of mapping flooded vegetation in the tropical wetlands. L-band signals penetrate better the dense forests than other microwave frequency bands (Figure 1.8). As a consequence, hydrological sciences need L-band sensors to map flooded areas below the canopy, mainly in large tropical basins as the Amazon and Congo. Passive microwave observations can be used to map surface water (Prigent et al., 2007, 2020; Parrens et al., 2017) but they have a coarse spatial resolution. On the contrary, L-band SAR achieves high-resolution requirement and penetrates well the vegetation, but very few data are available for analysis. The only corresponding missions were JERS-1 (Japanese Earth Resources Satellite, from 1992 to 1998), ALOS/PALSAR (2006-2011) and ALOS/PALSAR-2 (from 2014), all from the Japan Aerospace Exploration Agency (JAXA).

Large data gaps therefore exist for L-band SAR observations. Still, they were used to produce several maps of permanent and seasonal water in tropical river basins with consistent results. In the Amazon, JERS-1 mosaics acquired for the Global Rain Forest Mapping (GRFM) project were used to classify wetlands according to the vegetation type and inundation state (Hess et al., 2003, 2015). More interestingly, two maps were produce: one at low water stage in October-November 1995, and the second at high water stage in May-June 1996. They are shown in Figure 1.10a and Figure 1.10b, respectively. The flooded forests appear in white and are mainly located along the streams of the Amazon and its tributaries, and in the floodplains close to the Rio Negro in the north of the basin. The flooded herbaceous

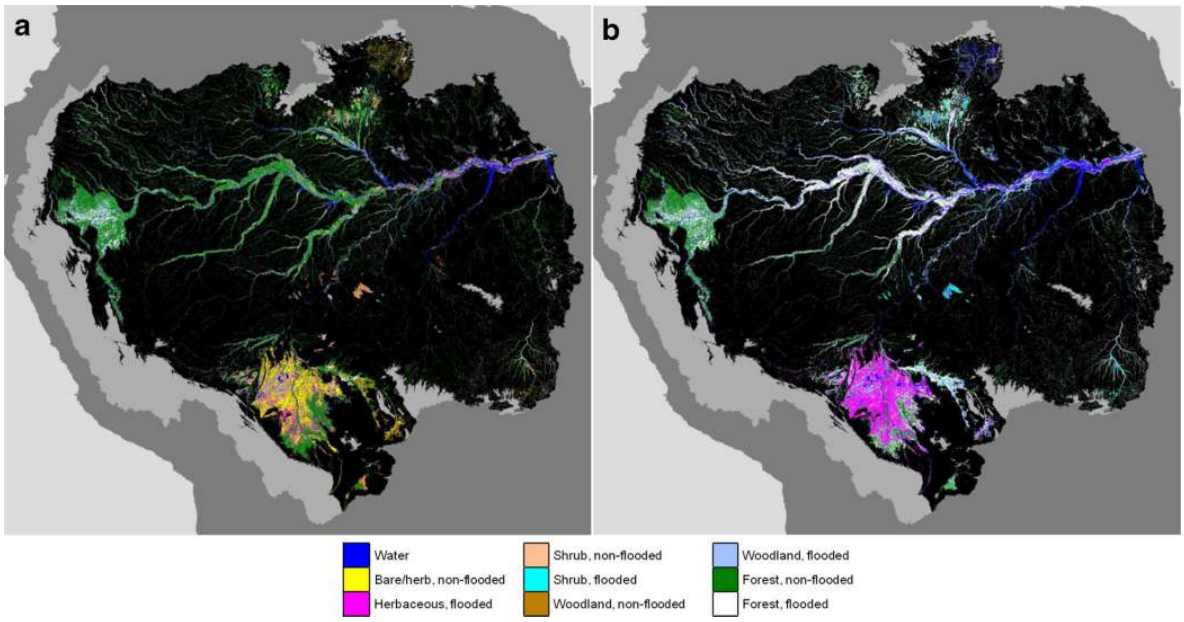

Figure 1.10: Dual-season classification of inundations in the Amazon basin based on JERS-1 L-band SAR mosaics, adapted from Hess et al. (2015). **(a)** Low water stage in October-November 1995, and **(b)** high water stage in May-June 1996.

appear in magenta, and are mainly located in the Llanos de Mojos in Bolivia (south of the Amazon basin). Both flooded forests and herbaceous show much more inundated areas at high water stage (right panel in Figure 1.10). These maps are distributed and due to their high-resolution and quality, still represent reference flood products for the Amazon basin. In the Cuvette Centrale of Congo, ALOS/PALSAR data were combined with MODIS-derived vegetation indexes to map the permanent and seasonal extent of water surfaces for different vegetation types (Betbeder et al., 2014). These studies prove the interest of L-band sensors for mapping flooded vegetation, but are not extended globally or on long timescales.

Although numerous studies have been recorded on mapping flooded vegetation (Tsyganskaya et al., 2018), consistent estimations based on SAR data exclusively are highly uncertain for now. The solutions were found in multi-mission products, merging active radar datasets together, with passive microwave observations, or with optical datasets.

Multi-mission SWE products

The combination of SAR and optical datasets allows to map floods at high spatial resolution, while improving temporal coverage and overall accuracy of the products (Tong et al., 2018). In particular, Martinis et al. (2022) combined Sentinel-1 and Sentinel-2 flood mapping approaches to produce over two consecutive years (2019 and 2020) a map of permanent water, and monthly seasonal water layers. This work is based on Sentinel-2 flood detection, with the input of Sentinel-1 SAR data to complement when the optical sensor have issues. It shows that during months with cloud cover, SAR data regularly take over optical data, with similar performances in detecting surface water. The extraction of permanent and seasonal water is also important, despite most of the masks being static or presenting a binary water against land classification approach. The seasonal layers compare well with the 32-year GSW monthly surface water based on Landsat, from Pekel et al. (2016). While this study addresses the need of high-resolution mapping systems for monitoring future flood disasters, it does not implement methods for retrieving flooded vegetation in the tropics. Moreover, it does not provide insights into interannual variations of SWE and SWS required for hydrological and biogeochemical models (Section 1.1 and Section 1.2).

The need for long time series of SWE was partially filled by combining passive and active microwave products, *i.e.* the Global Inundation Extent from Multiple Satellites (GIEMS, Prigent et al. (2007, 2020) and the Surface WAter Microwave Product Series (SWAMPS, Schroeder et al. (2015); Jensen and Mcdonald (2019)). Both supply water fraction estimations at 0.25° spatial resolution and use mostly the same inputs. These include the brightness temperatures from SSM/I and SSMIS at several wavelengths/ polarizations, the backscattering coefficients from ERS (and other scatterometers for SWAMPS), Normalized Difference Vegetation Index (NDVI) from the Advanced Very High Resolution Radiometer (AVHRR, plus MODIS for SWAMPS), meteorological data describing atmospheric variables, and a snow cover flag. The main differences come from two distinct retrieval strategies. GIEMS and its most recent version, GIEMS-2 (Prigent et al., 2020) are 3-step algorithms. First, the brightness temperatures at 19 and 37 GHz from the radiometers are calibrated into surface

emissivities. This was first performed by radiative transfer (GIEMS-1), and now GIEMS-2 uses a neural network trained against previous emissivity estimations from GIEMS-1. Then, the monthly climatology of the emissivities and ERS backscatter time series are computed and clustered to identify the water pixels. Finally, the water fraction in these pixels are computed using end-members. In a simple framework with two end-members (dry and water), the fraction of water (WF) can be estimated as (Sippel et al., 1994; Fily et al., 2003):

$$WF = \frac{E_p - E_{dry}}{E_{water} - E_{dry}} \tag{1.5}$$

where E_p is the emissivity at polarization p, E_{dry} is the emissivity of dry land, and E_{water} is the emissivity of water. For GIEMS-2, a linear relation is directly established between the top 20% and the bottom 20% emissivities, corresponding respectively to a 50% and a 1% Water Fraction (Prigent et al., 2020). The ERS backscatter and NDVI time series are also used in a multi-linear regression to correct the vegetation effect on microwave emissivities. The retrieval methodology of SWAMPS differs as it uses the Microwave Polarization Difference Index (MPDI) of SSM/I and SSMIS brightness temperatures at 19 GHz, instead of the emissivities calculated through radiative transfer. The fractions of water are directly estimated from the MPDI using vegetation and water-dependent end-members to calibrate the retrieval.

The first version of SWAMPS was subject to numerous issues, as listed in Pham-Duc et al. (2017). A strong contamination by the oceans and large erroneous estimations over arid regions were noted, and both corrected in the latest version of SWAMPS (Jensen and Mcdonald, 2019). However, significant problems still remain unsolved. SWAMPS presents

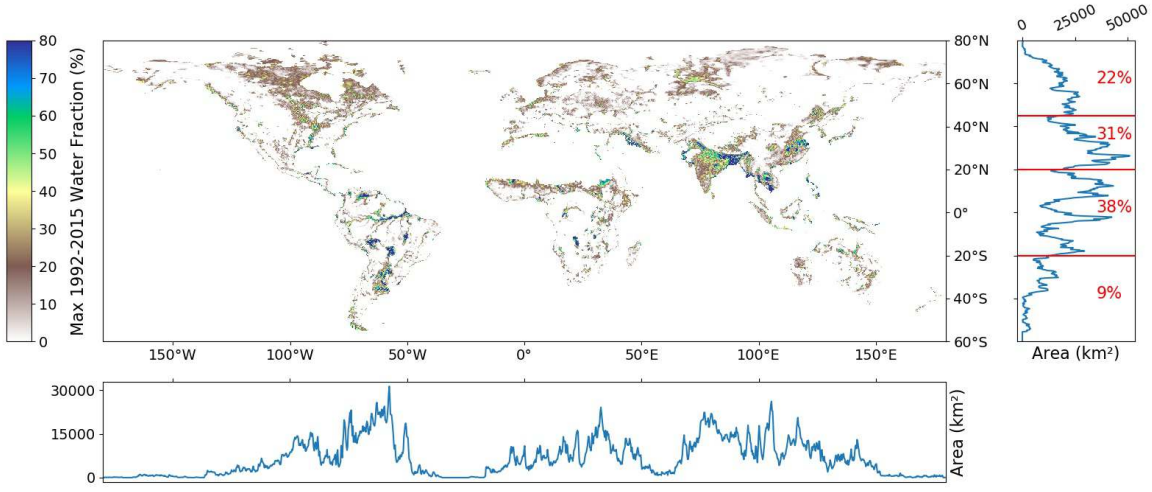

Figure 1.11: Maximum water fraction per pixel in the 1992-2015 records of GIEMS-2 (Prigent et al., 2020). The latitudinal and longitudinal profiles of SWE in km^2 are added at the right and below the world map, respectively. For the latitudinal profile, the percentage of total SWE contained in latitude bins are shown in red. The zonal separation are made with latitude thresholds of 20°S, 20°N and 45°N.

a very lower seasonal amplitude than GIEMS and other regional products. Small water fractions are observed over regions with no surface water, whereas the large floodplains show a lower signal than GIEMS, due to a wrong calibration of the retrieval algorithm. GIEMS-2 estimates are better in seasonality than SWAMPS and compare well with several regional estimations of SWE, including the JERS-1 classification over the Amazon basin from Hess et al. (2003), presented in Figure 1.10). It was also validated against other hydrological products like precipitation, river discharges and water levels. It can therefore be considered as the reference dataset for interannual and decadal estimations of surface water extent over the globe, including consistent SWE retrieval in the tropics.

Figure 1.11 presents a map of the maximum water fraction (in percentage) contained in the 1992-2015 GIEMS-2 records, as well as the latitudinal and longitudinal profiles of the corresponding area in km^2. It highlights a strong contribution of tropical and subtropical wetlands to the maximum water extent, dominated by large wetlands in South America, seasonal floods from the Sahel and irrigated croplands in Southeast Asia. It can be compared to Figure 1.7 for the maximum SWE in Landsat-based 1984-2020 GSW records, where the tropical component is almost absent. Both datasets describe the northern boreal contribution to the total water extent, which corresponds to the immense majority of SWE in GSW product. The large flood signal from tropical and subtropical wetlands in GIEMS-2 ($\sim$69% of total SWE between 20°S and 45°N) highlights the detection of water below forest canopies, in flooded savannas and paddy fields, which lack in optical-based estimations.

Still, GIEMS-2 owns several limitations. First, SSM/I and SSMIS Ku-band brightness temperatures are impacted by the vegetation. It is unclear whether the methodology based on NDVI and ERS backscatter to correct the vegetation effect on SSM/I and SSMIS brightness temperatures produces exact results. Few validation datasets exist over tropical wetlands, but GIEMS was validated in the Amazon against the JERS-1 classification from Hess et al. (2003) (see Figure 1.10) and showed consistent values for both amplitude and seasonality (Prigent et al., 2007). Then, an overestimation of flooded areas was noted under wet regimes, due to a confusion between SM and surface water (see Section 1.4.2). Finally, the very coarse spatial resolution at 0.25° is not sufficient to describe hydrological regimes with enough precision, which propagates into wetland methane emissions (Section 1.2.2) and land carbon sink (Section 1.2.1) uncertainties. Still, GIEMS-2 represents the most reliable dataset to study long-term SWE dynamics and derive SWS changes over time.

1.3.4 Determination of Surface Water Storage

The monitoring of SWS generally relies on independent estimations of SWE and water levels. Several approaches were used to retrieve the water volume anomalies. The most direct transforms SWE maps from GIEMS or other references to water level maps with the same spatial coverage and resolution, using radar altimetry observations. A more sophisticated way is the hypsographic approach, which uses a high-resolution DEM to derive time series of water volumes in every pixel. SAR (in particular at L-band) were also used to monitor directly the water levels, through the Interferometric SAR (InSAR) technique. An exhaustive review of

techniques for monitoring SWS over wetlands can be found in Papa and Frappart (2021).

1.3.4.1 Using radar altimetry and SWE

High-precision radar altimeters have been available for the last three decades, since the 1990s and the launch of Topex/Poseidon (T/P). These missions and their characteristics are listed in Table 1.3. Although they were first designed to monitor the sea surface height, they have also provided altimetry data over the continents. These data have been used to derive water levels on lakes, rivers and floodplains. These water levels are estimated at the intersection between altimetry ground tracks and water bodies, called *virtual stations* (VS). Several databases such as Hydroweb (Crétaux et al., 2011; Hydroweb) or DAHITI (Schwatke et al., 2015; DAHITI) now provide mono- and multi-mission (up to 30 years) altimetry-based time series of water levels. The spatial and temporal behavior of altimetry tracks vary, with revisit times ranging from ∼10 days (T/P and Jason 1/2/3/CS, on the same orbit) to 35 days (ERS, ENVISAT and SARAL on the same orbit too), and ground track separated by 80 km (ENVISAT orbit) to 315 km (Jason orbit) at the equator (Table 1.3). The larger is the ground track separation at the equator, the better is the time sampling.

In the 2000s, radar altimetry data were combined to wetland maps to estimate the volume of water stocked in rivers and floodplains. The first study used T/P and the dual-season flood extent maps in the Amazon derived from JERS-1, to retrieve the maximum and minimum water stocks in the Rio Negro sub-basin during year 1995-1996 (Frappart et al., 2005). A map of water levels is there computed over the flood extent by interpolating the water heights at VS, and is further used to derive water volumes over one hydrological year. A similar methodology was then applied to SWS retrieval in the Mekong basin for 1998-2003 (Frappart et al., 2006), using a combination of radar altimeters (T/P, ERS-2 and ENVISAT) and flood extent derived from the Satellite pour l'Observation de la Terre–Végétation (SPOT-VGT). It improved our regional understanding of the seasonal and interannual variations of SWS, but relied on visible and infrared observations.

This methodology is well-suited for retrieving SWS at the regional scale, but long-term determination of SWS changes relies on the availability of input datasets. Multi-mission radar altimetry time series can now be estimated for the last 30 years (Kitambo et al., 2022a). Also, GIEMS-2 provides monthly SWE estimations from 1992 to 2015, and up to now upon request. The combination of these two datasets was therefore logical, and was tested successfully in several distinct environments: the Rio Negro in the Amazon basin (Frappart et al., 2008, 2011), the entire Amazon basin itself (Frappart et al., 2012; Tourian et al., 2018), the Ob basin in boreal regions (Frappart et al., 2011), the Orinoco basin (Frappart et al., 2014), the Ganges-Brahmaputra basin (Papa et al., 2015) and the Congo basin (Becker et al., 2018). These studies cover both flooded vegetation in the major tropical wetlands, and large irrigated croplands in Southeast Asia. This methodology is however limited by it low level of detail, due to the coarse spatial resolution of GIEMS SWE estimations.

SWE estimations at higher spatial resolution were obtained using vegetation and water

indexes from the 8-day composite product of MODIS (Sakamoto et al., 2007). It provides data from 2000, which can be combined with radar altimetry to derive SWS in regions with low or sparse vegetation. Monthly maps of SWS were therefore generated over the Tonle Sap in Cambodia (Frappart et al., 2018)) and the whole Lower Mekong Basin (LMB; Pham-Duc et al. (2019)). Similar maps were generated over a totally different environment, the Mackenzie Delta in northern Canada (Normandin et al., 2018). This approach also permitted to quantify the SWS variations over two decades (2000-2019) in the Lake Chad and surrounding floodplains (Pham-Duc et al., 2020). This approach is limited to low vegetated areas because of MODIS spectral properties, and is therefore not applicable to tropical wetland hydrology.

1.3.4.2 Hypsographic curves

The idea of the *hypsographic curve* approach is to allocate the fraction of water estimated in a 0.25° pixel by GIEMS (or SWAMPS) into high-resolution sub-pixels extracted from a DEM. Figure 1.12 helps to understand the hypsographic relation. The plots in the first line (Figure 1.12a-c) present a GIEMS pixel over a river stream with high topography around, while the plots in the second line (Figure 1.12d-f) present a pixel with low elevation changes, undergoing large floods at high water stage. For each 0.25° pixel, the cumulative distribution of the elevation (*i.e.* the *hypsographic curve*, Figure 1.12b,e) is extracted from the DEM sub-pixels. It can be further expressed as a curve of the elevation depending on the cumulative

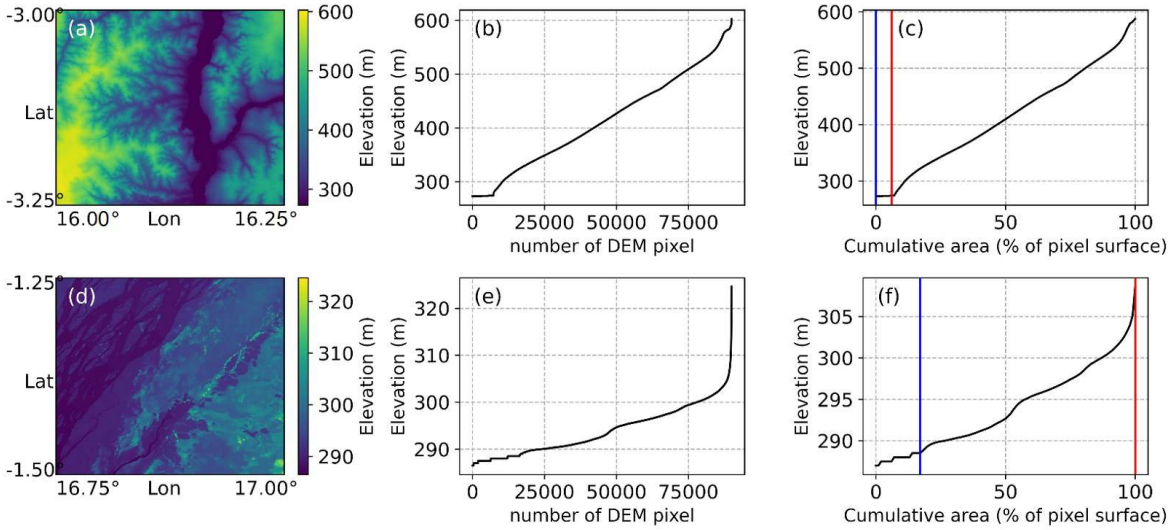

Figure 1.12: Hypsometric curve approach described for two GIEMS pixels in the Cuvette Centrale of Congo **(d,e,f)** and in mountainous areas in the south of the Cuvette Centrale **(a,b,c)**. **(a,d)** GIEMS pixels filled with the 90 m MERIT DEM, **(b,e)** cumulative distribution of the DEM elevation, and **(c,f)** same with the distribution transformed as cumulative area in the pixels. This explains the relation between the water fraction in every GIEMS pixel and the water levels, further derived as water volumes. The blue and red vertical lines show the minimum and maximum water fractions in GIEMS-2 records, respectively. Adapted from Kitambo et al. (2022b).

area (Figure 1.12c-f). Now for each GIEMS water fraction (with the long-term minimum and maximum shown in blue and red, respectively), the corresponding elevation of the water surface can be estimated. All the DEM sub-pixels with en elevation lower than this threshold are considered as flooded, and the water volume is further derived. This area-elevation-volume relation was assessed to retrieve the variations of both lake area and volume, based on satellite altimetry data (Tortini et al., 2020).

Over rivers and wetlands, the hypsographic curve approach was used to retrieve long-term SWS variations in the Amazon and Ganges-Brahmaputra basins, based on GIEMS and high-resolution DEMs (Papa et al., 2015; Salameh et al., 2017). A 25 years time series of SWS in the Congo basin was recently derived from GIEMS and several high-resolution DEMs, and compared to the traditional estimates based on GIEMS and radar altimetry (Kitambo et al., 2022b). The hypsometric curve approach shows a similar amplitude in the seasonal SWS signal. Its main advantages are: i) it is not based on the interpolation of water levels time series from altimetry VS which are poorly sampled, and ii) it provides a finer representation of the flood dynamics inside the large GIEMS pixels, due to the allocation of surface water to the lowest elevation areas.

The main uncertainties of the hypsometric approach are linked to the input products of SWE and elevation. While the SWE estimates need further improvements (see Section 1.3.3), high-resolution (30 m to 90 m) DEMs also present biases and uncertainties, in particular in densely vegetated areas. The MERIT DEM (Multi Error-Removed Improved-Terrain) at 90 m, derived from the Shuttle Radar Topography Mission (SRTM), and the ALOS World 3D at 30 m (AW3D30) recently released, have improved the performances of DEMs for hydrological applications. They are particularly needed for resolving channel connectivity and floodplain discharge / recharge in hydrological models (Yamazaki et al., 2011), and can be used efficiently for SWS retrieval (Kitambo et al., 2022b).

Finally, the hypsometric curve approach could be extended worldwide, as it only relies on a static DEM and SWE estimates available from 1992 to now with GIEMS-2. It would provide long-term time series of SWS for the needs of the hydrological and biogeochemical communities, whose performances and limitations should be further evaluated.

1.3.4.3 InSAR water level monitoring

The SAR Interferometry (InSAR) is a technique for determining the phase difference between two or more SAR images, acquired with different positions or at different times (Bamler and Hartl, 1998). It is commonly used to retrieve the topography with high accuracy, and therefore to produce DEMs. The surface height changes can also be derived from the interferometric phase difference. For water surfaces, it is used to compute changes in water level between two acquisition times (Alsdorf et al., 2001b). It is particularly useful in vegetated wetlands (Wdowinski et al., 2008) due to the double-bounce effect (Figure 1.9). However, it requires a frequent revisit especially at C-band, as the changes in phase are related to the wavelength of the electromagnetic wave. It is also not suitable over open water due to both the low

backscatter, and the loss of coherence (or decorrelation) between two consecutive SAR images (Gallant, 2015).

Based on the SAR sensors listed in Table 1.2, InSAR performances over wetlands have been widely investigated (Alsdorf et al., 2000, 2001a,b; Hong et al., 2010). Low coherence was observed over open water and higher over flooded vegetation, particularly for L-band SAR observations. InSAR-based monitoring of water levels and water stocks in tropical wetlands provides interesting although limited results. Several L-band JERS-1 SAR images were used to map the surface water changes and flood waves in a subset of the Amazon basin in the 1990s, and compared well with results from in-situ gauges (Alsdorf et al., 2007). A similar approach was used with ALOS/PALSAR and ALOS-2/PALSAR-2 data acquisitions from 2010 to 2015 to derive water levels, still in the Amazon, and presented good agreement with radar altimetry time series (Cao et al., 2018). The Congo basin also presents flooded forests adapted for InSAR water level retrieval. Water levels from PALSAR in the Cuvette Centrale of Congo were validated using ENVISAT time series (Lee et al., 2015). InSAR with PALSAR and ENVISAT were then used to retrieve SWS time series from 2002 to 2011 in the Cuvette Centrale (Yuan et al., 2017).

Finally, the KaRin instrument (Fjørtoft et al., 2013) onboard the future Surface Water and Ocean Topography (SWOT) mission will perform InSAR measurements to provide a global-scale monitoring of SWS with 21 days revisit (Biancamaria et al., 2016). Unlike current inSAR applications relying on the repeat pass interferometric configuration, KaRin will perform simultaneous measurements with two antennas separated by 10 m. It expected to show a ∼10 dB radiometric contrast between land and water, and a high interferometric coherence on water.

1.4 Limitations of the state of the art

Several techniques to retrieve SWE and SWS, two essential hydrological variables, were presented in Section 1.3.1, Section 1.3.2, Section 1.3.3.2 and Section 1.3.4. In this section, we discuss the general limitations of these sensors, techniques and products. The future advances expected in hydrological sciences due to new remote sensing tools and datasets are discussed in a following section.

1.4.1 Effect of the vegetation

1.4.1.1 Problems with optical datasets

Visible and infrared signals are reflected by the canopy and are not able to sense the water surface below vegetation layers. Moreover, they are not able to penetrate cloud cover which is frequent over tropical and polar regions. This has two consequences for surface water datasets from Landsat, MODIS and other optical sensors: i) they are frequently aggregated by time

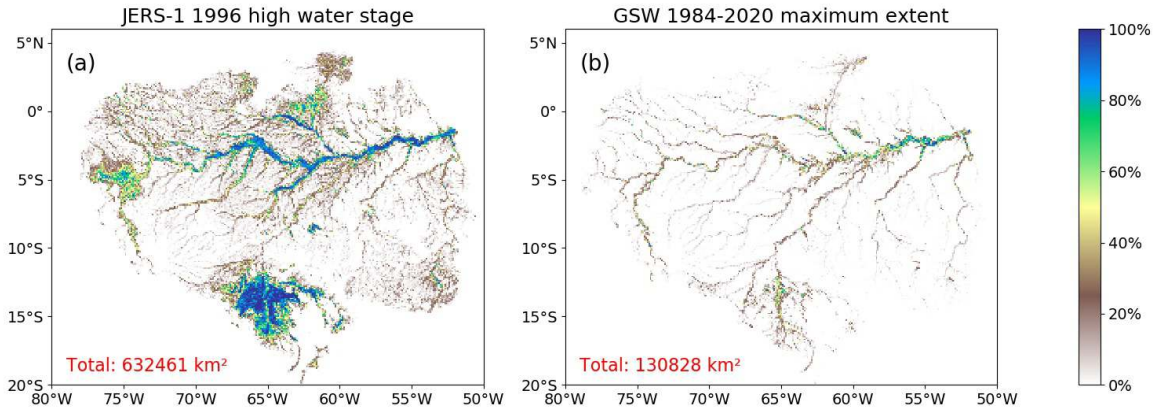

Figure 1.13: Maximum surface water extent from optical and microwave datasets. (a) SWE at high water stage in spring 1996 from Hess et al. (2003) based on JERS-1 L-band SAR, and (b) maximum SWE in the 1984-2020 GSW records (Pekel et al., 2016). Both datasets are gridded at 0.1° and the colorbar corresponds to the water fraction in every pixel. The maximum inundated surfaces in km^2 are shown in the bottom left corner of both figures.

to overcome weather issues, and ii) they do not detect water below the canopy in flooded forests. This largely limits their interest for studying tropical wetland hydrology.

Figure 1.13 shows the surface water extent in the Amazon basin from two datasets: the classification at high water stage of JERS-1 data in 1996 (Hess et al., 2003), and the 1984-2020 maximum extent in the Landsat-based GSW product (Pekel et al., 2016). The JERS-1 map is obtained only for year 1996 and shows very large inundated areas in Llanos de Mojos and Rio Branco floodplains, and along the stream of the Amazon and its main tributaries. On the contrary, the GSW product poorly detects these large inundated areas although it represents the maximum estimate for the entire 1984-2020 period. This example highlights the biases affecting optical-based flood products in tropical regions.

In Figure 1.13, the maximum inundated area is printed in red for both datasets. The total flooded surface reaches ∼632461 km^2 for the JERS-1 dataset from Hess et al. (2003), almost five times greater than the value from the GSW product. At global scale, the total inundated area reaches ∼6 millions km^2 for GIEMS, which is nearly twice more than optical datasets. There is a large difference in the zonal distribution of surface water between GIEMS (high fraction over the tropics, Figure 1.11) and GSW (low fraction over the tropics and boreal regions dominate, Figure 1.7). GIEMS estimates are not fully trustful, but both amplitude and seasonality were validated in the Amazon against the dual-season JERS-1 classification from Hess et al. (2003). Therefore, there are evidences that GSW and other optical-based flood products underestimate the global flooded areas, due to large underestimations of SWE in tropical wetlands and floodplains.

1.4.1.2 Influence of the vegetation on the microwave signals

The microwave signals are very useful to overcome the limitations of optical sensors in the tropics, as they work with cloud cover and by night. They are also able to penetrate the canopies, although the signals are partially attenuated by vegetation layers. This attenuation is lower at L-band, when compared to the lower wavelengths/higher frequencies X and C bands (see Figure 1.8). It is represented by the transmissivity of the vegetation (γ):

$$\gamma = \exp\left(\frac{-\tau}{cos(\theta)}\right) \tag{1.6}$$

where θ is the incidence angle of the signal, and τ is the Vegetation Optical Depth (VOD) at specific microwave frequency. L-band VOD (L-VOD) has been evaluated globally using SMOS (Wigneron et al., 2021) and SMAP (Konings et al., 2017) records. The transmissivity of the vegetation and therefore the VOD are used in models, for calibrating both passive microwave emissivities and active backscattering from soil and vegetation layers (Frappart et al., 2020). As it is an exponential term, a very low γ is found in dense forests (high τ) or at high incidence angles (because the signal propagates into a much larger distance in the vegetation layers). For active radar signals, the backscattering is described using the water cloud model (Attema and Ulaby, 1978):

$$\sigma^0 = \gamma^2\sigma^0_{soil} + \sigma^0_{vege} + \sigma^0_{soil+vege} \tag{1.7}$$

where $\sigma^0_{soil+vege}$ is the backscattering due to the interactions between the soil and the vegetation and is generally neglected. σ^0_{vege} corresponds to the volume scattering into the vegetation, while σ^0_{soil} corresponds to the backscattering from the soil, with the wave crossing twice vegetation layers and therefore attenuated by a factor γ^2. Similarly, microwave emissivities encompass a component from the soil and a component from the vegetation (e_{soil} and e_{vege}, respectively).

The attenuation of microwave signals by the vegetation has two consequences. First, L-band signals propagate better through the canopy and should be used for monitoring soil properties in forested areas. Secondly, the backscattering from dense forests must be corrected from an attenuation factor, *i.e.* the transmissivity of the vegetation. A complete parameterization of the vegetation in physical models for both passive and active microwave observations can be found in Frappart et al. (2020).

1.4.2 Confusion between soil moisture and surface water

Soil moisture (SM) is another disturbing geophysical parameter for surface water monitoring. SM enhances the reflection of microwave signals on land surfaces and decreases their emissivity, so the brightness temperatures observed by radiometers. A near-saturated soil can comport approximately like surface water for both passive and active radars. As a consequence, SM is a pitfall for flood detection algorithms.

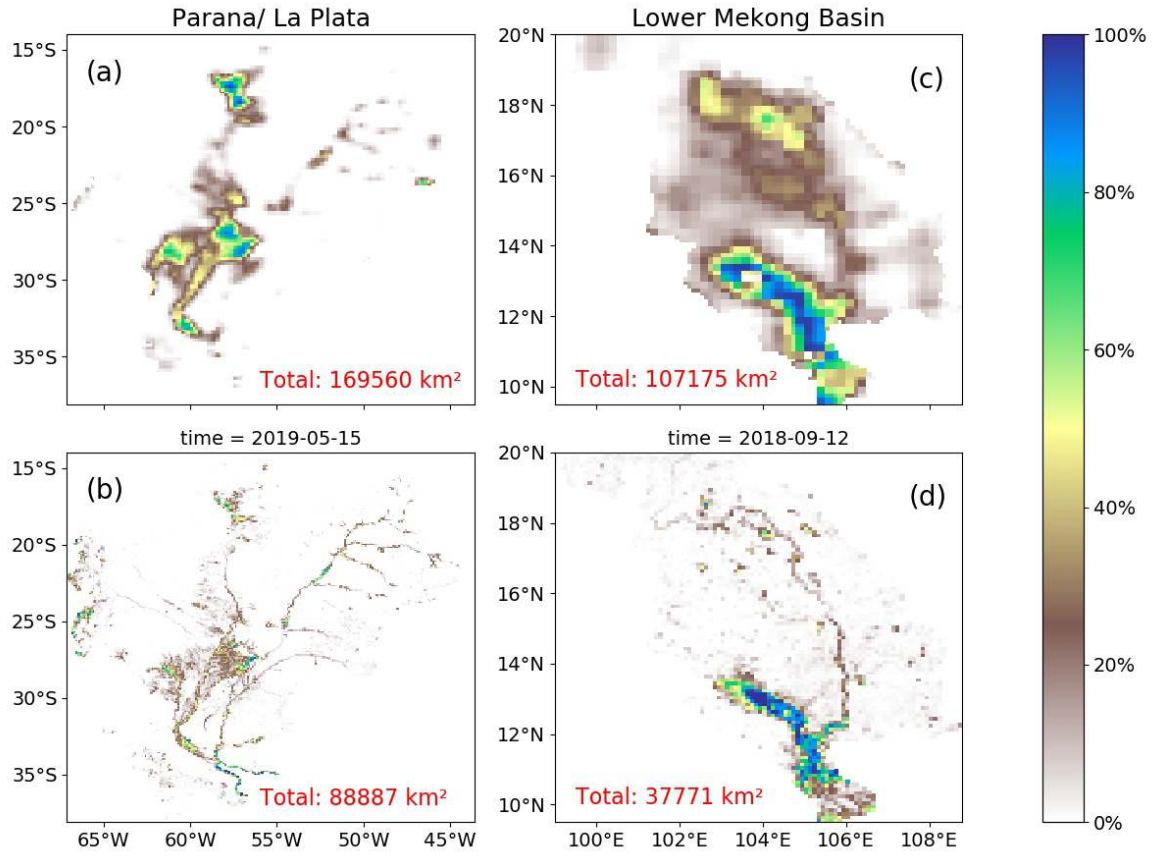

Figure 1.14: Comparison of SWE extracted from GIEMS **(a,c)** and MODIS **(b,d)** at maximum flood extent in 2018-2019, in: **(a,b)** the Parana basin and **(c,d)** the Lower Mekong Basin (LMB). MODIS dataset was gridded at 0.1° and the colorbar indicates the percentage of water in each pixel. The maximum inundated surfaces in km^2 are shown down each figure.

The confusion between surface water and near-saturated soils during wet seasons in global multi-satellite flood products is highlighted in Figure 1.14. The maximum water fraction in 2018-2019 in the Parana and Lower Mekong Basin (LMB) are shown for both GIEMS (Figure 1.14a,c) and MODIS (Figure 1.14b,d) SWE estimates. The MODIS dataset is obtained as in Frappart et al. (2018) and Normandin et al. (2018). The spatial patterns between GIEMS and MODIS SWE estimations over the Parana (Figure 1.14a-b) and LMB (Figure 1.14c-d) are very consistent. However, GIEMS presents two limitations based on these plots: i) a much lower level of detail than MODIS SWE due to lower spatial resolution (∼25 km against 500 m), and ii) twice more estimated flooded area than MODIS.

It is to be noted than although MODIS SWE is considered as the reference in Figure 1.14, it may underestimate floods in savannas and forests due to the spectral properties of the sensor. Still, it is clear that GIEMS overestimates floods at the higher water stage: almost all the pixels in the LMB contain a fraction of water (Figure 1.14c). Most of these regions likely present near-saturated soils but not surface water during rainfall events. These issues in

GIEMS have been also noted in other monsoon regimes, in the Ganges-Brahmaputra plains and the Sahel region. Finally, all low-resolution microwave sensors are subjected to a confusion between SM and surface water. In particular, this is a common issue in spaceborne GNSS-Reflectometry (see Chapter 2), and L-band radiometers (SMAP and SMOS) have troubles in disentangling the relative contributions of both variables in the large 35-50 km pixels.

1.4.3 Spatial and temporal sampling of microwave sensors

The global SWE products, *i.e.* GIEMS and SWAMPS, are based on SSM/I and SSMIS brightness temperatures and have a coarse spatial resolution of ∼25 km at the equator. Many other passive microwave sensors are available with a spatial resolution lower than ∼25 km, including L-band SMAP and SMOS with a resolution of 35-50 km (Table 1.1). Overall, microwave radiometers are very sensitive to soil moisture and surface water, even in the tropics, but are not able to sense small-scale spatial features, floodplain connectivity to rivers, tight streams, etc. Figure 1.14 highlights the poor level of detail in GIEMS SWE estimations.

High-resolution (<30 m) observations are available with optical and radar imagery. Because visible and infrared wavelengths are limited in the tropics, SAR are the only operational systems capable of sensing high-resolution surface water dynamics in vegetated areas. However, for many years, the revisit time of SAR was quite low (>20 days, see Table 1.2). This is problematic for detecting flash floods and accurately specifying inundation dynamics. Recent SAR missions with two (TanDEM-X) or four (COSMO SkyMed) satellites were designed to improve the spatial (∼1 m) and temporal (down to one day) resolutions of derived products. This advances the capability of NRT flood detection to support policymaking (see Section 1.3.3.2), but was limited at the local or regional scales. The first SAR-based operational global-scale flood product was supplied by the New ESA's twin satellites Sentinel-1A/1B. Multi-mission SAR and optical imagery datasets are capable of mapping floods at high spatio-temporal resolution and global scale by complementing each-other (Martinis et al., 2022).

1.4.4 L-band SAR data availability

Most of the SAR sensors are operating at C and X band (see Table 1.2), because it permits a higher resolution, better coherence and precision. However, C and especially X bands barely penetrate the dense forests. L-band SAR have lower spatial resolution but better penetration in the vegetation. It makes them suitable for monitoring flooded forests and for the characterization of biomass properties.

One of the major gaps in the knowledge of the water cycle is the functioning and dynamics of tropical wetlands and floodplains, and their interactions with the biogeochemical cycles (see Section 1.2 and Section 1.1). For this, the availability of L-band data is essential. Unfortunately, very few missions have been launched carrying L-band SAR instruments.

The JERS-1, ALOS/PALSAR-1 and ALOS/PALSAR-2 missions provide 30 years of discrete records. Two main data gaps exist between the launch of the satellites, from 1998 to 2006 and from 2011 to 2014. Furthermore, the acquisition was limited in terms of revisit time, and data acquisition was performed upon demand with no continuous records available.

Despite the limitations of the sensors, L-band SAR provided precious reference maps of surface water in the Amazon (Hess et al., 2003) and Congo (Betbeder et al., 2014) basins. The future L-band missions NISAR and ALOS-4/ PALSAR-3 will hopefully provide large-scale monitoring of the surface water with a >bi-monthly revisit to help filling current gaps.

1.5 Future advances expected

Two satellite missions will hopefully provide breakouts in tropical wetland hydrology in the next years: the Surface Water and Ocean Topography (SWOT) from NASA/CNES (Centre National d'Etudes Spatiales), to be launched by the end of 2022, and the NASA/ISRO (Indian Space Research Organization) SAR (NISAR), planned for 2024. They will complement to provide a full coverage of the Earth surface for both SWE and SWS monitoring.

NISAR will be the first SAR mission to perform bi-frequency measurements at both L and X bands. It will have a 12-day repeat cycle, sampling the Earth surface every 6 days on average with high resolution (3-10 m, see Table 1.2). It will provide for the first time, a continuous, L-band and high resolution monitoring of the inundation dynamics in tropical forests, which is essential to understand the regional water cycles and their contributions to regional and global biogeochemical processes. ALOS-2/ PALSAR-2 and its successor ALOS-4/ PALSAR-3 from JAXA will be continuing the historical Japanese L-band SAR missions at the same time. Both NISAR and PALSAR-3 will largely contribute to filling the gap in L-band SAR data.

The SWOT mission will provide unprecedented features for land hydrology (Biancamaria et al., 2016). Although it is described as a radar altimetry mission, its technology is pioneering. The main instrument is the Ka-band Radar Interferometer (KaRIn) which will measure surface water in a 120-km wide swath, with a 20 km gap across nadir. It will monitor rivers as small as 100 m wide, and water bodies of 250 $\times$ 250 m. Early experimental results show that Ka-band InSAR imagery have strong coherence on water bodies and separates well water and land surfaces (Fjørtoft et al., 2013). SWOT will provide the first global and consistent monitoring of SWS changes, with a complete inland water mask and the corresponding InSAR water levels at 100 m spatial resolution evaluated once every repeat cycle (21 days). It will make available water height, discharge, storage and extent of lakes, rivers and floodplains.

However, the SWOT mission does not include requirements over vegetated wetlands. It is up to now uncertain on what extent the Ka signal of KaRIn will be affected by the vegetation (Fjørtoft et al., 2013; Biancamaria et al., 2016). The capabilities of SWOT for monitoring the several hydrological variables enunciated above in flooded forests will have to be evaluated on further studies (Papa and Frappart, 2021). For SWS monitoring, its combination with

other data sources is still promising. NISAR will likely complement data gaps to understand tropical wetland hydrology. Spaceborne GNSS-R constellations can also map flooded areas below the canopy with high temporal sampling. In particular, the current NASA Cyclone GNSS (CYGNSS) and future ESA HydroGNSS missions are of high interest in the field of hydrology (see Chapter 2).

Chapter 2

Global Navigation Satellite System Reflectometry (GNSS-R)

Contents

2.1 Introduction

In this chapter, the Global Navigation Satellite System (GNSS) and the monitoring of GNSS reflected signals, called GNSS-Reflectometry (GNSS-R), are introduced. Note that another PhD thesis was previously conducted on GNSS-Reflectometry at the University of Toulouse - Paul Sabatier, in 2015 (Roussel, 2015). It widely described the GNSS systems, the reflectometry principles and techniques, and its applications. Consequently, only the fundamentals of both GNSS and GNSS-R required for the understanding of my PhD work are described in this manuscript. For additional theoretical notions, the readers are invited to dive into either the PhD thesis from Rodríguez-Álvarez (2011) and Roussel (2015), or other references like Zavorotny et al. (2014) and Darrozes et al. (2016).

2.2 The Global Navigation Satellite System

GNSS satellites are dedicated to the positioning on Earth and in space. It has been developed since the beginning of the US Global Positioning System (GPS) in the 1970s and supports both military and civil usages. GPS has long been the only working GNSS system (Kaplan and Hegarty, 1996), but other global constellations are now fully operational (Section 2.2.2). In particular, the Russian GLONASS, the European Galileo and the Chinese COMPASS/Beidou largely contribute to the amount (>100) of GNSS satellites on orbit. Recently, the regional constellations IRNSS and QZSS have also been set up to monitor specifically the Asian area. A large interest is observed worldwide for the services provided by GNSS beyond positioning, and contributes to other scientific applications such as radio-occultation and GNSS-R.

2.2.1 GNSS positioning

GNSS signals are emitted at L-band and Right Hand Circular Polarization (RHCP, see Figure 2.1). They comprise a *carrier wave* which is the basic signal at the fundamental frequency ($f = 1575.42$ MHz for GPS L1), modulated to carry both a navigation message and the Pseudo Random Noise (PRN) code (Enge, 1994; Kaplan and Hegarty, 1996). The navigation message contains the satellite ephemerides and the precise clock time, ionospheric corrections, and informations on the satellite status. The PRN code is specific to each GNSS spacecraft and permits its identification. It is also the basis of the GNSS positioning principle.

2.2.1.1 Code measurement

The GNSS code measurement consists in determining the distances between one receiver, and several GNSS transmitters. It is a multi-frequency system, as the combination of observables at two frequencies permits to compensate the ionospheric delays due to the interactions of the waves with ionised particles. The PRN code from each GNSS spacecraft is locally replicated

by the receiver, and correlated with the received signal. The delay between the emission and the reception of the signal (Δt_i^s) is there calculated, and gives the distance between the transmitter and the receiver. As the synchronization of the receiver and transmitter clocks is not perfect, the determination of Δt_i^s suffers from small biases. The distance measured with the code is thus called a *pseudorange* (R_p), and is computed as:

$$R_p^* = c\,\Delta t_i^s = c\,(t_R - t_E) + c\,(dt^s - dt_i) = \rho_i^s + c\,(dt^s - dt_i) \tag{2.1}$$

$$R_p = \rho_i^s + c\,(dt^s - dt_i) + d_{iono,f} + d_{tropo} + M_p + \epsilon_p \tag{2.2}$$

where c is the celerity of the wave, Δt_i^s is the measured code delay, t_R and t_E are the times of the wave emission and reception, and dt^s and dt_i are respectively the transmitter (GNSS satellite) and receiver clock biases; ρ_i^s is the geometric distance between the transmitter and the receiver that would have been measured without any clock bias; d_{tropo} and $d_{iono,f}$ represent the atmospheric corrections for the propagation of the signal into the troposphere and the ionosphere; M_p represents the effect of the multipath (see Section 2.3.1), and finally ϵ_p is the receiver noise.

The geometric distance ρ_i^s has three unknowns, which are the coordinates of the receiver. The clock biases represent a fourth unknown in Equation (2.2), thus at least four satellites are required to determine the 3D position of the receiver (Kaplan and Hegarty, 2017). The positioning through code measurement is used by navigation systems onboard vehicles or smartphones. However, its precision is not optimal, on the order of several meters minimum. To enable precise applications, carrier phase measurements must be performed.

2.2.1.2 Carrier phase measurement

The phase measurement is based on the determination of the phase difference between the GNSS carrier, and the local replica generated by the receiver. This phase difference can be determined very precisely (Kaplan and Hegarty, 2017). However, the number of whole wave cycles on the path from the GNSS satellite to the receiver is unknown, and is called the *integer ambiguity* (N_i^s).

The integer ambiguity can be determined in a *double phase difference* configuration, with two receivers observing the same two satellites. The unknowns coordinates and the integer ambiguities are first determined through a least square method. These *floating* ambiguities are fixed into integer values, and the coordinates of the receiver are estimated in a second iteration (Mervart, 1995; Teunissen, 1996). Once the integer ambiguity has been determined, it is known until the lock on the signal is lost by the receiver. During this period known as the *phase tracking*, the distance between the satellite and the receiver (L_i^s) is modeled as:

$$L_i^s = \rho_i^s + c\,(dt^s - dt_i) + \lambda N_i^s - d_{iono,f} + d_{tropo} + M_L + \epsilon_L \tag{2.3}$$

where M_L and ϵ_L correspond to the carrier phase multipath and noise, respectively. Equation (2.3) is used in most scientific applications requiring high precision. The position of

the receiver is determined in post-processing approaches. The *Differential GNSS* (DGNSS) method uses a reference station to determine the coordinates of a mobile/rover, reaching sub-centimetric precision. On the contrary, the *Precise Point Positioning* (PPP) does not require a reference and reaches a quite similar level of precision than the DGNSS when resolving the ambiguities (Ge et al., 2008).

2.2.2 GNSS constellations

GNSS is a set of constellations developed by different countries or continents. It started with the development of the GPS constellation by the US army, whose first Block 1 satellites were launched in 1978 (see Section 2.2.2.1). Since it was declared fully operational in 1994, GPS has long been the only GNSS constellation permitting a global positioning on Earth (Enge, 1994; Kaplan and Hegarty, 1996). Other current global GNSS systems are the Russian *Globalnaya Navigatsionnaya Sputnikovaya Sistema* (GLONASS, Section 2.2.2.2), the European Galileo (Section 2.2.2.3) and the Chinese Beidou/COMPASS (Section 2.2.2.4). These four constellations are now operational for global positioning on Earth with about 25-30 satellites in orbit each, and are used in GNSS-R applications. Other regional constellations dedicated to the Asian area exist, which are not described here: the Japanese Quasi-Zenith Satellite System (QZSS) and the Indian Regional Navigational Satellite System (IRNSS).

2.2.2.1 GPS

The GPS is currently composed of 31 satellites flying in Medium Earth Orbit (MEO), at an altitude of ∼20,200 km and with a 11h58 min revisit time. They are arranged in six equally-spaced orbital planes with 55° inclination, ensuring at least four visible satellites from any point on Earth (Kaplan and Hegarty, 2017). The space segment is composed of several *blocks* of satellites following the main evolutions of the system in time:

- **Block I**: it comprises the first 11 satellites, launched from 1978 to 1985 for validating the GPS concepts. They are not anymore active.
- **Block II**: it comprises the first nine operational satellites, launched in 1989 and 1990. They are not anymore active.
- **Block IIA** (*Advanced*): it comprises 19 satellites, launched from 1990 to 1997. The last active spacecraft from Block IIA stopped functioning in 2016.
- **Block IIR** (*Replenishment*): it comprises 21 satellites launched from 1997 to 2009. The last 8 are a sub-group named Block IIR-M (*Modernized*), as they emit a new civilian code (L2C) on L2 frequency. Seven Block II-R and seven Block IIR-M satellites are still operational.
- **Block IIF** (*Follow-on*): it comprises twelve satellites launched from 2010 to 2016, all operational now. They emit a new civilian signal on the GPS L5 frequency.
- **Block III**: among ten satellites, the first five were launched since 2018 and are now operational. They emit a new civilian signal on GPS L1 frequency (L1C).

Table 2.1: Current operational GPS satellites in orbit, with their civilian frequencies.

GPS block	N° satellites	Date of launch	L1 C/A code	L1/L2 P(Y)	L2C	L5	L1C
Block IIR	7	1997-2004	yes	yes	no	no	no
Block IIR-M	7	2005-2009	yes	yes	yes	no	no
Block IIF	12	2010-2016	yes	yes	yes	yes	no
Block III	5	2018-	yes	yes	yes	yes	yes

The current active GPS satellites and their civilian frequencies are listed in Table 2.1. The GPS signal is composed of three carrier frequencies (two for the oldest satellites without L5), respectively the GPS L1 (f = 1575.42 MHz, λ = 19.05 cm), L2 (f = 1227.60 MHz, λ = 24.42 cm) and L5 (f = 1176.45 MHz, λ = 25.48 cm). They are modulated by:

- The **civilian code C/A** (*Coarse Acquisition*) is a short sequence modulating the GPS L1.
- The **military code P** encrypted with the code W, forms the code P(Y) for military usages. It modulates both the GPS L1 and L2.
- A new **code M** modulating both L1 and L2 for military usages.
- New civilian codes **L1(C)** and **L2(C)** modulating the GPS L1 and L2, respectively.
- The **navigation message** as a binary signal transmitted every 30 minutes.

For the most recent satellites, and provided that the receiver is able to track all the frequencies, four GPS civilian codes are available on three carrier frequencies: the L1 C/A, L1(C), L2(C) and L5. Multi-frequency data are interesting for remote sensing applications, especially GNSS-R (see Chapter 3).

2.2.2.2 GLONASS

The GLONASS constellation has been developed during the same period than GPS, with a first satellite launched in 1982 and the system declared fully operational in 1993 (Langley, 1997). However, it suffered from low maintenance after the fall of the Soviet Union. Following the launch of new satellites, it only became effective again for global positioning in 2011 (Springer and Dach, 2010). GLONASS nominal configuration includes 24 + 3 spares satellites at an altitude of ∼19,140 km, arranged on three orbital planes with an inclination of 64.8° and a revolution period of 11h15 min. As of September 2022, 26 satellites were in orbit with 22 operating and the others in maintenance (GLONASS constellation status, https://www.glonass-iac.ru/en/sostavOG/). There are several generations:

- **Glonass** (first generation): they were launched from 1985 to 1990 and composed the first operational GLONASS constellation. They are now out of service.
- **Glonass-M**: they were launched from 2003 to 2016 and constituted the global operational positioning system from 2011.
- **Glonass-K**: they were launched from 2011 to 2018 and included for the first time the CDMA (see below).
- **Glonass-K2**: they are deployed since 2017 and are still in development.

Table 2.2: Current operational GLONASS satellites in orbit and their civilian signals. The acronym of each signal must be red as follow: the two first letters correspond to the carrier frequency (L1, L2 and L3), the third letter to the availability of the signal (O stands for Open access and S for restricted access signals, not shown here), and the last letter indicates the signal type (F is for FDMA, C is for CDMA).

Generation	Date of launch	Signal type	L1 signals	L2 signals	L3 signals
Glonass	1982-2005	FDMA	L1OF	-	-
Glonass-M	2003-2016	FDMA	L1OF	L2OF	-
Glonass-K	2011-2018	FDMA, CDMA	L1OF	L2OF	L3OC
Glonass-K2	2017-	FDMA, CDMA	L1OF, L1OC	L2OF, L2OC	L3OC

The structure of the historic GLONASS signals differs from GPS, BeiDou/COMPASS and Galileo. In a Code Division Multiple Access (CDMA) strategy, each GPS (or GNSS) satellite emits a signal at a fundamental frequency, and the modulation by the PRN code allows its identification. On the contrary, GLONASS uses a Frequency Division Multiple Access (FDMA) strategy where each spacecraft emits its own, distinct frequency (Langley, 1997). For the GLONASS L1, L2 and L3, these frequencies are:

$$f_{L1} = 1602 + k \times 0.5625 \ MHz \tag{2.4}$$

$$f_{L2} = 1246 + k \times 0.4375 \ MHz \tag{2.5}$$

$$f_{L3} = 1201 + k \times 0.4375 \ MHz \tag{2.6}$$

where the channel of the satellite k is an integer number between -7 and 6 (two spacecrafts on opposite sides of the Earth can share the same channel). A code C/A is modulated on the L1 carrier and a code P is modulated on both L1 and L2, but unlike GPS, these two codes are freely available for civilian usage. Since 2011 and Glonass-K, a CDMA transmission has also been implemented in addition to the FDMA strategy (Urlichich et al., 2011). It concerned first the new L3 signals, and now all carrier frequencies in Glonass-K2 generation. Table 2.2 summarizes these characteristics and the caption explains in detail the acronyms used.

2.2.2.3 Galileo

Galileo comprises 30 satellites at an altitude of ∼23,222 km, arranged on three orbital planes with 56° inclination and 14h21 min of revolution period. It ensures higher coverage of the polar regions than the GPS. The first four operational satellites were launched in 2011 and 2012, with following batches making the constellation operational. Galileo provides an open navigation service (OS) for all users, a commercial service (CS) with higher positioning precision, two security services (*Safety Of Life* and *Search And Rescue*), and a public regulated service (PRS) which is encrypted.

Galileo satellites are identified using the CDMA strategy just as GPS. The signal is composed of four carriers which are listed in Table 2.3, and modulated for the different services listed above (Hein et al., 2006). The commercial service and the public regulated service have

Table 2.3: Overview of the Galileo carrier frequencies and the services available. Os stands for Open Service, CS for Commercial Service, and PRS for Public Regulated Service

Carrier	Frequency	Wavelength	OS	CS	PRS
E1	1,575.42 MHz	19.03 cm	yes	no	yes
E6	1,278.75 MHz	23.44 cm	no	yes	yes
E5a	1,176.45 MHz	25.48 cm	yes	no	no
E5b	1,207.14 MHz	24.83 cm	yes	no	no

their own signals. Note that the E1 and E5a carriers have the same fundamental frequency than the GPS L1 and L5, respectively. This enhances the interoperability of both systems, and increase the amount of L5/E5 data to be collected for scientific applications.

2.2.2.4 BeiDou/COMPASS

The BEIDOU/COMPASS constellation is certainly the most original as it combines both a global and a regional system. BeiDou originally covered only China with the BeiDou-1 geostationary satellites. To extend and replace this regional system, China launched the 35 BeiDou-2 and BeiDou-3 (also named COMPASS) satellites, designed for global positioning. Five satellites are on a geostationary orbit over China, three on an Inclined Geosynchronous Orbit (IGSO), and the last 27 are on a usual MEO like other GNSS spacecrafts (Yang et al., 2019). Since 2020 the constellation is now completed, and provides operational GNSS solutions.

As for GPS and Galileo, the BeiDou signals use the CDMA approach. Each carrier is modulated by the navigation message and a ranging code used for positioning. The carrier frequencies and wavelengths are summarized in Table 2.4. The BeiDou B1C frequency is the same that GPS L1 and Galileo E1, which enhances the GNSS interoperability with about 90 satellites emitting a carrier at 1575.42 GHz. Interestingly, BeiDou B2a has the same frequency than the Galileo E5a and the GPS L5. Finally, the BeiDou B3 is close to Galileo E6.

Table 2.4: Overview of the civilian signals for the COMPASS/BeiDou constellation

Carrier	Frequency	Wavelength	BDS-2 (15 satellites)	BDS-3 (18)
B1I	1,561.10 MHz	19.20 cm	yes	yes
B1C	1,575.42 MHz	19.03 cm	no	yes
B2a	1,176.45 MHz	24.83 cm	no	yes
B3	1,268.52 MHz	23.63 cm	yes	yes

2.3 Principle of GNSS-R

The principle of GNSS-Reflectometry (GNSS-R) was presented theoretically by Hall and Cordey (1988), and developed for ocean altimetry applications in Martin-Neira et al. (1993).

Another technical report introduced the first scatterometry analysis (Katzberg and Garrison, 1996), and finally Garrison and Katzberg (1997) presented the first GNSS-R experiment using data acquired by an airborne system. For early studies, GNSS-R was in fact only based on the GPS constellation. It remained fairly confidential until the early 2000s, mainly due to a lack of precision. New GNSS constellations and an improved accuracy of the retrieval methods have increased the interest for GNSS-R techniques and applications. Also, GNSS-R instrumentation is very low-cost when compared to other remote sensing instruments, as it only requires a receiver to collect the GNSS reflected signals. Furthermore, it is based on GNSS constellations which are essential for both civil and military uses. It is thus viable for long-term uses permitting to improve the technique over the years.

The early GNSS-R experiments were concentrated over the ocean, for both sea level (Anderson, 2000; Ruffini et al., 2004; Larson et al., 2009a) and sea state/wind speed determination (Garrison and Katzberg, 2000; Garrison et al., 2002; Cardellach et al., 2003). The applications are now diversified, especially over land: soil moisture (Larson et al., 2008b; Camps et al., 2016), river and lake altimetry (Beckheinrich et al., 2014; Li et al., 2018), snow depth (Larson et al., 2009b; Rodriguez-Alvarez et al., 2012; Nievinski and Larson, 2014), sea ice detection and thickness (Alonso-Arroyo et al., 2017; Schiavulli et al., 2017), biomass (Carreno-Luengo et al., 2020) and vegetation height retrieval (Zhang et al., 2017; Munoz-Martin et al., 2022). Near global applications are made possible by the launch of space GNSS-R missions like CYGNSS over the last decade. GNSS-R techniques benefit from a very high temporal resolution compared to other remote sensing systems, with sub-second measurements for in-situ receivers and sub-daily revisit for spaceborne observations.

In this section, the properties of the GNSS signals reflected on the Earth's surface are presented. The GNSS-R techniques, sensors and applications are further described in Section 2.4 and Section 2.5. The determination of water levels using geodetic receivers (Section 2.4.2.3) and the detection of surface water using spaceborne GNSS-R observations (Section 2.5.2.3) are particularly highlighted, to pave the way to the next chapters presenting my thesis work.

2.3.1 GNSS multipath

The *multipaths* represent the fraction of a GNSS signal reflected by the Earth surface (land, ocean and ice sheets), vegetation or human constructions, contributing in the signal recorded by the receiver. Multipath degrades the precision of GNSS positioning, for both the code and phase measurements. They extend the wave travel and produce errors up to tens of meters in the code pseudo-range. They also produce an additional phase shift in the more precise phase measurements, and provoke errors on the order of 1-2 wave cycles (∼20-40 cm) or more (McGraw and Braasch, 1999).

For these reasons, the filtering of multipaths for GNSS positioning is important. The geodetic antennas are designed to favour the direct signals with a strong antenna gain at Right Hand Circular Polarization (RHCP) on the higher hemisphere, and an antenna gain at nadir as low as possible. For the high-precision positioning requirements, the use of *choke*

rings antennas is recommended as they eradicate the low elevation signals. Post-processing strategies with a large amount of observations is used to detect and filter observations with high residues. On the contrary, the *kinematic* (instantaneous positioning) modes are very sensitive to multipaths because they use few observations to determine the position of the receiver.

The limitation of positioning is an opportunity for GNSS-R methods, exactly based on the exploitation of the multipaths. GNSS-R is sometimes described as a technique using the *Signals of Opportunity* (SoP), that is to say the unwanted signals considered as noise in GNSS positioning. Hence, the properties of these reflected signals are very interesting for remote sensing applications.

2.3.2 Polarization of the reflected signals

The polarization of an electromagnetic wave describes the evolution of the electric field $\vec{E}$ in a plane perpendicular to the direction of propagation. It can be either linear, elliptic or circular - the latter being a particular case of the elliptic polarization. For the linear polarization, the vector $\vec{E}$ has a constant direction in time in the perpendicular plane. For the elliptic and circular polarizations, $\vec{E}$ describes an ellipse or a circle in the plane during the propagation of the wave. Looking at the direction of propagation, the polarization is *right handed* when the ellipse is travelled with clockwise rotation, and is *left handed* otherwise. Figure 2.1 supplies a graphical representation of the RHCP and LHCP signal propagation.

A linearly polarized wave can systematically be written as the sum of a LHCP and a RHCP waves of similar amplitude (Cheng, 1993), which is essential afterwards. For the GNSS multipaths, the power of the incident/direct RHCP signal is always greater or equal (for a perfect «mirror») to the power of the reflected signal. This one is composed of both a RHCP and a LHCP components. The Fresnel reflection coefficients Γ_{RHCP} and Γ_{LHCP} indicate the fraction of the incident signal amplitude that is returned into the RHCP and LHCP components, and are expressed as (Ulaby et al., 2014):

$$\Gamma_{RHCP} = \frac{\Gamma_H + \Gamma_V}{2} \tag{2.7}$$

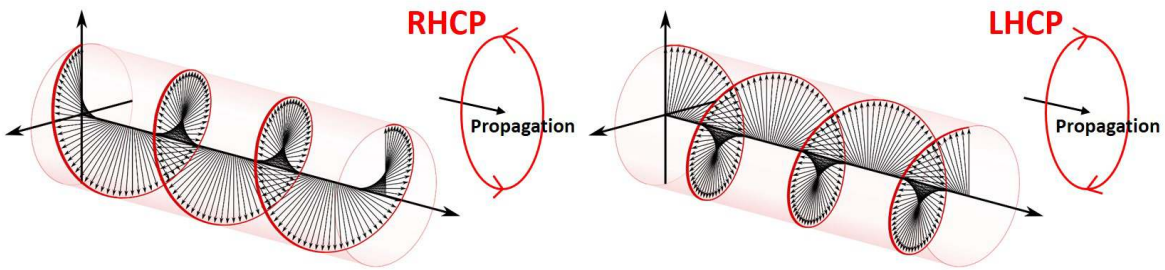

Figure 2.1: Right Hand Circular Polarization (RHCP) and Left Hand Circular Polarization (LHCP) of an electromagnetic wave. Adapted with images from https://en.wikipedia.org/

$$\Gamma_{LHCP} = \frac{\Gamma_H - \Gamma_V}{2} \tag{2.8}$$

$$\Gamma_H = \frac{\sin(\theta_e) - \sqrt{\epsilon - \cos^2(\theta_e)}}{\sin(\theta_e) + \sqrt{\epsilon - \cos^2(\theta_e)}} \tag{2.9}$$

$$\Gamma_V = \frac{\epsilon \sin(\theta_e) - \sqrt{\epsilon - \cos^2(\theta_e)}}{\epsilon \sin(\theta_e) + \sqrt{\epsilon - \cos^2(\theta_e)}} \tag{2.10}$$

where θ_e (degree) is the GNSS satellite elevation angle and ϵ (without units) is the complex dielectric constant of the reflecting surface, related in particular to the soil moisture. Γ_H and Γ_V are the Fresnel coefficients for the horizontal and vertical components of an electric field linearly polarized, respectively. In Equation (2.7) and Equation (2.8), Γ_{RHCP} and Γ_{LHCP} are called the co-polarization component (RHCP), and the cross-polarization component (LHCP) of an electromagnetic wave.

There is a particular elevation angle value where the RHCP and LHCP components of the reflected signal are equal ($\Gamma_{RHCP} = \Gamma_{LHCP}$), because the reflection coefficient at vertical polarization vanishes. It is called the *Brewster's angle* (θ_b), and its value depends on the dielectric properties of the reflecting surface. The reflected signal at θ_b is therefore linearly polarized. For $\theta_e < \theta_b$, the RHCP component dominates the reflected signal such as $\Gamma_{RHCP} > \Gamma_{LHCP}$. The polarization of the reflected signal is therefore elliptic and right handed, tending towards a circular RHCP polarization for $\theta_e = 0°$ (because $\Gamma_{LHCP} = 0$). On the contrary, the LHCP component dominates the reflected signal for $\theta_e > \theta_b$, which is completely circular at $\theta_e = 90°$ because $\Gamma_{RHCP} = 0$.

The polarization of the direct and reflected signals are important for the design of GNSS-R receivers. Single geodetic antennas acquire both RHCP and LHCP signals (see Section 2.4.2), with antenna gain patterns permitting an optimized filtering or the multipaths and enhancing the RHCP direct signal component. On the contrary, dual-antenna configurations (Section 2.4.1) have one zenith looking antenna for acquiring direct RHCP signals, and nadir-looking antennas for recording the multipath components.

2.3.3 Coherence of the reflections

The coherence of a reflection depends on the reflecting surface properties, like the root-mean-square (RMS) roughness. It is an important parameter of GNSS-R signals that modulates the waveform of an acquisition. It is used to study geophysical properties, in particular to detect the presence of water (Section 2.5.2.3).

2.3.3.1 Coherent and diffuse scattering regimes

Once emitted by the satellite, a GNSS signal propagates to the Earth's surface which acts like a reflecting mirror. The shortest path of a reflected signal from the transmitter to the receiver intersects the Earth surface at a particular location, the *specular point*. The receiver

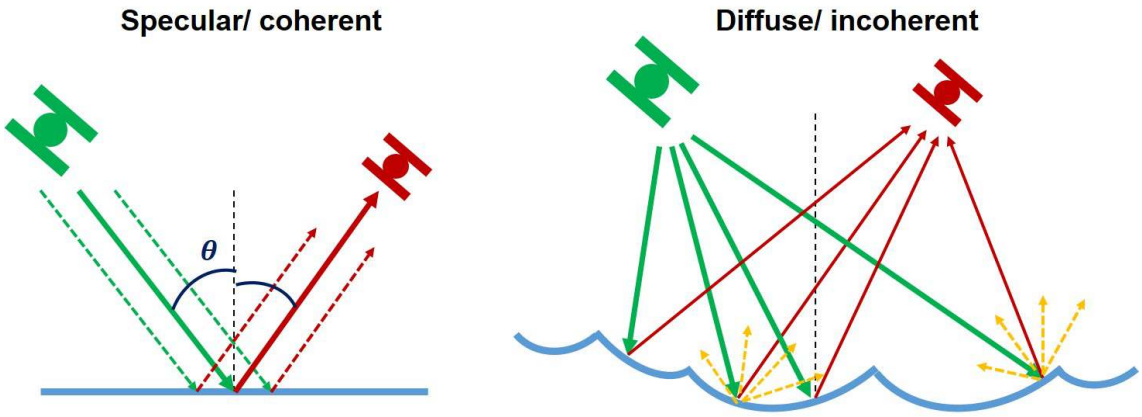

Figure 2.2: Specular (or coherent) reflection over a smooth reflecting surface, and diffuse (or incoherent) scattering over a rough reflecting surface.

first records the contribution from the specular point, which forms the leading edge of the delay waveform. It further records the contributions from the surrounding areas with increasing delay shifts. For a smooth reflecting surface, the reflected signal contains only contributions from areas near the specular point. When it is rough (such as ocean and most of the non-flooded land areas), the reflected signal comprises the contributions from all the surface facets that are favourably oriented, as illustrated in Figure 2.2. It represents therefore a much wider contributing area.

In a bistatic radar configuration like GNSS-R, the forward-scattered signals are composed of both a coherent and an incoherent components, as opposed to the classical monostatic radars where the backscattering is dominantly incoherent (Munoz-Martin et al., 2020c). The coherent component of GNSS-R signals represents the contribution from areas near the specular point, while the incoherent component is due to the diffuse scattering mechanism. The power of the reflected signal $Y(t_0, \tau, f_d)$ can be written as the sum of the power from both components, respectively $Y_{coh}(t_0, \tau, f_d)$ and $Y_{inc}(t_0, \tau, f_d)$ (Voronovich and Zavorotny, 2018):

$$Y(t_0, \tau, f_d) \; = \; Y_{coh}(t_0, \tau, f_d) \; + \; Y_{inc}(t_0, \tau, f_d) \tag{2.11}$$

where t_0 is the time of the acquisition, τ the delay between the direct and reflected signals, and f_d the Doppler frequency shift of the reflected signal (see Equation (2.17)). The coherent and the incoherent components can both dominate the signal, depending on the surface geophysical properties and particularly the roughness (relative to the wavelength). The power of a specular reflection is higher than the power of a reflection dominated by diffuse scattering, because it does not spread along multiple directions.

2.3.3.2 Contribution of the Fresnel zones

Although a simple specular reflection model is easier to understand, the reflection over a smooth surface is better described using wave optics. According to the principle of Huygens-

Fresnel, each point at the intersection of the wavefront and the reflecting surface is itself the source of secondary spherical wavelets (as illustrated in Camps and Munoz-Martin (2020)). These wavelets are scattered along multiple directions and their sum form a new wavefront, which is further called the reflected signal.

The multiple Fresnel zones are used to characterize the GNSS-R scattering regime. Inside a Fresnel zone, the reflections in two points show a path difference of less than half a wavelength, *i.e.* a phase shift of less than 180°. The First Fresnel Zone (FFZ) is thus an ellipse centered on the specular point (whose phase is 0°), extending up to a phase equal to 180°. It is mostly used to describe the spatial resolution of GNSS-R coherent observations as over smooth water surfaces (Camps, 2019). Its semi-minor (r_b) and semi-major (r_a) axes are, for a ground-based GNSS-R receiver (Hristov, 2000; Larson and Nievinski, 2013):

$$r_{b,gound} = \sqrt{n\lambda \left(\frac{h}{\sin(\theta_e)} + \left(\frac{n\lambda}{2\sin(\theta_e)} \right)^2 \right)} \tag{2.12}$$

$$r_{a,ground} = \frac{r_{b,gound}}{\sin(\theta_e)} \tag{2.13}$$

where $n = 1$ for the FFZ, λ is the signal wavelength, h is the receiver antenna height, and θ_e is the satellite elevation angle. The expression of the Fresnel zones changes for a spaceborne GNSS receiver and is given as (Camps, 2019):

$$r_{b,space} = R_{n,space} = \sqrt{n\lambda \frac{R_t \times R_r}{R_t + R_r}} \tag{2.14}$$

$$r_{a,space} = \frac{r_{b,space}}{\cos(\theta_i)} \tag{2.15}$$

where R_t and R_r are the transmitter and receiver ranges, *i.e.* the distance from the transmitter and the receiver to the specular point, and θ_i is the incidence angle at the specular pint. Overall, the higher is the incidence, the larger is the FFZ ellipse and the lower is the peak of the received power. The other Fresnel zones are represented by rings that add around each other, while the path difference increases by half a wavelength and the phase shift increases by 180°.

When the reflected signal comes from a wider area than the FFZ, as over rougher surfaces, the integration of signal samples that are in opposition of phase causes fluctuations in the received power (Camps, 2019). This is especially the case when transitioning from water to land, where ripple effects are observed (Camps and Munoz-Martin, 2020). The modeling results in Loria et al. (2020) suggest that it could also be the case over inland water (see Figure 3 and Figure 4 of this article). The heterogeneity of land surfaces can cause a strong scattering from targets away from the specular point, and therefore induce these fluctuations of the received power.

2.3.3.3 Rayleigh parameter

The spreading of GNSS-R diffuse waves around the theoretical specular direction, or the phase variations of these scattered waves, can be characterized using the Rayleigh roughness parameter (Beckmann and Spizzichino, 1987). It is expressed as a function of the surface RMS (Root Mean Square) height σ_h, *i.e.* the RMS deviation of the reflecting surface according to its average height:

$$R_a = 2\pi \, \sigma_h \, \frac{\cos(\theta_i)}{\lambda} \tag{2.16}$$

with θ_i and λ the incidence angle and GNSS wavelength, respectively. R_a is proportional to the surface roughness, so high values typically represent a regime of completely incoherent scattering. A weaker or moderate diffuse scattering is conversely associated with $R_a \leq 1$ (Alonso-Arroyo et al., 2015; Voronovich and Zavorotny, 2017). This happens with low wind speed over water surfaces, which is more frequent for inland water bodies than for the oceans. The coherent component of the signal coexists with the incoherent scattering under such conditions, and is attenuated or extincted for higher surface roughness.

2.4 GNSS-R techniques

The theoretical concept of GNSS (or GPS, at the time) reflectometry was introduced by Hall and Cordey (1988) and then demonstrated accidentally in Aubert et al. (1994). The first remote sensing study based on the reflection of GPS signals was conducted by ESA, to prove the feasibility of GNSS-R ocean altimetry (Martin-Neira et al., 1993). The early GNSS-R systems used two antennas, one for monitoring the direct GPS signals, and another for monitoring the multipaths. The gain of each antennas is optimized for its purpose, and the two signals are analyzed separately. Later, advances included the possibility of using a single antenna to record both the direct and reflected signals (Kavak et al., 1998). Even if the information of each signal is lost, these receivers can be very low-cost and are therefore useful to be deployed elsewhere on ground stations. However, single-antenna techniques are not suitable for airborne or spaceborne applications (see Equation (2.23) and Section 2.4.2 for more informations).

Up to now, both the single and double antenna configurations have seen an increasing number of scientific contributions, new techniques and applications, and a rapid development from the experimental to the operational phases. Both systems are described in the next subsections and in Figure 2.3. The aim is particularly to introduce ground-based acquisitions using a geodetic antenna, which led to the results presented in Chapter 3, and spaceborne GNSS-R systems using two (or three) antennas as the Cyclone GNSS (CYGNSS) constellation, which led to the results presented in Chapter 4 and Chapter 5.

2.4.1 GNSS-R with two antennas

2.4.1.1 Configuration of the acquisition

As it is shown in Figure 2.3, the double-antenna GNSS-R configuration is composed of one zenith-looking RHCP antenna A_R to record the direct signals, and one nadir-looking antenna A_L to record the multipaths. The LHCP component mostly dominates the reflected signal except for a satellite elevation angle lower than the Brewster's angle (see Section 2.3.2), so A_L is usually a LHCP antenna. The receivers are able to process separately the waveforms of both signals.

With evidence regarding the geometry of the acquisition (Figure 2.3), the reflected signal travels additional distance when compared to the direct signal. The distance traveled by one wave can be estimated either through the code (Section 2.2.1.1) or phase (Section 2.2.1.2) measurements, the second technique being much more precise. In a GNSS-R *conventional* configuration (cGNSS-R), the direct and the reflected signals are correlated with a local replica of the PRN code, to estimate the distances traveled by both waves (Zavorotny et al., 2014). The additional delay of the reflected signal is extracted and is used to derive the additional distance traveled $\Delta\delta$. The receiver height h can then be computed from $\Delta\delta$, the distance between the center of phase of the two antennas d, and the elevation angle θ_e.

Another configuration is possible, where the direct and the reflected signal are correlated between each-other directly. This is called the *interferometric* GNSS-R, or iGNSS-R. This type of receivers does not require the knowledge of the PRN code to correlate the two waves and compute $\Delta\delta$ (Zavorotny et al., 2014). However, iGNSS-R is noisier than cGNSS-R as the correlation is based on the direct signal which may have a low Signal-to-Noise Ratio (SNR). Also, the reflected signals need to be filtered in the iGNSS-R method to match the direct signals, using the antenna beam and the different delay and Doppler.

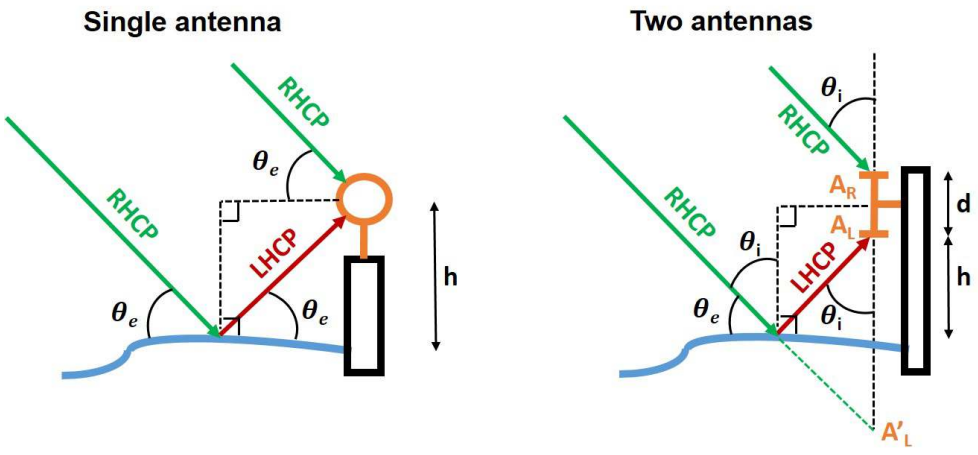

Figure 2.3: Geometry of a bistatic configuration between a GNSS transmitter and a GNSS-R receiver, with a single or two antennas.

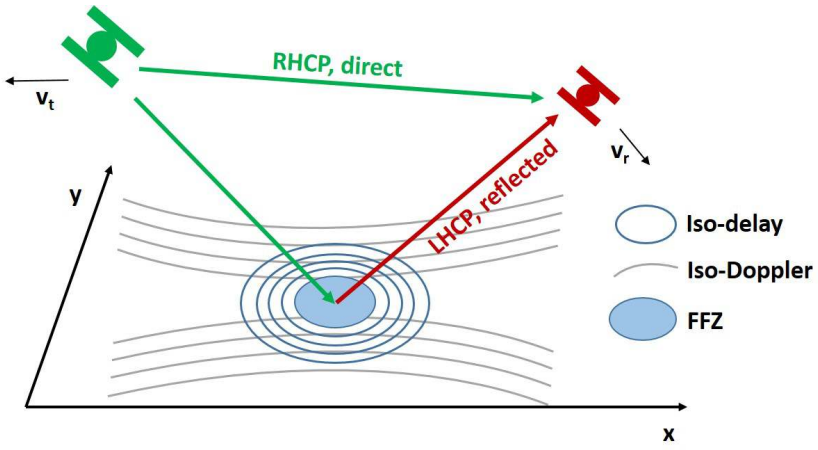

Figure 2.4: Explanation of the Delay and Doppler shifts used to produce Delay Doppler Maps (DDM), with iso-Delay and iso-Doppler lines around the specular point for a bistatic GNSS-R acquisition with both the transmitter and the receiver in motion.

2.4.1.2 Iso-delay, iso-Doppler and Delay Doppler Maps

As the multipaths travel a longer total distance, they are delayed when compared to the direct signal. This delay is minimum at the specular point, and increases as the scattering from surrounding areas arrives at the receiver. Because the emitting GNSS satellite is in motion (and also the GNSS-R receiver for airplanes or spacecrafts), the Doppler effect causes a frequency shift in the received signals. The changes in delay and Doppler when moving away from the specular point are illustrated in Figure 2.4.

The Delay Doppler Maps (DDM) observables represent the received power in two dimensions, along the bins of delay and Doppler shifts. Figure 2.5 presents four DDM from the

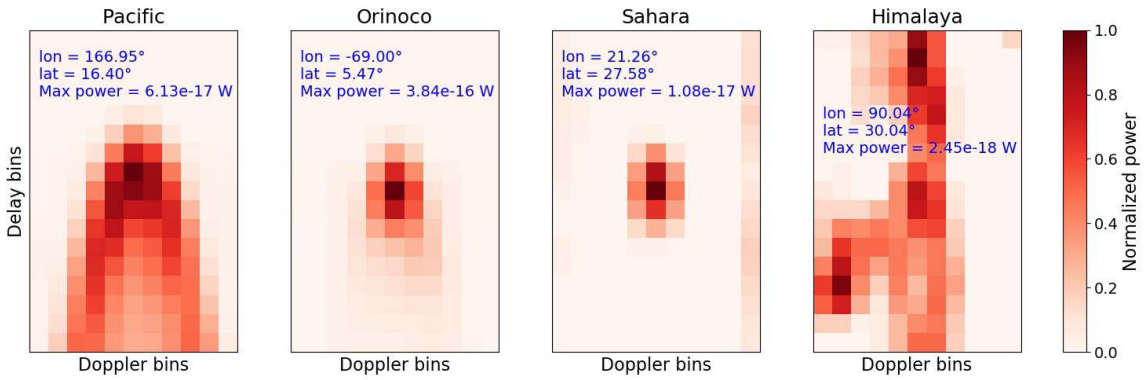

Figure 2.5: Examples of CYGNSS Delay Doppler Maps (DDM) over a rough ocean (Pacific), smooth water during floods in the Orinoco, smooth desert in the Sahara, and mountainous region in the Himalayas. The power in each DDM was normalized between 0 and 1 for the purpose of visualization. The coordinates and the maximum power are printed in blue.

Cyclone GNSS (CYGNSS, see Section 2.5.2) mission. The reflections along an iso-delay ellipse are mapped in the corresponding delay bin (Figure 2.5). Similarly, a Doppler bin corresponds to the reflections along the corresponding iso-Doppler line. The rougher is the reflecting surface, the more regions away from the specular point contribute to the reflected signal, with consequent delay and Doppler shifts. The peak amplitude and the power spread in a DDM are thus good indicators of the surface's geophysical properties. The first DDM in Figure 2.5 was acquired over the Pacific, and looks like a horseshoe with a large spread of the reflecting power typical of rough surfaces. The second DDM was acquired over the flooded Llanos de Orinoco, and is highly coherent with a low power spread. The peak power is there higher than for any other case. The third DDM presents a coherent reflection acquired over the Sahara, due to a low surface roughness. However, the peak power is lower than over calm inland water bodies. And lastly, the fourth DDM was acquired over the Himalayas, highlighting the limitation of CYGNSS in mountainous regions. The several peaks and the low received power are due to the topography, and indicate that the DDM does not contain the specular point.

2.4.1.3 Signal processing

The correlation of the reflected signal $S_r(t)$ with a local replica of the PRN code a^* for cGNSS-R acquisitions takes the same form that the correlation of the direct signal with such replica. The received power is integrated coherently over a short period of time T_c (T_c =1 ms for CYGNSS receivers). At time t_0, with a delay shift τ and a Doppler shift f_d, the reflected power $Y(t_0, \tau, f_d)$ is computed as (Zavorotny and Voronovich, 2000; Voronovich and Zavorotny, 2018):

$$Y(t_0, \tau, f_d) \;=\; \frac{1}{T_c} \int_{t=t_0}^{t_0+T_c} S_r(t)\, a^*\,(t-\tau)\, e^{-j2\pi(f+f_d)\,t}\, dt \tag{2.17}$$

For iGNSS-R acquisitions, Equation (2.17) is rewritten by replacing the local PRN replica a with the direct signal $S_d(t-\tau)$. The thermal and speckle noises are important after this coherent integration. The received power is therefore averaged over a period of time, corresponding to N_i samples. This incoherent average was set to 1 s for CYGNSS data acquired before July, 2019, and to 0.5 s after. The corresponding averaged power takes the form:

$$\langle |Y(\tau, f_d)|^2 \rangle \;=\; \frac{1}{N_i} \sum_{n=1}^{N_i} |Y(t_n, \tau, f_d)| \tag{2.18}$$

For spaceborne GNSS-R missions, two data products can be distinguished: the Level 1 DDM obtained after the incoherent averaging following Equation (2.18), and the Level 0 raw Intermediate Frequency (IF) obtained with coherent integration (Equation (2.17)), but before the averaging. The raw IF are rare (Li et al., 2022), because they require high computational and storage capabilities due to their 1 ms sampling. They are primarily downlinked to ground stations to support mission quality controls. They also have large interest for the scientific community. They contain the phase information that is further lost during the incoherent averaging of the signal. The complex information from coherent raw IF tracks permits to reach a spatial resolution as high as 100 m (Li et al., 2021, 2022), while the DDM products

are limited at hundreds to thousands of meters. Moreover, a precise phase-delay altimetry can be performed using raw tracks as proposed in Cardellach et al. (2004). Spaceborne GNSS-R phase altimetry was further used to retrieve sea ice thickness (Li et al., 2017) and water level variations in a lake (Li et al., 2018). Still, DDMs are the main product of all the past and current spaceborne GNSS-R missions due to a very lower data volume.

2.4.1.4 Modelling of the coherent reflections

As it was previously explained, the GNSS reflected signals over smooth water surfaces are largely dominated by the coherent component of the signal. Moreover, a high received power is observed over water bodies when compared to surrounding areas, facilitating their delimitation Figure 2.5. The modelling of the coherent component of the reflected signal is therefore essential for studying surface water with GNSS-R. From Equation (2.11) and according to the literature (De Roo and Ulaby, 1994; Ulaby et al., 2014; Voronovich and Zavorotny, 2018), the coherent power can be expressed as follows:

$$\langle |Y_{coh}(t_0, \tau, f_d)|^2 \rangle \;=\; \left(\frac{\lambda}{4\pi}\right)^2 \frac{P_t G_t G_r}{(R_t + R_r)^2} \, |\chi(\tau, f_d)|^2 \, \Gamma(\theta_i) \tag{2.19}$$

where $P_t G_t$ is the GNSS Equivalent Isotropically Radiated Power (EIRP) obtained from the transmitted power P_t and the transmitter antenna gain G_t, G_r is the receiver antenna gain, R_t and R_r are the distances from the specular point to the transmitter and from the specular point to the receiver (*ranges*), $\chi(\tau, f_d)$ is the Woodward Ambiguity Function (WAF), and $\Gamma(\theta_i)$ is the specular reflectivity of the surface which depends on the incidence angle θ_i. From Equation (2.19), the peak of the coherent power and the surface reflectivity can be expressed as:

$$P_{coh} \;=\; \left(\frac{\lambda}{4\pi}\right)^2 \frac{P_t G_t G_r}{(R_t + R_r)^2} \, \Gamma(\theta_i) \tag{2.20}$$

$$\Gamma(\theta_i) \;=\; |\Gamma_{LHCP}|^2 \, \gamma^2 \, \Psi \tag{2.21}$$

where $\Psi = \exp(-4R_a^2)$ depends on the Rayleigh coefficient (Equation (2.16)) and accounts for the attenuation of the coherent component by surface roughness, Γ_{LHCP} is the Fresnel reflection coefficient for the reflected signal (mostly LHCP, Equation (2.8)), and γ is the transmissivity of the vegetation computed following Equation (1.6), using the Vegetation Optical Depth at L-band (τ_L or L-VOD). The reflectivity is thus a function of several geophysical parameters of the reflecting surface. With the hypothesis of a fully coherent reflection, P_{coh} is the peak power in the reflected DDM (Clarizia et al., 2019; Eroglu et al., 2019). So, the surface reflectivity $\Gamma(\theta_i)$ can be inverted from GNSS-R observations to study either SM, biomass, surface roughness, or the presence of surface water which causes a high coherent power, as I did in Chapter 4 and Chapter 5.

2.4.2 GNSS-R with a single antenna

Some methods have been developed to analyze the direct and reflected signals collected by a single antenna. Figure 2.3 presents the geometry of this acquisition. This method is simpler in terms of installation and processing chain compare than the two-antenna configuration. In this case, the direct and reflected signals are correlated with each-other in order to extract the observables of GNSS-R acquisitions. The single antenna GNSS-R acquisition can be configured in two manners: using a conventional geodetic antenna (also named the *Signal-to-Noise Ratio* or SNR technique), and using the Interference Pattern Technique (IPT). Their principles and applications are described in the following subsections, in particular the SNR-based altimetry measurements (Section 2.4.2.3). Note that the path difference $\Delta\delta$ between the direct and the reflected signals is equivalent to:

$$\Delta\delta \;=\; 2\,h \times \sin(\theta_e) \tag{2.22}$$

where h is the relative antenna height above the reflecting surface, and θ_e is the satellite elevation angle. $\Delta\delta$ must be lower than the length of one PRN code chip to guarantee the correct correlation of the direct and reflected signals to produce coherent interferences (see Section 2.4.2.1 and Section 2.4.2.2). As the GPS PRN code is composed of 1023 chips and has a length of 1 ms, the maximum path difference $\Delta\delta_{max}$ between the direct and the reflected signals is:

$$\Delta\delta_{max} \;=\; \frac{10^{-3}}{1023} \times c \;=\; 293\,m \tag{2.23}$$

where c is the wave celerity ($c = 3{\times}10^8$ m·s^{-1}). The single antenna acquisitions are only suitable if $\Delta\delta < \Delta\delta_{max}$, *i.e.* whatever the incidence angle for $h < 145.5m$. For $h = 300m$, only the reflected signals with $\theta_e < 30°$ can be monitored. So, the SNR and IPT techniques are generally limited to antenna heights of 300 m at the most. This excludes airborne and spaceborne platforms that are conversely equipped with double-antenna configurations.

2.4.2.1 Interference Pattern Technique

The geometry of the acquisition in the IPT technique is slightly modified when compared to the conventional SNR method (see Section 2.4.2.2). Both the direct and reflected signals are recorded and coherently added by the receiver. The received power fluctuates, with constructive and destructive interferences depending on the satellite elevation angle (see Figure 2 in Rodriguez-Alvarez et al. (2009) and Figure 6 in Rodriguez-Alvarez et al. (2010) for examples). These oscillations have a minimum of amplitude called *notch*, which occurs at the Brewster's angle (Rodriguez-Alvarez et al., 2009).

The antenna used for the IPT is not a conventional geodetic receiver. It is oriented towards the horizon, making the antenna gain patterns symmetric for the direct and reflected signals. It is also linearly polarized for two reasons:

- The linear Fresnel reflection coefficients Γ_v and Γ_h (see Equation (2.9) and Equa-

tion (2.10)) show higher variations with the incidence angle than the circular Γ_{RHCP} and Γ_{LHCP}. It provides additional information about the reflecting surface.
- The antenna is able to record simultaneously the direct and the reflected signals, instead of a RHCP antenna which recording only the direct signal and the RHCP component of the multipaths.

2.4.2.2 Signal-to-Noise Ratio

The more conventional SNR (Signal-to-Noise Ratio) method uses zenith-looking geodetic antennas. The advantage when compared to the IPT method is the potential use of GNSS stations from national and international geodetic networks, with free access to a large existing database. However, the processing of the signal is a bit trickier which lead to a higher noise in the observations. The asymmetry of the antenna gains from the direct and reflected signals needs to be compensated. Also, these antennas are designed to filter the multipaths because they decrease the precision of the positioning (Section 2.2.1). Still, the interferences of the direct and reflected waves create the same oscillations of the received power, which are transposed into the SNR.

In the Rinex files, the SNR is a common observable used to evaluate the quality of a GNSS acquisition. The SNR is extracted from the I (*in-phase*) and Q (*quadrature*) components of the signal recorded by the receiver (Georgiadou and Kleusberg, 1988). It is computed with the same time sampling than GNSS observations. Lets consider the direct signal of amplitude A_d and phase ϕ_d, and the reflected signal of amplitude A_m with a phase difference ψ to the direct signal. The receiver records a composite signal of amplitude A_c and phase ϕ_c, which is the vectorial combination of the direct signal with all the multipaths. The SNR is thus expressed as follows (Bilich et al., 2008; Larson et al., 2008a):

$$SNR^2 = A_c^2 = A_d^2 + A_m^2 + 2A_dA_m\cos(\psi) \simeq A_d^2 + 2A_dA_m\cos(\psi) \tag{2.24}$$

Equation (2.24) is simplified as we assume $A_d \gg A_m$, as the multipath component is attenuated due to the reflection and a lower antenna gain. In the SNR time series, the direct signal has therefore a high amplitude and low frequency. Its contribution was modeled using a second order polynomial fit depending on the satellite elevation angle (Larson et al., 2013a; Roussel, 2015). The multipaths then add a low amplitude and high frequency signal. Figure 2.6 presents the example of a raw SNR time series (a), with the decomposition for the descending pass of the contributions from the direct (b, red curve) and reflected (c) signals. It highlights that the amplitude of the multipath oscillations depends on the elevation angle.

2.4.2.3 GNSS-R altimetry using the SNR

The multipath component of the SNR depends on the receiver antenna height h over the reflecting surface. As can be seen in Figure 2.3, h directly impact the distance travelled by the reflected wave, and consequently modifies the interference patterns between the direct

and reflected signals. The phase difference ψ is directly linked to the difference of distance travelled $\Delta\delta$ (Georgiadou and Kleusberg, 1988):

$$\psi = \frac{2\pi}{\lambda}\Delta\delta = \frac{4\pi h}{\lambda}\sin(\theta_e) \tag{2.25}$$

with θ_e the satellite elevation angle and λ the GNSS signal wavelength. In Equation (2.25), the reflection surface is supposed to be planar. Figure 2.6 shows that once the direct signal is removed (Larson et al., 2013a; Roussel et al., 2015), the multipath component presents high frequency oscillations whose amplitude depends on the satellite elevation angle. The frequency of these oscillations f_ψ is estimated as the derivative of the phase difference:

$$f_\psi = \frac{1}{2\pi}\frac{d\psi}{dt} = \frac{2\dot{h}}{\lambda}\sin(\theta_e) + \frac{2h}{\lambda}\cos(\theta_e)\dot{\theta_e} \tag{2.26}$$

where $\dot{h} = \frac{dh}{dt}$ is the vertical velocity of the reflecting surface, and $\dot{\theta_e} = \frac{d\theta_e}{dt}$ is the elevation angular velocity. Equation (2.26) can be simplified by making a change of variable $x = \sin(\theta_e)$ (Larson et al., 2008a):

$$\tilde{f} = \frac{1}{2\pi}\frac{d\psi}{dx} = \frac{2}{\lambda}\left(\dot{h}\frac{\tan(\theta_e)}{\dot{\theta_e}} + h\right) \tag{2.27}$$

The frequency of the multipath oscillations is there expressed as a function of $\sin(\theta_e)$. It is composed of two components: the relative antenna height above the reflecting surface (h), and the vertical velocity ($\dot{h}$). For static surfaces, *i.e.* when $\dot{h} < 1e^{-6}$ (Larson et al., 2013a;

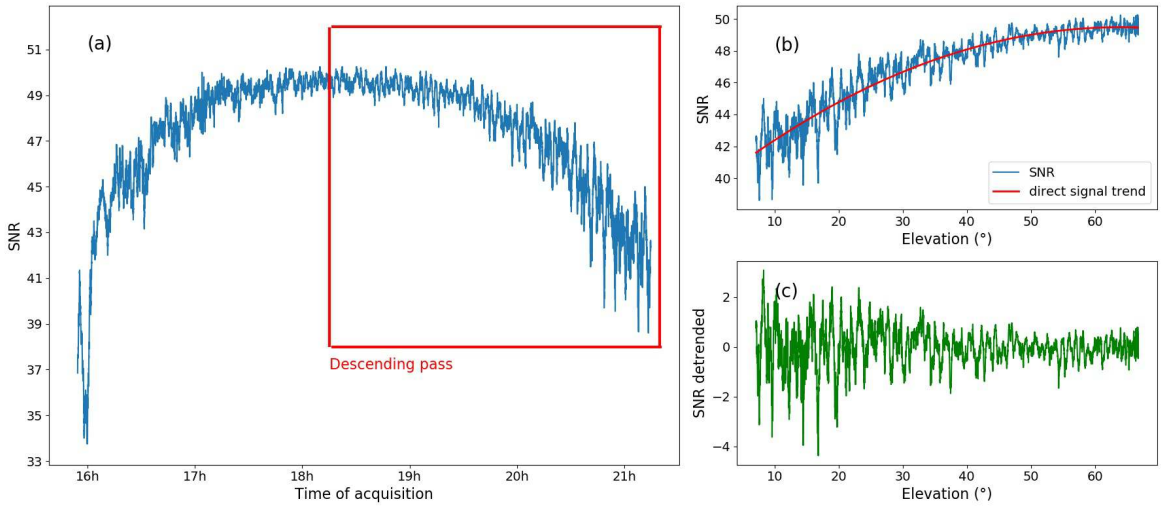

Figure 2.6: Multipath contribution in the SNR time series, for a GPS track from satellite G07 in Podensac, the 17th of October, 2016. **(a)** Complete time series of SNR, **(b)** SNR vs elevation angle for the descending pass with the 2nd order polynomial fit corresponding to the direct signal, and **(c)** detrended SNR showing the multipath oscillations.

Roussel, 2015), it can be simplified as:

$$\tilde{f} = \frac{2h}{\lambda} \tag{2.28}$$

Starting from the detrended signal in Figure 2.6c, a Lomb-Scargle Periodogram (LSP) is usually employed to extract the dominant frequency in the spectrum analyzed (Larson et al., 2009b, 2013b,a; Roussel et al., 2015). The relative antenna height is then directly inverted for SNR altimetry.

Equation (2.28) and the SNR technique have been widely used to retrieve sea surface height (SSH) from coastal geodetic receivers, that are available worldwide (Anderson, 2000; Larson et al., 2013a,b; Löfgren and Haas, 2014; Strandberg et al., 2016; Vu et al., 2019; Purnell et al., 2020; Tabibi et al., 2020; Geremia-Nievinski et al., 2020). All these studies show great accuracy in retrieving SSH, with RMS error of few centimeters. However, the constant height assumption ($\dot{h} < 1e^{-6}$) is obviously not valid for most ocean environments and introduces large errors in the SSH retrieval (see Figure 2.7). Different contributions dealt with this limitation by using tide models in the retrieval process (Larson et al., 2013a; Geremia-Nievinski et al., 2020), by modelling the SSH variations using B-spline functions (Strandberg et al., 2016; Geremia-Nievinski et al., 2020), and also by correcting *a posteriori* SSH estimates for the error introduced by the rate of change (Larson et al., 2013b). Still, most of the studies were limited to places with a low tidal range and low significant wave height (SWH).

Because the SNR method has the potential for monitoring macro-tidal environments, water level variations in rivers and SSH variations during storms, a new technique was needed for these applications. The dynamic SNR method was proposed by Roussel (2015) based on Equation (2.27). It directly estimates both h and $\dot{h}$ using multi-GNSS satellite data, by resolving a system of n equations with two unknowns using the least square method. It showed good results in coastal environment (see Figure 2.7 with a tidal range of ∼4 m), but not in monitoring inland water level variations extreme events like flash floods and storms. Another study used a similar technique with the addition of a third unknown in the system of equations representing the acceleration $\ddot{h}$ (Tabibi and Francis, 2020). It allowed the detection of abrupt water level changes (∼10 m) in a reservoir.

These methods were reproduced and improved in my thesis work to extend GNSS-R water level monitoring to complicated inland water environments, with elevation and azimuth masks, narrow rivers, high vertical velocity $\dot{h}$, and a low number of GNSS satellites visible. The corresponding results are presented in Chapter 3.

2.4.2.4 Other applications

Both the IPT and the SNR methods can be used to estimate various geophysical parameters, especially over land surfaces. Soil moisture (SM) has been widely studied as the signals are very sensitive to changes in the dielectric properties of the reflecting surface caused by

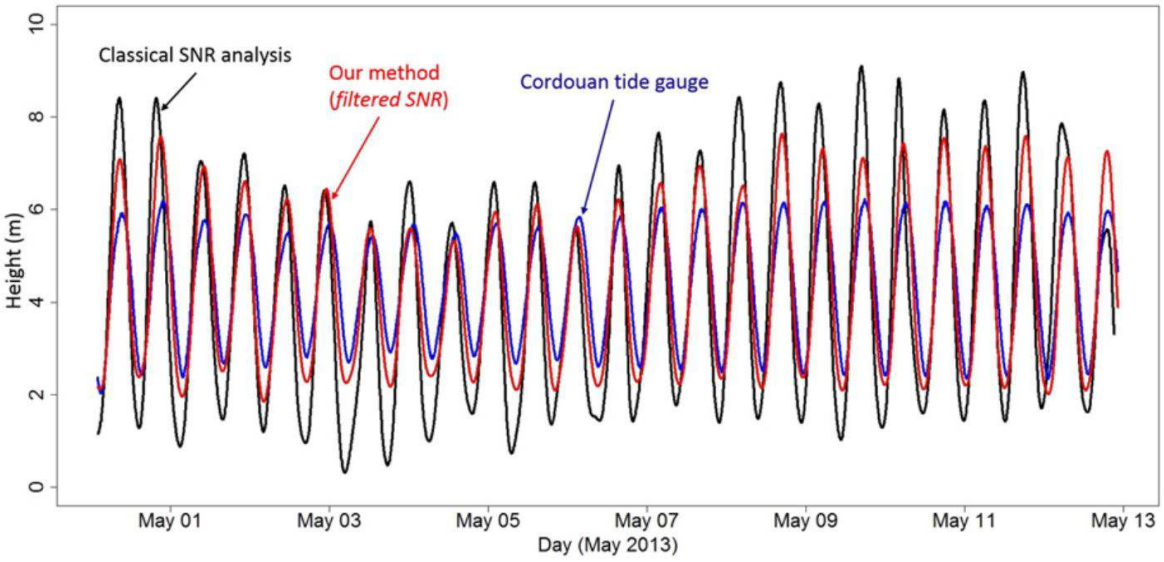

Figure 2.7: Comparison of the SNR-based SSH time series with the tide gauge in Cordouan (blue curve), for a classic retrieval with constant heigh assumption (black curve), and for the dynamic estimation of both h and $\dot{h}$ (red curve). The figure and results are from Roussel et al. (2015).

variations of the soil moisture content, through the Fresnel reflection coefficients (see Equation (2.10) to Equation (2.7)). The multipath SNR (SNR_m) is extracted after the removal of the direct signal component (see the previous section and Figure 2.6), and is further modelled as:

$$SNR_m \;=\; A_m \cos\left(\frac{4\pi h}{\lambda}\sin(\theta_e) \;+\; \phi_m\right) \tag{2.29}$$

where A_m and ϕ_m are the amplitude and the phase of the multipath component, respectively. These two parameters have shown sensitivity to the SM content, especially for the phase which was used by further studies to retrieve SM variations (Larson et al., 2008a, 2009a; Chew et al., 2013; Roussel et al., 2016; Zhang et al., 2018). Empirical relations have been developed to calibrate SM estimates based on SNR_m. The antenna height h needs to be previously determined using Equation (2.28), as the vertical velocity is usually negligible. Small variations of h in time are also recorded as the penetration of the radar signal in the ground changes with the SM content.

The IPT also permitted to successfully sense SM variations, over a bare soil (Rodriguez-Alvarez et al., 2009) and a more complex terrain (Rodriguez-Alvarez et al., 2010). For this, both the amplitude and the position of the *notch* can be used. The latter is advantageous, as it is independent from the surface roughness and only varies depending on the dielectric properties of the soil, *i.e.* SM in a specific place (Rodriguez-Alvarez et al., 2009). The possibility of using both the vertical and horizontal polarizations also makes the IPT power independent from the soil roughness, and improves the SM estimates (Alonso-Arroyo et al., 2014a).

Rodriguez-Alvarez et al. (2010) modelled the air, vegetation and soil structures and their interactions with the electromagnetic wave to retrieve SM in vegetated crops, with typical errors of 2-5%. They also derived the plant height with a few centimeters accuracy, and the topography around the receiver with a few tens of centimeters accuracy. The combined study of vegetation growth and SM has also been performed using the SNR method (Zhang et al., 2017). However, the possibility of retrieving topography is specific to the IPT, as the conventional SNR method assumes a planar reflector.

Another application of the IPT is the retrieval of significant wave height (SWH) over the oceans, exploiting the loss of coherency of GNSS signals measured by the Rayleigh parameter (Alonso-Arroyo et al., 2014b). Finally, the snow depth was estimated using both the IPT (Rodriguez-Alvarez et al., 2012) and the SNR (Larson et al., 2009b) methods. The latter offers the possibility to continuously monitor snow depth (as well as other variables), with the use of national or international geodetic networks (Wan et al., 2022).

2.5 GNSS-Reflectometry from space

GNSS-R has three distinct spatial scales for studying geophysical parameters: the local scale with in-situ receivers, the regional scale using receivers onboard airplanes, and at the global scale using satellite platforms. The applications of local GNSS-R datasets have been described in the previous section. Airborne GNSS-R is also of great interest to study essential environmental variables at the regional scale. As stated in Section 2.4.2, they can only be performed using double-antenna receivers. Besides the first GNSS-R demonstrators onboard aircrafts (Martin-Neira et al., 1993; Aubert et al., 1994), several campaigns were programmed since the early 2000s, especially by ESA. The applications include ocean altimetry (Ruffini et al., 2004; Cardellach et al., 2011) and wind speed (Cardellach et al., 2011) retrieval. Over land, the sensing of soil moisture and biomass was assessed (Cardellach et al., 2011; Egido et al., 2014). For the cryosphere, airborne GNSS-R can be applied to sea-ice, snow and ice sensing (Cardellach et al., 2011).

Despite their interest, this manuscript does not provide more insights into the airborne GNSS-R techniques and applications, as I did not use such data during my PhD work. The in-situ, single antenna SNR method used in Chapter 3 was described in Section 2.4.2. Now, this section introduces the principal spaceborne GNSS-R missions and their applications over land. It is based on the waveform acquisition using two antennas and the signal processing techniques described in Section 2.4.1.3. It is a state-of-the-art for the second part of my PhD work presenting surface water detection based on CYGNSS data.

2.5.1 Early missions

The first spaceborne GNSS-R experiment was performed by a satellite from the British Disaster Monitoring Constellation (UK-DMC), launched in 2003. A second nadir-pointing antenna

was added to the GPS receiver, which was modified to downlink the first spaceborne GNSS-R observations (Gleason et al., 2005). A further analysis of the DDM from UK-DMC helped in demonstrating the feasibility of spaceborne GNSS-R sensing of the sea state (Clarizia et al., 2009).

The UK TechDemoSat-1 (TDS-1) was launched in 2014 to demonstrate the accurate monitoring of wind speed around the globe with GNSS-R. The Space GNSS Receiver - Remote Sensing Instrument (SGR-ReSI) acquiring GNSS reflections onboard TDS-1 was a preliminary version of the Delay Doppler Mapping Instrument (DDMI) carried onboard the next CYGNSS mission (Section 2.5.2). It was the first instrument to provide an amount of DDM large enough to analyze spaceborne GNSS-R observations over various land and ocean surfaces. Furthermore, it downlinked several short raw Intermediate Frequency (IF) tracks (see Section 2.4.1.3). The raw IF contain phase information and allow high-resolution spaceborne GNSS-R applications (Li et al., 2021). They are an object of investigation, prior to the launch of the future HydroGNSS mission (see Unwin et al. (2021) and Section 2.5.3), with a particularly interesting contribution from TDS-1 data (Li et al., 2022).

TDS-1 DDM were used for ocean altimetry (Clarizia et al., 2016) and wind speed retrieval (Foti et al., 2015), paving the way to the next CYGNSS mission. Interestingly, TDS-1 has a Sun-synchronous orbit at an altitude of 635 km and with 98.8° elevation, permitting data acquisitions over the high latitudes. This led to applications over the cryosphere as, for example, sea ice detection (Yan and Huang, 2016; Alonso-Arroyo et al., 2017), sea ice concentration (Yan et al., 2017) and thickness (Yan and Huang, 2020). Using the raw IF samples, a precise phase altimetry over sea ice was also demonstrated (Li et al., 2017).

Over land, the calibrated surface reflectivity showed high sensitivity to SM (Chew et al., 2016), with variations of several decibels between tracks during the dry and the wet season over the same region. Camps et al. (2016) also highlighted the sensitivity of TDS-1 observations to SM, especially over low vegetated areas with a NDVI value below 0.1. The sensitivity to SM was found to decrease with an increasing NDVI, except for dense vegetation in wet regions. However, the Spearman correlation coefficients were low: 0.63 in the best case with few vegetation, and below 0.35 otherwise. To conclude, TDS-1 showed sensitivity to several land geophysical parameters, but the amount of data was not sufficient for large scale or continuous monitoring.

2.5.2 CYGNSS

2.5.2.1 Technical aspect

TDS-1 was followed by the launch of the NASA Cyclone GNSS (CYGNSS) mission in December 2016. It was the first operational GNSS-R mission, for monitoring the formation and the evolution of tropical cyclones through wind speed retrieval (Ruf et al., 2016). It has provided 25 km resolution sub-daily wind speed estimations since its early days (Clarizia and Ruf, 2016). CYGNSS has a very high revisit due to the association of 8 spacecrafts in a

constellation of Low Earth Orbit (LEO) micro-satellites. The DDMI instruments are able to record DDM sampled every 1 s (0.5 s after July 2019), and also to downlink raw IF data. Each DDMI has four channels, thus the entire constellation can make up to 32 observations per second (up to 64 after July, 2019). Over a 25 km pixel, the mean and median revisit times are 3 and 7 hours, respectively (Ruf et al., 2016). CYGNSS made a large amount of Level 1 DDM freely available for land and ocean applications in the inter-tropical band, between -38° and +38° latitude as highlighted in Figure 2.8.

The spatial resolution of CYGNSS has to be specified. Over a rough ocean where the diffuse scattering dominates, it is estimated to a coarse 25 km. This gives the spatial sampling of CYGNSS ocean products (Clarizia and Ruf, 2016). The incoherent scattering regime is also found over many land surfaces. However, over calm water bodies, very wet soils with low roughness and flat arid surfaces, the coherent component of the reflected signal can dominate with a low Rayleigh coefficient. The spatial resolution of a coherent observation corresponds roughly to the FFZ, so about 0.5-1 km depending on the elevation angle (see Camps (2019) and Section 2.3.3.2). For a 1 s incoherent averaging of the received power (Section 2.4.1.3), it is elongated in the along-track direction up to 6.6 km. Still, the strong power observed over rivers with few hundred meters width highlights the sensitivity of CYGNSS observations to a small fraction of water inside this large footprint.

The more topography there is, the weaker are the coherence and the amplitude of the reflected signal. Mountainous regions represent thus an inherent limitation of GNSS-R measurements. Moreover, early CYGNSS DDM (in 2017) were of poor quality above a 600 m threshold, because the estimation of the specular point location was based on a mean sea surface model (Gleason et al., 2019). This led several authors to apply to filter out data above 600 m (Eroglu et al., 2019; Rodriguez-Alvarez et al., 2019; Yan et al., 2020). The Level 1 DDM calibration procedure was further improved with a DEM to account for the elevation and topography effects (Gleason et al., 2020). High quality DDM are now collected up to 2500-3000 m high, with lower quality over the Tibetan plateau and the Altiplano. CYGNSS observations are thus unavailable at high latitudes and high altitude, so the cryosphere and permafrost components cannot be studied.

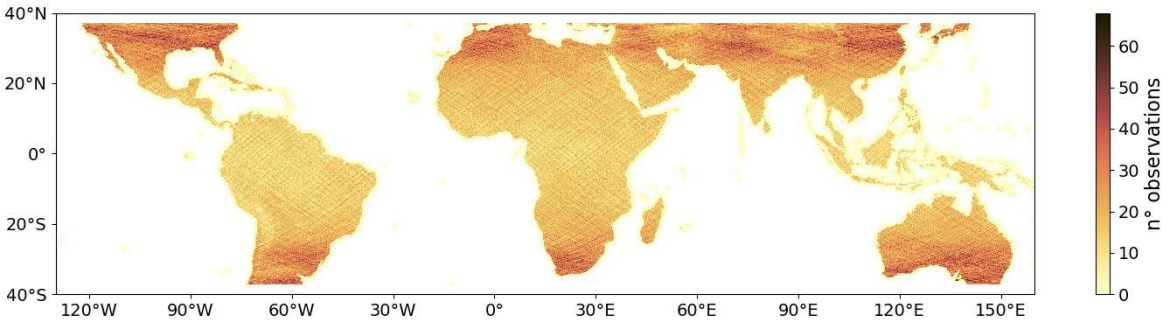

Figure 2.8: Number of CYGNSS observations per 0.1° pixel over land during one month (August 2018).

2.5.2.2 Land applications

Over land surfaces, the main application of CYGNSS data has been soil moisture estimates (see Figure 2.9). Several studies have retrieved SM with accuracy ranging from 0.04 to 0.07 cm^3/cm^3 (Chew and Small, 2018; Kim and Lakshmi, 2018; Al-Khaldi et al., 2019; Carreno-Luengo et al., 2019; Clarizia et al., 2019; Eroglu et al., 2019; Senyurek et al., 2020; Yan et al., 2020). Most of them assumed a coherent reflection over wet soils, which is certainly wrong for SM values under 0.4-0.5 cm^3/cm^3 (Collett et al., 2022). Only Al-Khaldi et al. (2019) supposed an incoherent scattering regime that fits better the soils with low moisture (see Figure 2.9). The large amount of CYGNSS bistatic observations make it possible to improve either the spatial (Yan et al., 2020), or the temporal resolution of SMAP and SMOS SM estimates. So, the University Corporation for Atmospheric Research (UCAR) product provides sub-daily SM estimates in the SMAP 36 km EASE-grid, to sense very rapid variations of this essential climate variable with low spatial resolution (Chew and Small, 2020a).

Many other applications have emerged to take advantage of this huge dataset over land. The specific case of inland water detection will be discussed in the following section. The vegetation properties were also studied, as the GNSS-R coherent component is gradually attenuated with a factor equal to the square transmissivity of the vegetation according to Equation (2.21). This even causes the extinction of the coherent signal for very dense vegetation, from the model point of vue (Loria et al., 2020). However, Carreno-Luengo et al. (2020) have shown that CYGNSS-derived parameters - the trailing edge (TE) and the reflectivity (Γ) - are sensitive to variations of Above Ground Biomass (AGB) up to the range of 250-350 Mg/ha in the Amazon and Congo basins. The robust regressions found in this study with no saturation of the signal (see Figure 9 and Figure 10 in (Carreno-Luengo et al., 2020)) suggest the possibility of AGB retrieval using CYGNSS. Also, Pierdicca et al. (2021) retrieved L-VOD from CYGNSS and TDS-1 data, using neural networks (NN) and the SMAP VOD as input for training the NN and validation. Finally, CYGNSS reflectivity shows spatial variations but is very stable temporally over the deserts. This was used to retrieve surface roughness parameters in a 0.03° grid (Stilla et al., 2020). It offers the possibility to calibrate CYGNSS reflectivity through the known dielectric constant of deserts, and further derive SM from the physical model in Equation (2.21).

2.5.2.3 Detection of surface water

After the retrieval of SM, the detection of surface water has emerged as one of the top land applications of CYGNSS dataset. It takes advantage of the very coherent signal over smooth inland water bodies (Loria et al., 2020; Collett et al., 2022), where surface roughness is lower than over the ocean. The coherent component of the received signal was thus logically used in several studies to derive the surface reflectivity Γ and map inundation extent. The earlier ones presented a separation of flooded and non-flooded areas, before and after the passing of hurricanes, typhoons or other extreme rainfall events (Chew et al., 2018; Morris et al., 2019; Wan et al., 2019; Ghasemigoudarzi et al., 2020; Rajabi et al., 2020). Although the changes in reflectivity were largely related to the presence of water, these studies were mostly conducted

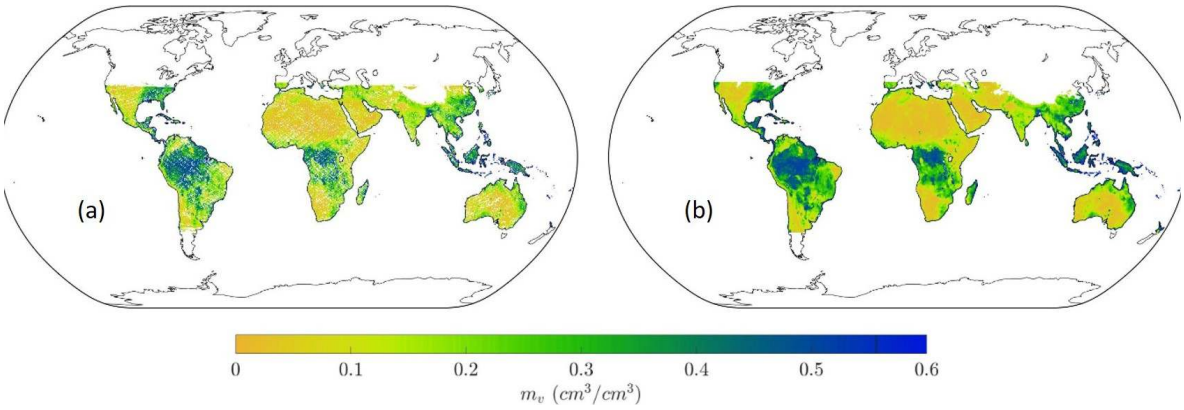

Figure 2.9: Example of SM retrieval using CYGNSS, from Al-Khaldi et al. (2019). **(a)** SM from CYGNSS, **(b)** SM from SMAP.

over specific targets, using a threshold to extract inundated areas. These methodologies are evidently not applicable at larger spatial scales, where the changes in land cover, surface roughness and biomass density contribute in the GNSS reflections.

The influence of the vegetation on flood detection was studied in a small subset of the Amazon basin (Jensen et al., 2018). These results were used in Rodriguez-Alvarez et al. (2019) for classifying open water and flooded vegetation over the same study area, based on a machine learning approach and a training/ validation dataset derived from PALSAR-2. It demonstrated the ability of CYGNSS to detect floods below the canopy, although the accuracy of flooded forest detection was lower than the accuracy of open water detection.

Instead of focusing on inundation dynamics, other studies tried to detect long-lasting surface water with a low time sampling. Gerlein-Safdi and Ruf (2019) used CYGNSS reflectivity to produce annual water masks at 0.1° over small test regions. Their method estimates the mean value of reflectivity in one pixel, and compares it with the average and the standard deviation (std or σ) of reflectivity in a very large box of size 150 × 150 pixels. A threshold on σ helps in separating water and non-flooded pixels, and the map is then refined using a random walker segmentation. It was slightly improved in a further study to perform the same analysis temporally, *i.e.* by comparing the monthly mean value of a pixel with the average and std of its time series (Gerlein-Safdi et al., 2021). This allowed to create a monthly map of water bodies over the Pantanal and Sudd wetlands with a spatial resolution of 0.01° (∼1 km), quite high for an approach based on CYGNSS. The detection of inundations with this method is weakened by changes in the geophysical properties (*i.e.* roughness, SM, biomass) and the flood estate from one pixel to another. Moreover, it also relies on spatial interpolation to fill large data gaps. It was only applied over small test regions, while fitting the algorithm parameters for an optimal surface water detection.

Recent improvements in the detection of surface water have been achieved using the coherence of CYGNSS observations over land instead of the surface reflectivity. In particular, Al-Khaldi et al. (2021b) have proposed an estimation of the coherence based on the shape of CYGNSS Level 1 DDM. The DDM Power Spread Detector (DPSD) method is based on the

Power Ratio (PR), defined as:

$$PR = \frac{C_{in}}{C_{out}} \tag{2.30}$$

$$C_{in} = \sum_{i=-1}^{1} \sum_{j=-2}^{2} DDM\left(\tau_M + i, f_M + j\right) \tag{2.31}$$

$$C_{out} = \sum_{i=1}^{N_\tau} \sum_{j=1}^{N_f} DDM(i,j) \; - \; C_{in} \tag{2.32}$$

where C_{in} and C_{out} are the integrated power received inside and outside a 3 delay × 5 Doppler region centered on the specular point, respectively (see Figure 2.10a,c). Also, τ_M and f_M are the delay and Doppler indices of the specular point, and N_τ and N_f are the number of delay and Doppler bins. A noise exclusion threshold is also applied in the computation of C_{out}, and

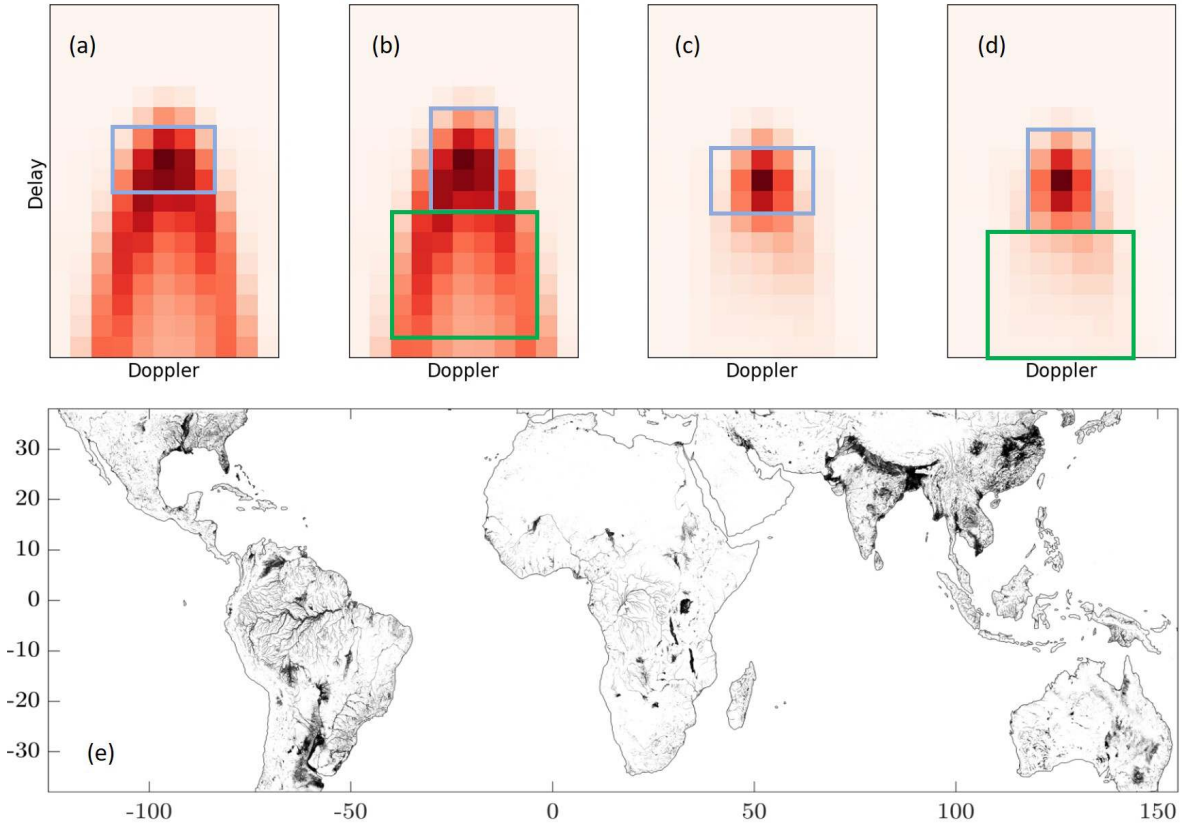

Figure 2.10: Use of CYGNSS coherence to study surface water, with incoherent (**a,b**) and coherent (**c,d**) DDM examples taken from Figure 2.5 (normalized, no units). **(a,c)** Computation of the PR following the DPSD method in Al-Khaldi et al. (2021b,a): C_{in} is computed inside the (3 × 5) delay-Doppler blue box, and C_{out} outside this box. **(b,d)** Computation of the PHPR according to Wang et al. (2022): C_{peak} is computed inside the (5 × 3) delay-Doppler blue box, and $C_{horseshoe}$ inside the (6 × 7) delay-Doppler green box. **(e)** Annual 1-km water mask extracted with the DPSD method, adapted from Al-Khaldi et al. (2021a).

the PR finally represents an estimate of the power spread in the DDM, related to the geophysical properties of the reflecting surface. A PR above 2 was found to be associated to surface water (Al-Khaldi et al., 2021b). The DPSD was used to map water bodies in the full coverage of CYGNSS mission (Al-Khaldi et al. (2021a) and Figure 2.10), using either a 1 year/1 km, a 3 months/3 km or a 2 weeks/6 km grid. A large amount of false alarms were recorded over the vegetated areas due to the choice of the reference dataset (GSW product from Pekel et al. (2016)), which makes difficult to assess the performances in monitoring flooded vegetation. Another similar metric was proposed to (slightly) increase the performances of the algorithm (Wang et al., 2022). It is named the Peak to Horseshoe Power Ratio (PHPR) and defined as $PHPR = C_{peak}/C_{horseshoe}$, with:

$$C_{peak} = \frac{\sum_{i=-2}^{2}\sum_{j=-1}^{1} DDM\left(\tau_M + i, f_M + j\right)}{5 \times 3} \tag{2.33}$$

$$C_{horseshoe} = \frac{\sum_{i=3}^{N_\tau}\sum_{j=-3}^{3} DDM\left(\tau_M + i, f_M + j\right)}{6 \times 7} \tag{2.34}$$

where C_{peak} is the integrated power received inside a 5×3 delay-Doppler region centered on the specular point, and and $C_{horseshoe}$ represents the power received in the 6×7 delay-Doppler region behind the peak value (see Figure 2.10b,d). Both the DPSD and the PHPR methods have shown to be effective for mapping surface water. The two methods are illustrated in Figure 2.10, with the yearly DPSD water mask from Al-Khaldi et al. (2021a). Still, no estimation of the temporal variations of wetland and inundation extent has been performed using CYGNSS, although the data have the potential for it. It is the core of my thesis work, that is presented in Chapter 4 and Chapter 5.

Lastly, some studies showed the detection of open water and flooded vegetation using CYGNSS raw IF tracks (Li et al., 2021, 2022; Collett et al., 2022; Russo et al., 2022; Chapman et al., 2022). The main advantage is to obtain information at very higher spatial resolution than the Level 1 DDM product, using the phase just after 1 ms coherent integration (see Section 2.4.1.3). Also, the coherence derived from raw IF tracks shows lower sensitivity to the vegetation cover and other disturbing variables than the reflectivity. The complex waveform observations permit to detect water bodies down to a hundred meters, with a 1000 Hz sampling rate (Li et al., 2022). The coherence is found over vegetated areas with an AGB up to 300 Mg/ha (Li et al., 2021). If there is no surface water, higher coherence is found over saturated soils with $SM \geq 0.45\, cm^3/cm^3$ (Collett et al., 2022). Overall, the raw IF tracks permit a high resolution mapping of inundation dynamics when compared to the coarse resolution of the parameters derived from L1 DDM (Chapman et al., 2022). However, few raw IF samples exist as they are downlinked to investigate specific targets and to perform quality controls of the CYGNSS products. Thus, they cannot be used for operational surface water mapping. A large dataset with complex phase information will be provided for the first time by the future ESA HydroGNSS mission, to be launched in 2024 (Unwin et al., 2021).

2.5.3 Other systems and perspectives

Other GNSS-R spaceborne sensors were already launched. They include the Passive Reflectometry and Interferometry System - In orbit Demonstrator (PARIS-IoD), which demonstrated ocean altimetry with a precision of about 5 cm (Camps et al., 2014). It is named as the pioneer GNSS-R system proposed by ESA in 1993 (PARIS; Martin-Neira et al. (1993)), and helped in planning the next high precision GNSS-R altimetry missions. In particular, the GNSS REflectometry, radio occultation, and scatterometry onboard the International Space Station (GEROS-ISS) was, as indicated by its name, a GNSS-R experiment onboard the ISS in the framework of ESA programs. It was primarily dedicated to the monitoring of both sea surface height (SSH) dynamics using GNSS-R altimetry, and ocean mean square slope related to surface roughness, wind speed and direction (Wickert et al., 2016).

Another recent space GNSS-R mission is the ^{3}CAT-2 nano-satellite designed by the Universitat Polytècnica de Cataluña (UPC). Its payload is the P(Y) and C/A ReflectOmeter (PYCARO), able to track multi-GNSS, bi-frequency (L1 + L2) and bi-polarization signals to provide measurements of ocean and ice altimetry, wind speed, soil moisture and biomass (Carreno-Luengo et al., 2016). Further evolutions of this instrument include the PYCARO-2 receiver tested on the ground for determining snow and ice thickness (Munoz-Martin et al., 2020d), the Flexible Microwave Payload 1 (FMPL-1) onboard ^{3}Cat-4 (Munoz-Martin et al., 2018), and the FMPL-2 onboard the new Federated Satellite Systems/^{3}Cat-5 (FSSCat) mission developed at UPC (Camps et al., 2018). FFSCat won the 2017 ESA Sentinel Small Satellite (S^3) Challenge and the Copernicus Masters competition. The FMPL-2 payload includes a cGNSS-R receiver and a L-band microwave radiometer to provide low resolution soil moisture, ice thickness and sea ice detection/coverage measurements over the polar regions (Munoz-Martin et al., 2020a). The mission was successfully launched on September 2020 (Munoz-Martin et al., 2020b), and the early GNSS-R and radiometric data from FMPL-2 were evaluated using NN approaches for several applications: soil moisture (Munoz-Martin et al., 2021), wind speed and sea surface salinity (Munoz-Martin and Camps, 2021), and sea ice concentration and extent (Llaveria et al., 2021).

The experience gained from CYGNSS and other missions highlights good performances for retrieving both land and ocean geophysical parameters, but also pointed out the limitations of current GNSS-R observations. Future spaceborne missions can help in overcoming these limitations. The next ESA HydroGNSS mission (Unwin et al., 2021) will be the first dedicated to study the entire water cycle, including inland water. Unlike CYGNSS, HydroGNSS will also be able to collect information near the poles, with lower revisit time than elsewhere. Both the hydrosphere and cryosphere will benefit from the launch of HydroGNSS, planned for 2024. HydroGNSS targets four essential climate variables (ECVs) as its primary objectives: SM, inundation/wetlands, AGB and permafrost with soil freeze/thaw.

HydroGNSS coverage will depend on the number of satellites in orbit. For one satellite and with a 0.25° spatial resolution, a global coverage is expected in 30 days with a mean revisit time of ∼4 days (Unwin et al., 2021). The two-satellite configuration is required for optimizing the performances of the mission, as it provides a 15-day global coverage and a

below 3-day mean revisit. Future scenarios including more satellites will enhance the interest of HydroGNSS for studying rapid variations of the water cycle. Whatever the number of satellites, the interesting new features of HydroGNSS will be:

- Acquisition of Level 1 DDM for both GPS L1 and Galileo E1, offering a double coverage.
- L1 DDM acquired at both RHCP and LHCP polarizations. The two polarizations and the ratio RHCP/LHCP can help in disentangling the effects of several geophysical parameters, *e.g.* soil roughness, moisture and biomass.
- It is planned to experiment the acquisition on a second frequency, *e.g.* the GPS L5/ Galileo E5 signals. These reflected signals may be weaker but can provide a higher resolution due to higher wavelength.
- A new channel allowing to capture coherent amplitude and phase information before the Level 1 incoherent averaging. This provides information with higher spatial resolution (∼100 m in Li et al. (2022)) and should allow precise phase altimetry. These observations will be scheduled but are to represent a much larger dataset than TDS-1 and CYGNSS raw IF tracks. Furthermore, they will include dual frequency and dual polarization samples to investigate reflections over specific targets.

Chapter 3

Inland water levels with GNSS Interferometric Reflectometry

Contents

3.1 Introduction

This section presents the results of the first part of my PhD published in *Remote Sensing* in May 2021 (Zeiger et al., 2021). Although the main topic of this thesis was the monitoring of land geophysical parameters using CYGNSS, we decided to begin with an in-situ GNSS Interferometric Reflectometry (GNSS-IR) case study. This permitted to experience the SNR GNSS-R technique additionally to CYGNSS waveform data. It is also interesting to link the use of GNSS-R datasets for hydrological studies at local, regional and global scales.

The data used in the following article were acquired in 2016 in Podensac by a team composed of scientists from CNES, Environnements et Paléoenvironnements Océaniques et Continentaux (EPOC) and Observatoire Midi-Pyrénées (OMP). The objective of this study was to evaluate the possible retrieval of extreme water height variations using GNSS-R, during the passage of the tidal bores in the Garonne River. Tidal bores are very high-frequency waves propagating upstream the estuaries and rivers under appropriate conditions. For more informations about the spatial structure and the formation of tidal bores, see Bonneton et al. (2015) and Martins et al. (2017). In the Garonne River, this phenomena had a period of approximately 3 seconds and an amplitude of 1.5 - 2 m (see Figure 3.1). The datasets included GNSS positioning with a reference station on the riverbanks and a mobile onboard a buoy on the river. The buoy acquisition permitted to retrieve surface level variations during the propagation of the waves, as in a previous GNSS acquisition campaign in 2015 (Frappart et al., 2016). Pressure and acoustic measurements were also used as reference water heights for this study. For GNSS-R investigations, the GNSS reference station acquired during ∼12 hours multi-frequency (L1, L2 and L5) SNR time series for GPS and GLONASS satellites tracked.

After some investigations, I concluded that the idea of retrieving water height variations during the propagation of tidal bores was a pipe dream. The frequency of the multipath SNR oscillations permits in theory to retrieve h, following Equation (2.28) (Larson et al., 2013a) or Equation (2.27)) (Roussel et al., 2015). As this frequency is usually quite low, at least ∼30 minutes to 1 hour of SNR time series are required for its accurate determination. This, with

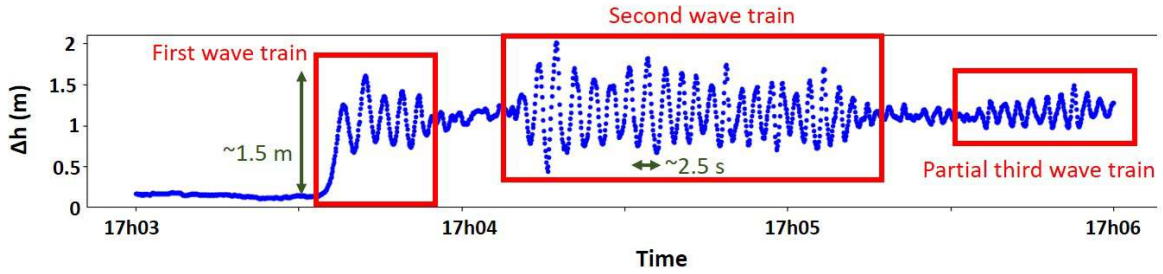

Figure 3.1: Evolution of the water heights during the propagation of the tidal bore in Podensac, from DGNSS processing of the mobile buoy position on the river. The large amplitude (∼1.5 m) and small period ($2.5 - 3s$) of the tidal bore waves are illustrated.

the water level rising and dropping at a rate of ∼1 m/s during the tidal bore propagation (see Figure 3.1). For evident physical reasons, the SNR data and current retrieval methodologies cannot be used to estimate water level variations during such high-frequency phenomena. Alternately, we then focused on retrieving water heights during one asymmetric tide cycle in the Garonne River, showing a large amplitude but a lower frequency than tidal bores. For this, the dynamic SNR method proposed by Roussel et al. (2015) was adapted to the particularities of inland water environments. Asymmetric tides showed higher vertical velocities than what was already observed using GNSS-IR (Larson et al., 2013b,a; Roussel et al., 2015). Moreover, fewer satellites are visible on the river due to elevation and azimuth masks. Indeed, the width of the Garonne in Podensac is approximately 150 m, and the riverbanks are largely vegetated including trees up to 20-30 m high.

The methods, results and discussion about this work are presented in the following publication. I will then comment these findings and my implication in this work. I also added to this chapter some unpublished results on how the tidal bore propagation affects the SNR time series (Section 3.4). Finally, the perspectives for the monitoring of inland water using the SNR technique are discussed.

3.2 Publication in *Remote Sensing*

remote sensing

MDPI

Article

SNR-Based Water Height Retrieval in Rivers: Application to High Amplitude Asymmetric Tides in the Garonne River

Pierre Zeiger [1,*], Frédéric Frappart [1], José Darrozes [2], Nicolas Roussel [3], Philippe Bonneton [4], Natalie Bonneton [4] and Guillaume Detandt [4]

1 LEGOS, UMR 5566, OMP, Université Paul Sabatier, 31400 Toulouse, France; frederic.frappart@legos.obs-mip.fr
2 GET, UMR 5563, OMP, Université Paul Sabatier, 31400 Toulouse, France; jose.darrozes@get.omp.eu
3 Navigation & ATM Chains, Airbus Defence and Space, 31400 Toulouse, France; nicolas.n.roussel@airbus.com
4 EPOC, UMR 5805, CNRS, Université de Bordeaux, 33600 Pessac, France; philippe.bonneton@u-bordeaux.fr (P.B.); natalie.bonneton@u-bordeaux.fr (N.B.); guillaume.detandt@u-bordeaux.fr (G.D.)
* Correspondence: pierre.zeiger@legos.obs-mip.fr

Abstract: Signal-to-noise ratio (SNR) time series acquired by a geodetic antenna were analyzed to retrieve water heights during asymmetric tides on a narrow river using the Interference Pattern Technique (IPT) from Global Navigation Satellite System Reflectometry (GNSS-R). The dynamic SNR method was selected because the elevation rate of the reflecting surface during rising tides is high in the Garonne River with macro tidal conditions. A new process was developed to filter out the noise introduced by the environmental conditions on the reflected signal due to the narrowness of the river compared to the size of the Fresnel areas, the presence of vegetation on the river banks, and the presence of boats causing multiple reflections. This process involved the removal of multipeaks in the Lomb-Scargle Periodogram (LSP) output and an iterative least square estimation (LSE) of the output heights. Evaluation of the results was performed against pressure-derived water heights. The best results were obtained using all GNSS bands (L1, L2, and L5) simultaneously: R = 0.99, ubRMSD = 0.31 m. We showed that the quality of the retrieved heights was consistent, whatever the vertical velocity of the reflecting surface, and was highly dependent on the number of satellites visible. The sampling period of our solution was 1 min with a 5-min moving window, and no tide models or fit were used in the inversion process. This highlights the potential of the dynamic SNR method to detect and monitor extreme events with GNSS-R, including those affecting inland waters such as flash floods.

Keywords: GNSS-R; reflectometry; dynamic SNR; asymmetric tides; water height

Citation: Zeiger, P.; Frappart, F.; Darrozes, J.; Roussel, N.; Bonneton, P.; Bonneton, N.; Detandt, G. SNR-Based Water Height Retrieval in Rivers: Application to High Amplitude Asymmetric Tides in the Garonne River. *Remote Sens.* **2021**, *13*, 1856. https://doi.org/10.3390/rs13091856

Academic Editor: Renato Frasson

Received: 7 April 2021
Accepted: 4 May 2021
Published: 10 May 2021

1. Introduction

The Global Navigation Satellite System (GNSS) has been used for decades for positioning and navigation purposes. Developments of GNSS constellations, more than twenty years ago, led to other remote sensing applications such as GNSS meteorology [1] and radio-occultation [2]. The L2 civilian (L2C) and L5 frequencies offer new tools for sensing atmospheric and Earth surface conditions; specifically, the L5 band improves altimetric performances of GNSS and Global Navigation Satellite System Reflectometry (GNSS-R) retrievals [3].

Recent developments include the improvements and use of GNSS-R as a consistent technique for retrieval of various land and ocean parameters. The operating principle for the use of these opportunistic signals was first proposed in 1988 [4], then experimented in 1993 for GPS constellation with an ocean altimetry case study [5]. These early experiments used a dual-antenna device, one up-looking to acquire the direct signal, and one down-looking to acquire the multipath reflected on the Earth's surface.

The use of a single geodetic antenna to combine both direct and reflected signals was introduced later [6]. The constructive and destructive interferences due to the coherent superposition of direct and reflected paths are transcribed in Signal-to-Noise Ratio (SNR) oscillations. This allowed successful retrieval of various environmental parameters such as soil moisture [7–11], snow depth [12,13], vegetation height [8,14,15] and, of course, ocean altimetry [16–21]. The technique is known as the Interference Pattern Technique (IPT), or GNSS Interferometric Reflectometry (GNSS-IR).

Nowadays, the ability and the interest of GNSS-IR to monitor sea level (SL) changes as a tide gauge is being assessed, especially in coastal areas where the capability of radar altimeters is limited, due to insufficient spatial and temporal sampling, and performance is degraded due to the presence of land contamination in the radar echo [22]. Nevertheless, some limitations of IPT were identified, mostly coming from the retrieval algorithms, especially when the tidal range was high. Most of the methods use a constant height assumption to retrieve the relative antenna height h, related to the frequency of the multipath oscillations [23]. As a consequence, water stages can be estimated only when the elevation rate ($\dot{h}$) is small, i.e., in micro to meso tidal conditions. The complete time series has to be reconstructed using a representation/interpolation of SL as a B-spline, tide models, or fitting a cubic spline on raw heights derived from SNR data [22]; a correction for the rate of change in SL is also necessary [17].

Few studies have been performed, to our knowledge, for retrieving water levels from GNSS-IR over rivers [24]. The major reason is the narrow width of rivers compared to the coastal ocean cases, which drastically limits the number of available reflections over the water surface. The majority of the SNR retrieval approaches only use low-elevation tracks to estimate the dominant frequencies [22], which are mainly masked or flagged in rivers. Some other reasons depend on the location of the study area along the river stream, presence of asymmetric tides in the downstream part of the rivers, occurrence of tidal bores at low tides and, for low-flow conditions over the same areas, occurrence of flash floods, which rapidly modify the river stage in the upstream part of the rivers. The constant height assumption [23] is not adapted to retrieve water levels during such complex and rapid phenomena. First the rate of elevation is higher, and asymmetry in tides prevents us from using models to estimate a correction based on the rate of change [17,25]. Pressure and acoustic measurements from the bottom of the river are commonly used to estimate water heights [26,27], and provide a consistent determination of tides. The case of tidal bores is more complicated, as they are highly nonhydrostatic phenomena, and pressure water levels are reconstructed under the hydrostatic assumption. Other instrumentation can help in better understanding the evolution of water levels and the spatial structure of tidal bores, i.e., GNSS measurements using a mobile antenna onboard a buoy [28], and LiDAR [27].

The dynamic SNR method was proposed to increase the range of applications of the IPT technique to all tidal environments, considering the vertical velocity of the reflecting surface (or elevation rate, $\dot{h}$) that is simultaneously estimated with the relative antenna height h using a Least Square Estimation (LSE) [29]. This allowed it to retrieve the temporal variations of the SL in macro-tidal environments (i.e., with a tidal range higher than 4 m) [19,29,30]. Furthermore, this method introduces a moving window to estimate dominant frequencies along the satellite tracks, while previous approaches estimate only one frequency per track at low elevations [22]. The use of the dynamic SNR method can provide more information when few GNSS satellite tracks are available, and integrates reflections from satellites with high elevation.

In this study, we analyzed SNR data acquired using a GNSS geodetic receiver on a platform over the Garonne River (France), 130 km upstream of its mouth, during one semidiurnal tide. The dynamic method was first used to retrieve the water levels. Due to the asymmetry of the tides and the complexity of the study area (small river, few GNSS satellites visible, vegetation on the riverbanks), it was adapted by introducing two processes for filtering the frequencies estimated with LSP before height inversion. Validation was performed using pressure water levels in the Garonne River.

2. Study Area and Datasets

2.1. Study Area

The study area is located along the Garonne River in Podensac, southwest France (see Figure 1a,b). The confluence between the Garonne and Dordogne rivers forms the Gironde estuary in the Bay of Biscay. Tidal waves propagate in both rivers up to 160 km from the mouth of the estuary, with a tidal range exceeding 6m in Podensac [26,27]. Podensac is about 120 km upstream the mouth of the estuary and is the place where both the tidal range and the asymmetry of the tides are maximal due to a narrow river configuration.

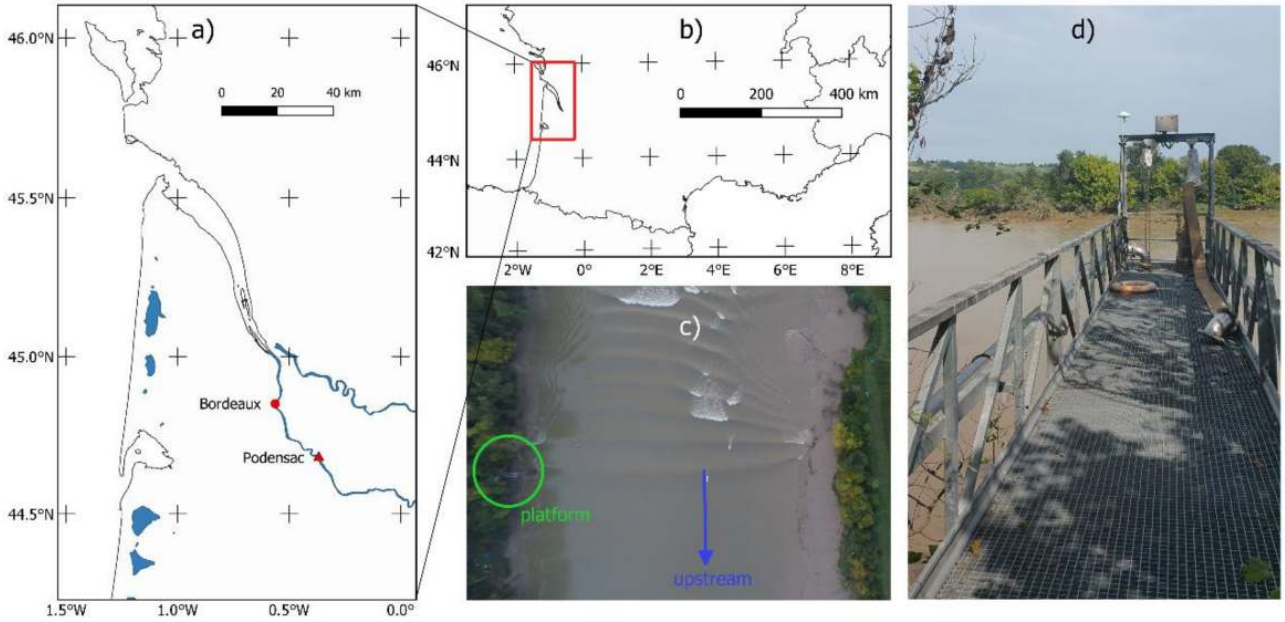

Figure 1. Location of the study area. (**a**) Location of Podensac on the Garonne River; (**b**) the Gironde/Garonne/Dordogne estuary in southwest France; (**c**) drone image of the Garonne River taken by V. Marie (EPOC), showing the first waves of the tidal bore, its direction of propagation, and the platform location; (**d**) photo of the Garonne River from the platform with the GNSS antenna installed. The narrowness of the river and the vegetation on riverbanks are visible in both images.

Tides in alluvial estuaries, such as the Gironde/Garonne/Dordogne estuary, are highly asymmetric [27,31], unlike coastal areas where GNSS-R classical retrieval methods compare well to tide gauges [22]. While symmetrical patterns are found in most studies, asymmetric tides in Podensac compose a challenging dataset to explore the ability to retrieve high amplitude water level variations using GNSS-R. Rising tides last around three hours, while falling tides are slower, around nine hours, to complete a semidiurnal cycle. During faster rising tides, elevation changes materialized by the vertical velocity $\dot{h}$ reach 0.5 mm $\cdot$ s^{-1} to 1 mm $\cdot$ s^{-1}. Moreover, the Garonne River is affected by tidal bores, whose intensities are particularly high during spring tides combined with low discharge [26]. Extremely rapid variations of water height during this phenomenon (>1 m $\cdot$ s^{-1} [27]) affect the performances of GNSS-R retrieval methods onsite. The structure of tidal bores is shown in Figure 1c. More information about the tides and the formation and dynamics of tidal bores in the Garonne and the Dordogne Rivers can be found in the literature [26,27,31].

The Garonne River in Podensac forms a straight line oriented from south-southeast (azimuth = 160°, upstream) to north-northwest (azimuth = 340°, downstream). The platform we used for SNR acquisition was on the left side of the river, whose width is around 150 m. We removed data in a 10° buffer corresponding to the potential interactions of reflected signals with the riverbank and the vegetation (including trees up to 20 m high: see Figure 1c,d). As no reflection was acquired on the north due to GPS orbits, the final azimuthal range was 10–150°. We also masked the data when elevation was below 5° (reflected signals affected by the vegetation on the opposite riverbank) or above 70° (reflections on the platform).

These constraints drastically reduce the number of satellites tracked, and particularly the number of low-elevation tracks available for analysis.

2.2. GNSS-IR Data

A GNSS station composed of a conventional Leica AR10 antenna and GR25 receiver was installed on a platform above the river, ~4 m to 10 m above, depending on the tides (Figure 1d). It was used as a reference station to process the data acquired by a mobile GNSS receiver onboard a buoy using Differential GNSS (D-GNSS) technique, to retrieve water heights during tidal bore occurrence [28]. This reference station was also employed to collect SNR information at L1 (1575.42 MHz for GPS and 1602 MHz for GLONASS with a 562.5 kHz frequency shift between each carrier), L2 (1227.60 MHz for GPS and 1246 MHz for GLONASS with a 437.5 kHz frequency shift between each carrier), and L5 (1176.45 MHz for GPS) GNSS bands from GPS and GLONASS satellites reflected over the river.

The SNR acquisition was performed during a time span of 15 h, starting on October 17, 2016, at 2:35 p.m. UTC and ending on October 18, 2016, at 5:50 a.m. UTC. The nominal sampling rate was set to 20 Hz to further evaluate the ability of SNR-based retrieval methods to detect tidal bore propagation. A first tidal bore occurred on October 17th at 05:03 p.m., and a second on October 18th at 05:30 a.m. UTC (see Figure 2). In this study, we only focused on tide retrieval. Thus, 1 Hz resampled SNR data were used as inputs over the entire time span to retrieve the evolution of water heights during asymmetric tides. The output of our method produced a regularly spaced antenna height time series. Sampling was set to one minute for this case study.

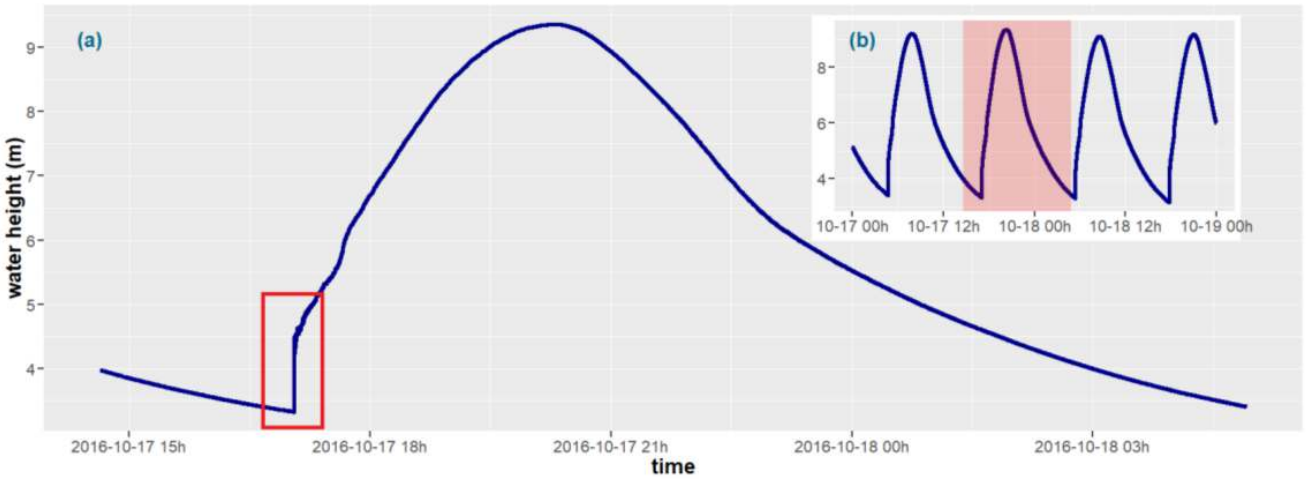

Figure 2. 1-Hz resampled pressure water level time series of the Garonne River at Podensac. (**a**) During the GNSS-R SNR acquisition with red rectangle indicating tidal bores; (**b**) during four consecutive tide periods. These figures are representative of the entire pressure water level time series and do not show the tidal oscillations over a longer period, as the acquisition was performed during spring tides only when both the tidal range is maximum and tidal bores can form.

2.3. Validation: Pressure Data

Several instruments were installed on the riverbed from 12 October to 20 October 2016, to monitor tidal bore occurrences during spring tides, including an Acoustic Doppler Current Profiler (ADCP, Nortek Signature 1000) and a pressure sensor (Ocean Sensor Systems). Water heights estimated with pressure measurements using the hydrostatic assumption were used for comparison with the GNSS-R SNR height estimation (Figure 2). The specifications from the constructor indicated an accuracy of 0.05% on the pressure measurements, corresponding to an accuracy of about 0.005 m on the water height determination at high tides in Podensac. Nonetheless, the maximum error could reach up to 0.10 m in extreme conditions. The nominal sampling rate of pressure measurements was 10 Hz. For validation purposes, we smoothed the pressure water height while averaging

values over a one minute interval to conform to the output sampling of GNSS-R height time series.

The whole anchorage containing the pressure and ADCP sensors was only referenced in longitude, latitude. No ellipsoidal height was provided, as it was installed to study temporal evolution of relative water height. For this reason, absolute comparisons between GNSS-R antenna heights and relative pressure water heights were impossible. Only relative comparisons, neglecting potential bias, were performed in this study.

3. Methods

3.1. Preprocessing

The SNR values represent the addition of the interferences of direct and reflected (multipath) signals components, which is written following [7]:

$$SNR^2 = A_d^2 + A_m^2 + 2A_dA_mcos(\psi), \quad (1)$$

where A_d and A_m are the amplitudes of the direct and multipath signals respectively, and ψ is the phase different between both signals. The term ${A_d}^2$ can be neglected since $A_m \ll A_d$ due to the attenuation upon reflection and the design of GNSS antennas. The direct component was modelled using a low-order polynomial function of elevation [23] then removed from the raw SNR time series. Here, as in most of the previous studies mentioned above, we used a 2nd order polynomial function. The remaining detrended SNR time series represents the multipath component corresponding to the reflected signals. Only specular points over the river were selected.

Due to the river characteristics in the study area (orientation and width, see Section 2.1), an azimuthal mask was applied to only select the tracks with an azimuth angle ranging from 10° to 150°. All elevation angles between 5° and 70° were retained, and higher values were masked by the platform on the top of which the antenna was installed. A small sample of reflected signals from low elevation could be affected by the presence of vegetation on the opposite riverbank. Nevertheless, we decided not to apply a more constraining mask on low elevations to maximize the number of available tracks for the inversion process.

Classical SNR-based retrieval methods use a limited elevation range (1–15°) to maximize the area of the first Fresnel zone [22]. This allows a more accurate height retrieval, as reflected signals are related to the crests of sea waves approximating a homogeneous surface. One frequency and height determination are performed for each single low-elevation track. In contrast, we used the dynamic SNR method [29] with a multisatellite approach, and no elevation mask was necessitated. This is a key point as very low-elevation tracks were not available due to vegetation masks and as we had a lesser number of satellites visible with constraining azimuthal configuration. We also performed a multifrequency analysis to improve the accuracy of the time series.

3.2. Dynamic SNR Inversion

According to [32], the multipath relative phase ψ of the detrended SNR time series is:

$$\psi = \frac{2\pi}{\lambda}\delta = \frac{4\pi h}{\lambda}sin(\varepsilon) \quad (2)$$

with δ the path delay, λ the signal wavelength, ε the satellite elevation and h the receiver height. The frequency of the multipath oscillations can be obtained deriving Equation (2) against time [29]:

$$f = \frac{d\psi}{dt} = \frac{4\pi\dot{h}}{\lambda}sin(\theta) + \frac{4\pi h}{\lambda}cos(\theta)\theta. \quad (3)$$

By making a change of variable $x = sin(\theta)$, we obtain:

$$\widetilde{f} = \frac{d\psi}{dx} = \frac{4\pi}{\lambda}\left(\dot{h}\frac{tan(\theta)}{\theta} + h\right). \quad (4)$$

The term with $\dot{h}$ can be neglected when $\dot{h} < 1 \times 10^{-6}$ m $\cdot$ s^{-1} [23]. Equation (4) is then simplified and the dominant frequency $\widetilde{f}$ is directly related to the antenna height h. This static case simplifies the computation assuming a very low elevation rate. In our case study, vertical velocities were much larger and the influence of the $\dot{h}$ component had to be considered.

Due to the characteristics of the study area, similar algorithms as the ones used for SL inversion (see [22] for a recent review) could not be used. Two main reasons constrained the retrieval of water heights with GNSS-R SNR in Podensac. First, the diurnal amplitude of the tides, which can reach up to 6 m, and the asymmetric tidal pattern, are responsible for high $\dot{h}$ especially during rising tides (see Figure 2), which is a major issue for most algorithms. Static SNR approaches applied to estimate SL use tide modelling to correct the $\dot{h}$ component or to reconstruct a complete time series from peaks and troughs (low $\dot{h}$). This is reliable when tides are symmetric, but not with asymmetric patterns. Second, the shape of the river in Podensac introduces a major constraint as reflections out of a 10–150° range fall out of the river (see Section 3.1). In most of the studies, the azimuthal range was near 360° and only observations with low elevations (<15°) were considered to maximize the extent of the first Fresnel zone and the accuracy of GNSS-R height retrieval. By contrast, in our study case it was mandatory to use a multisatellite, multielevation and multiGNSS frequency approach due to the lower number of available tracks.

The dynamic SNR method [29] offers an alternative technique which fits well with the problems we faced. A flowchart showing the initial method in blue, with our specific improvements in red, is presented in Figure 3. The basic idea was to estimate the dominant frequencies in a regularly spaced interval of time/elevation for all satellite tracks using a Lomb-Scargle Periodogram (LSP). These frequencies gathered into a moving window, filled an overdetermined system of equations following Equation (4), then h and $\dot{h}$ were jointly estimated using a Least Square Estimation (LSE, see Figure 3a). The parameter Δt, set by users, sampled the output time series of antenna heights by shifting the moving window in time. This approach was multisatellite because, from experience, the system of equations could only be solved if estimations from at least two distinct satellites were involved.

3.3. Improvements on the Dynamic SNR Approach

If we consider the dynamic approach as proposed by [29], the computed antenna heights exhibited very rough results (see Section 4.1). This was mainly due to the presence of noise in the dominant frequencies estimated using the LSP. As we considered the influence of $\dot{h}$ in the determination of $\widetilde{f}$, a larger range of frequencies were analyzed. This may affect the quality of retrieved $\widetilde{f}$ in several ways: low-frequencies due to the contribution of upper reflecting surfaces to the multipath signals (riverbanks, vegetation, the platform, boat stationed on the river) and high-frequency phenomena affecting the reflecting surface (tidal bores, boat wake: See Figure 1c). While the theoretical LSP outputs would present one single peak with high power, these aim to produce multi-peaks with equivalent power outputs, and a clear determination of the dominant frequencies is impossible. Illustrations of both a single peak and multipeaks outputs from LSP for the same satellite track are given in Figures 3b and 3c respectively.

Enough tracks are needed to perform a dynamic SNR inversion, so the application of more constraining elevation masks is impossible, as this would remove a major part of the data. Instead, we propose a method to filter out the noisy dominant frequencies in two steps (see red charts in Figure 3a). First, we look to filter out all frequencies associated with multipeaks considered as not reliable. Let $\widetilde{f_a}$ be the main peak in LSP with power P_a. We consider $\widetilde{f_a}$ being a multipeaks output if a frequency $\widetilde{f_b} \neq \widetilde{f_a}$ is found with peak power P_b that verifies:

$$P_b \geq k \times P_a, \tag{5}$$

with $0 \leq k \leq 1$. The value of parameter k is set empirically (see the results section). If k is low, a less noisy result is expected, but this may lead to gaps in the height determination. A 99% prediction interval (PI) is calculated based on single peak frequencies. This helps to extract the expected values in multipeaks LSP outputs and to fill the gaps resulting from filtering. If only one frequency is found in the PI, it is considered as the correct dominant frequency and extracted. In other cases, the multi-peaks are removed and a NA value is set at time index.

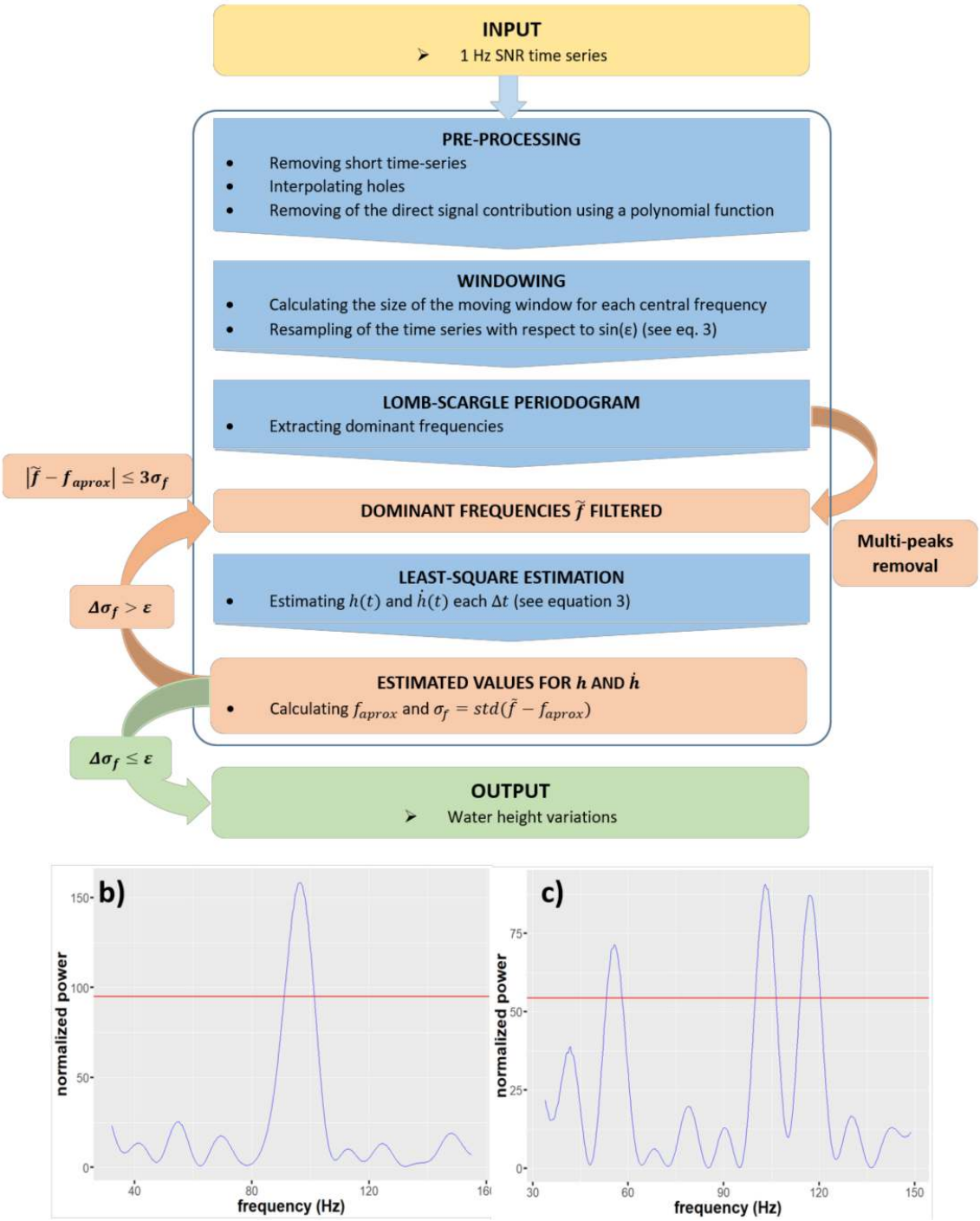

Figure 3. Flowchart of the dynamic SNR method with our improvements adapted from [29]. (**a**) Processing chain with the addition of a two-step filtering of the dominant frequencies (in orange); (**b**,**c**) respective examples of a single-peak output and a multipeak output from LSP for the same satellite track (G01). Red line materializes the level of filtering which depends on parameter k (here k = 0.6), and $\tilde{f}$ ~100 Hz.

The second step consists of the implementation of an iterative least square approach to estimate h instead of the former single LSE estimation. It is necessary because of the persistence of noise after filtering out the multipeaks. For each iteration, approximated values of h and $\dot{h}$ are computed with LSE [29]. Corresponding approximate frequencies f_{aprox} are calculated following Equation (4), and the standard deviation $\sigma_f = std\left(\widetilde{f} - f_{aprox}\right)$ is deduced from initial and estimated frequencies. Absolute errors between $\widetilde{f}$ and f_{aprox} are calculated and $\widetilde{f}$ is removed when:

$$\left|\widetilde{f} - f_{aprox}\right| > 3\sigma_f. \quad (6)$$

This corresponds to the filtering of the 1% worst values according to the Gaussian distribution. The standard deviation σ_f decreases while iterating, and a convergence criterion is set on its value. This allows removal of most of the outliers, with successive estimations of the approximated values of h, $\dot{h}$ and $\widetilde{f}$.

3.4. Validation

Pressure-derived water heights were used to validate the GNSS-R antenna heights. The absolute ellipsoidal height of the pressure sensor was not available which was the major issue in the validation procedure. Only a relative comparison between the dynamic SNR output and the pressure dataset was possible.

We further assumed that bias was zero, so the means of both datasets were adjusted for validation. Two metrics were then estimated: unbiased Root Mean Square Deviation (ubRMSD), adapted because of the no-bias assumption, and Pearson's correlation coefficient (R). We compute ubRMSD and R for the entire water height time series, and we also calculated metrics on subsampled times series to study the influence of both the vertical velocity and the number of satellites visible on the SNR heights.

Following the adapted dynamic SNR algorithm, we retrieved 1-min regularly spaced time series of antenna heights using a 5-min moving window. To set the empirical parameter k, several determinations were performed ranging from k = 0.5 (high filtering of multipeaks) to k = 1 (no filtering case). Qualitative and quantitative conclusions are highlighted in the following section. Results for separate L1, L2, L5 and combination of GNSS bands are also presented.

4. Results

4.1. Preliminary Filtering of the Dominant Frequencies

The estimation and filtering of dominant frequencies is a key point of our method to reduce noise and achieve a convincing height inversion. A good estimation of parameter k is necessary to first reject massive multi-peak outputs in LSP, then the iterative LSE approach filters out remaining noise in the dominant frequencies. With k too low, the loss of information leads to gaps in the output or to a miscalculation of h. With k too high, the filtering of the noise would be insufficient and cause variability in the output time series. The optimal value of k was empirically determined. It depended on the environment, particularly on the variations of the reflecting surface, the width of the river and the high-frequency phenomena affecting the SNR time series. In our case, the combination of these factors required strong filtering.

Figure 4 presents the frequencies estimated before and after these two filtering operations for a L1-only retrieval. As the occurrence of tidal bores (here at ~5:03 p.m.) affects the frequency estimation in the moving window, we decided to remove estimations using a one hour buffer around the phenomena. The second gap from 9:00 to 10:00 p.m. was due to a stop in the acquisition recording. Table 1 presents the statistical results obtained between each configuration of the adapted dynamic SNR algorithm (user-determined value of parameter k, GNSS bands used and minimal number of satellites for the inversion process) and the validation dataset. The values for R, ubRMSD, the maximum error and

the number of points calculated with LSE using the configuration (N_h) are shown. Further analysis based on statistical values refer to this table.

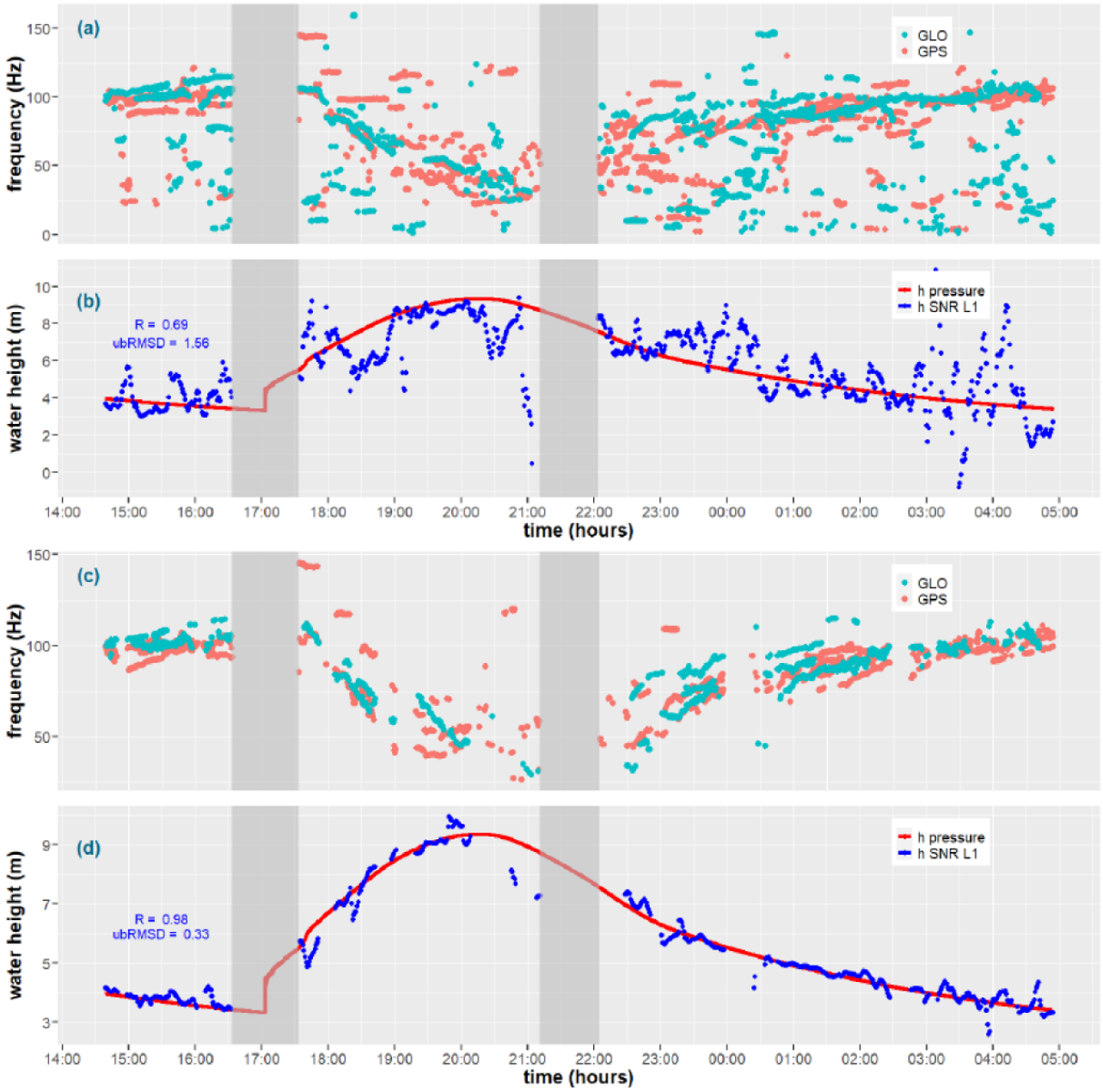

Figure 4. Dominant frequencies extracted from GPS (green) and GLONASS (orange) satellites using the LSP, and river heights inverted with the dynamic SNR method (blue) compared to pressure water levels (red). (**a**,**b**) $\widetilde{f}$ and h from raw LSP output respectively; (**c**) frequencies filtered out after multipeak rejection with parameter k = 0.6 and iterative LSE; (**d**) concordant time series of water levels estimated with iterative LSE. Grey areas are masks due to tidal bore occurrence (17 h) and data gaps (21–22 h).

The raw frequencies extracted with LSP (Figure 4a, with k = 1 and no iterative LSE process) show very noisy results particularly for the GLONASS constellation. The concordant estimation of h using LSE is fickle and highly affected by noise, as this is shown both graphically (Figure 4b) and statistically: R = 0.69, ubRMSD = 1.56 m, and the maximum error is 8.37 m. This presents the case of the nonadapted dynamic SNR method as proposed by [29] for SL retrieval. It confirms that the specific environment on rivers introduces noise in the SNR acquisition which has to be filtered. On the contrary, a much better result

was observed with our adapted solution (Figure 4c), considering k = 0.6 and iterative LSE process with at least four satellites visible. The heights inverted using these filtered frequencies compared well to the reference (Figure 4d), with R = 0.98, ubRMSD = 0.33 m, with a maximum error of 1.59 m. All maximum errors listed in Table 1 are high but this was mainly due to a poor satellite configuration as we will discuss in Sections 4.3 and 5.2.

Table 1. Performances of the adapted dynamic SNR inversion of h depending on the filtering level k, the minimal number of satellites, and the bands of frequency used. The best statistical results for each GNSS band/k parameter configuration were obtained with at least four satellites and new iterative LSE, while the worst results were obtained without the implementation of iterative LSE. Bold values highlight the best results for L1-only, L2-only and L1/L2/L5 height retrievals.

GNSS Bands Used	k	Iterative LSE	Min Number of Satellites	N_h—Number of deTerminations of h	Maximum Error (m)	R (Pearson)	ubRMSD (m)
L1	1	No	/	738	8.37	0.69	1.57
		Yes	2	735	10.45	0.73	1.64
			4	620	10.60	0.91	0.91
L1	0.90	No	/	733	3.82	0.90	0.85
		Yes	2	730	3.34	0.96	0.54
			4	643	3.32	0.97	0.47
L1	0.75	No	/	723	8.64	0.85	1.09
		Yes	2	720	8.46	0.88	0.96
			4	606	2.93	0.97	0.44
L1	**0.60**	No	/	715	20.64	0.83	1.19
		Yes	2	688	8.57	0.93	0.76
			4	**529**	**1.59**	**0.98**	**0.33**
L1	0.50	No	/	688	4.56	0.91	0.83
		Yes	2	660	5.28	0.93	0.72
			4	465	2.81	0.98	0.35
L2	**0.60**	No	/	702	21.60	0.77	1.48
		Yes	2	686	12.05	0.84	1.21
			4	**476**	**1.59**	**0.99**	**0.32**
L1, L2, L5	**0.60**	No	/	742	2.38	0.95	0.62
		Yes	2	741	2.16	0.98	0.44
			4	**662**	**2.08**	**0.99**	**0.31**

The optimal k for the Podensac study case is 0.6 (Table 1). This value seems adapted to Podensac, but rivers with larger width, SL determinations and antenna configurations with a large azimuthal range should accommodate with a higher value of k tending towards 1 (no filtering case). Several determinations were made using the L1 GNSS band, while varying its value from 0.5 to 1 and with or without an iterative LSE approach. When $k < 0.5$ too many frequencies were filtered out so height inversion was affected. Results without iterative LSE implementation were systematically worse whatever the value of k. The influence of the minimal number of satellites will be presented and discussed later. Finally, when considering an iterative LSE approach to filter out noise, a high correlation and low ubRMSD were found for k = 0.9, k = 0.75, k = 0.6 and k = 0.5, with the best values obtained for k = 0.6. The maximum error also helped us in analyzing the results: the lower value was found for k = 0.6 (maximum error = 1.59 m), with much larger values for all other outputs (maximum error $\geq$ 2.81 m).

The two-step filtering method drastically improved the quality of the height retrieval. One last parameter should be considered too: the number of determinations of h (N_h), i.e., the number of systems of equations successfully resolved using LSE while shifting time (with the same sampling parameters for every time series). This informs the potential loss of data when the level of filtering is high. When compared to the no-filtering case (738 points), the best results using L1 for k = 0.9, k = 0.75, k = 0.6 and k = 0.5 had N_h = 643

(87.1% of the total), N_h = 606 (82.1%), N_h = 529 (71.7%) and N_h = 465 (63.0%), respectively. This highlights the cost of the filtering approaches, with only ~70% of the time series being calculated with the best parameters.

4.2. Comparison Between L1, L2 and L5 GNSS Frequencies

The first results were obtained using the GPS and GLONASS L1 frequency only. To ensure the stability of the solution, the L1 band was first chosen because few satellites during the acquisition (2016) also operated at the L5 frequency. The quality of SNR measurements from L2C is good [33,34], but few GPS satellites (from Block IIR-M, 2005) are equipped with it, and the L2 P(Y) code is more sensitive to noise due to weaker power. Furthermore, penetration of the L1 signal (λ = 0.1903 m) compared to L2 (λ = 0.2445 m) and L5 (λ = 0.2540 m) signals in the vegetated riverbanks may change due to the wavelength difference. To prevent effects on studying reflections from heterogeneous surfaces, we first did a band-by-band inversion then compared results to the multiband analysis. The respective performances of L1 and L2 bands for GNSS-R altimetry are a subject of discussion [34].

The interest of the technique was assessed using the L1 band only (Figure 4d). Figure 5 presents final results using all bands: joint estimation of h (Figure 5a) and $\dot{h}$ (Figure 5b), and number of distinct GPS and GLONASS to solve the system of equations (Figure 5c). Figure 5a also superimposes the L1 only, L2 only and L1 + L2 + L5 height retrieval with the reference time series. Extended statistical results are found in Table 1. A qualitative analysis highlights the consistency of all the solutions calculated. Single band time series (L1 in orange, L2 in green) look noisy, especially when few satellites were visible: at ~9:00 p.m. (L1), ~11:00p.m. (L2) or ~0:30 a.m. (L1). On the contrary, they correlated well to each other and to the all-band solution (in blue) when optimal conditions were found, i.e., a high number of satellites and a low $\dot{h}$: before 4:30 p.m., and from 0:30 a.m. to 2:30 a.m. Looking globally, the three-time series were comparable with a bias below 0.03 m, and the wavelength gap from L1 to L2 and L5 bands did not affect the height retrieval.

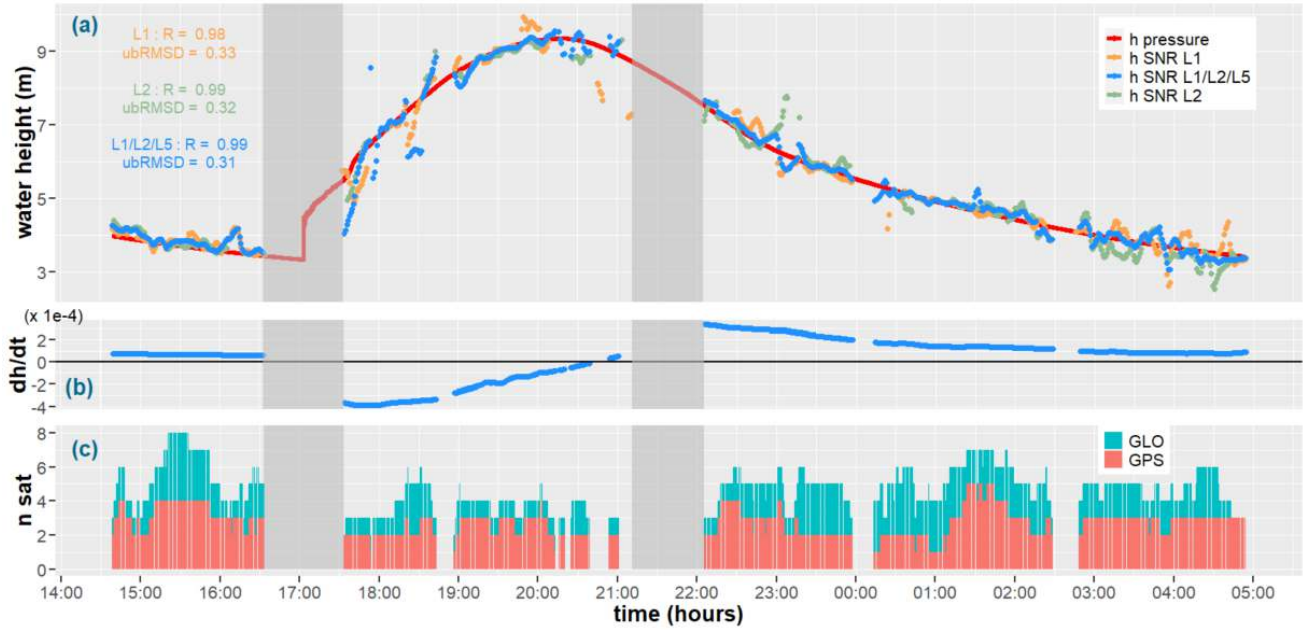

Figure 5. Final results using the adapted dynamic SNR inversion. (**a**) Comparison of h calculated using L1 only (orange), L2 only (green), and L1 + L2 + L5 (blue) frequencies for GPS and GLONASS satellites with pressure water levels (red); (**b**) output $\dot{h}$ with L1 + L2 + L5 bands; (**c**) number of GPS and GLONASS satellites for the calculation of h and $\dot{h}$. The value of $\dot{h}$ was derived from the relative antenna height h, thus it was negative during rising tides as the relative antenna height decreased.

Statistical results led to comparable conclusions. The correlation coefficients R were greater than 0.98 for every time series, and ubRMSD ranges from 0.31 m to 0.35 m, the best performance being obtained using all bands. The maximum errors were comparable for L1 and L2 (1.59 m), and surprisingly larger for the all-bands solution (maximum error = 2.08 m). This corresponds in reality to individual noisy points when few satellites were visible, e.g., between 6:00 p.m. and 7:00 p.m. (see Section 5.2 for details). Using the L1 band only, we obtained N_h = 535. Although presenting excellent statistical results, N_h = 484 for the L2-only case, confirmed that the absence of the L2 civilian frequency for some GPS satellites led to data gaps, especially when few satellites were visible. Finally, when combining all the L1, L2 and L5 bands in a unique system of equations, N_h = 664 was obtained. Compared to this result, the output time series of antenna height contained 19.4% fewer points for the L1-only case and 27.1% less for the L2, for quasi-identical statistical results. This loss of information was particularly accentuated when the configuration of the study area provided a restricted number of available satellite tracks for SNR based inversion, as in the Podensac case.

4.3. Influence of the Number of Satellites and Elevation Rate in the LSE Inversion

The number of GNSS satellites used to solve the system of equations with LSE (n_{sat}) appears to be a key point for the quality of the height retrieval. Following Equation (4), $\dot{h}$ also impacted the estimated frequencies and heights, as high $\dot{h}$ limits the application of static SNR retrieval methods [23]. In Figure 5c, the number of GPS (n_{gps}) and GLONASS (n_{glo}) satellites used for height inversion through the windowing of the dominant frequencies is presented. In Figure 5b, the vertical velocities $\dot{h}$ jointly estimated with h using LSE are shown.

Figure 5 is not sufficient to help us in separating the effects of $\dot{h}$ and n_{sat} on height retrieval. The best qualitative results for all bands configurations were obtained from 2:30 p.m. to 4:30 p.m., and after 11:00 p.m. This corresponds to both relatively low vertical velocities ($\dot{h} < 2 \times 10^{-4}$ m $\cdot$ s^{-1}) and high number of satellites: $n_{sat} = n_{gps} + n_{glo} \geq 4$. On the contrary, the worst results were obtained from 5:30 p.m. to 7:00 p.m., and correspond to high $\dot{h}$ combined with a poor satellites configuration. Due to the study area characteristics (see Section 2.1), the number of satellites visible was frequently lower than three, and $\dot{h}$ reached 4×10^{-4} m $\cdot$ s^{-1}. It was then necessary to dissociate both influences by studying each separately.

Figure 6 shows the R and ubRMSD values as a function either of the number of satellites (n_{sat}, Figure 6a) or the vertical velocity of the reflecting surface ($\dot{h}$, Figure 6b) in the case of a L1 + L2 + L5 combined SNR inversion. The deterioration of the height retrieval was clear for $n_{sat} < 4$. For $n_{sat} = 2$ and $n_{sat} = 3$, the correlations were 0.89 and 0.97 respectively, and the ubRMSD values were 0.97 m and 0.54 m. For $n_{sat} \geq 4$, the statistical results were homogeneous with R $\geq$ 0.96 and ubRMSD $\leq$ 0.32 m. We also notice that ubRMSD values fall to 0.12 m and 0.06 m respectively when n_{sat} is equivalent to 7 and 8 (on a very small sample). Following the same methodology, we calculated maximum errors of 0.39 m and 0.14 m for $n_{sat} = 7$ and $n_{sat} = 8$ respectively. Although the maximum errors values listed in Table 1 are high, a very coherent result was observed when a sufficient number of satellites was observed to ensure the stability of the solution. The major influence of the number of satellites in the dynamic SNR height retrieval was assessed. On the contrary, the variations of $\dot{h}$ from ~0 to 4×10^{-4}m $\cdot$ s^{-1} did not affect indisputably the quality of the height retrieval. For every $\dot{h}$ class extracted, ubRMSD values ranged from 0.35 m to 0.44 m and R from 0.87 to 0.99. A lower correlation was found for high $\dot{h}$ but ubRMSD was good, and the sample size is small. We conclude that using the dynamic SNR method, the influence of $\dot{h}$ (in the studied range) on the results was low.

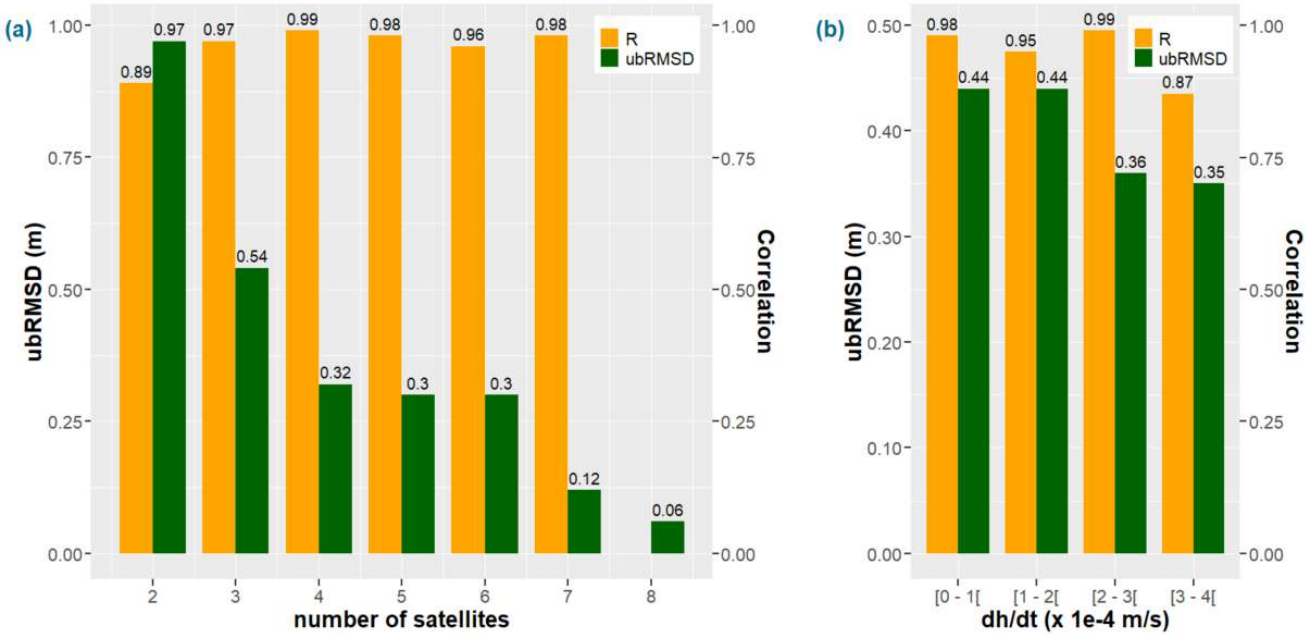

Figure 6. Statistical results depending on the number of satellites and the vertical velocity. (**a**) R and ubRMSD computed according to the number of satellites for the inversion of h and $\dot{h}$; (**b**) Idem according to the vertical velocity class (intervals computed with range 1×10^{-4} m.s^{-1}).

5. Discussion

5.1. Retrieving Water Heights in Rivers with GNSS-R

When compared to previous studies on GNSS-IR SL or water height retrieval, our work distinguishes itself by the complexity of the study area: constraining azimuthal masks, narrow river, vegetation and complex phenomena affecting the river surface. This led to two main difficulties. First, a fewer number of available satellite tracks than for SL studies, which employ only elevations lower than 15° [22], or use a receiver installed on top of a lighthouse for dynamic SNR inversion [29], with a near-360° azimuthal range and a multielevation analysis. Second, a deterioration of the quality of frequency analysis with LSP highly affected by noise, leading to wrong estimations of the heights whatever the retrieval algorithm.

The visibility of the maximal number of satellites is a key point for all GNSS-IR height retrievals, and is limited for all-case studies on the rivers. Ideally, the receiver should be installed in a meander with land on the north, to maximize the azimuthal range with reflections on the river. The configuration in Podensac is not ideal for GNSS-R acquisition regarding this criterion: the Garonne River forms a straight line and the platform on which the GNSS antenna is installed is on the southwest riverbank. On the other hand, the shape of the river makes it a perfect case study for the spatial structure of tidal bores [26,27,31] by conventional means including the pressure sensor that was used as a reference. Anyway, while working on rivers a reduction of the number of available satellite tracks is inevitable.

The second point, with much noise in the raw LSP frequency output (Figure 4a), may be considered as the particularity of the Podensac site. First, the occurrence of tidal bores during spring tides introduces high-frequency perturbations unrelated to the water height in the SNR time series and had to be removed from the analysis. Then, the narrow width of the river combined with the presence of vegetation on the riverbanks masks the low-elevation tracks, which are the most adapted for altimetry [22]. With a river cross-section of about 180 m only, riverbanks plus trees up 30 m on the opposite side, potentially affect reflections up to a 10–15° elevation. Furthermore, a small sample of the low-elevation tracks may be affected by the vegetation as a secondary reflecting surface, and low-frequency perturbations are then introduced in the signal. This affects only one or all GNSS bands, as L1 can slightly penetrate the canopy deeper due to the wavelength difference between L2 and L5.

5.2. Influence of the GNSS Band, the Number of Satellites Visible and the Vertical Velocity

We computed SNR heights using multiple configurations of our adapted dynamic method (see the Results section). Characterizing the differences introduced by a change in several parameters, i.e., the GNSS bands used, the number of satellites visible, n_{sat} and the vertical velocity $\dot{h}$, may lead to improvements in the quality of SNR-based height retrievals on rivers.

First, the use of the GNSS L1 or L2 band only did not show a specific deviation when compared to the multiband estimation (Figure 5a). Best results correlation ranged from 0.98 to 0.99 and ubRMSD ranged from 0.33 m to 0.31 m (Table 1). The single-band estimations contained some speckle noise in the output results, mainly due to a low number of distinct satellites visible, which was smoothed by the addition of L2 and L5 bands in the multifrequency analysis. Theoretically, the L2 P(Y) code is noisier and not appropriate for altimetry, while only GPS satellites from Block IIR-M (2005) and later blocks emit at L2C. Our results show that the noise in L2 SNR time series was removed from dominant frequency's estimation, with consequently a lesser number of retrieved heights for L2 than for L1. The stability of the results was guaranteed whatever GNSS band combination we used: the bias between L1 and L2 estimations was 0.04 m.

When considering the maximum errors between SNR based heights and pressure water levels, we noticed that the values for a single band output (maximum error = 1.59 m for both L1 and L2) were lower than the value for the all-bands output. These values were high but if we look more in detail, they correspond to bad determinations when few satellites were visible, e.g., from 5:30 p.m. to 7:00 p.m. (see Figure 5a). At this time the L1 heights seemed correct but there was mostly no L2-only output due to poor satellite configuration and the filtering operations on noisy frequencies. We think that the addition of speckle L2 (and eventually L5) frequencies in the L1-only system of equations reduced, in this particular case, the stability of the solution. However, looking both qualitatively (see Figure 5a) and quantitatively (ubRMSD = 0.31 m), the solution with all GNSS bands performed slightly better on the entire time series. The combination of distinct GNSS frequencies reduced the variability of the output, especially when a sufficient number of distinct tracks was available. Moreover, when compared to the time series inverted with all GNSS bands, the L1-only and L2-only solutions contained 19.4% and 27.1% less determinations, respectively. We, therefore, recommend a multifrequency analysis, which should be combined with an increasing number of distinct tracks in the SNR data, to maximize the number of independent observations in height inversion. A perspective for further studies is the potential of the L5 band for altimetry [3]. The low amount of GPS/Galileo satellites equipped with L5 emitters in 2016 prevented us from retrieving a L5-only height estimation, but it would be interesting for further work.

Results depending on the elevation rate looked very stable. In Figure 6b, results were separated in four classes depending on the value of $\dot{h}$ ranging from 0 to 4×10^{-4} m $\cdot$ s^{-1}. Analyzing these small, subsampled time series, we observed that the deterioration of statistical indicators (R and ubRMSD) did not seem related to the increase of $\dot{h}$ value as expected. Correlations were great and ubRMSD ranged from 0.35 m to 0.44 m with lower values for high $\dot{h}$ during the rising tide. Our method offers the possibility to cover rapid events up to at least 5×10^{-4} m $\cdot$ s^{-1} using GNSS-IR data.

With the major problem of GNSS-R retrievals on rivers being the azimuthal and elevation masks (see Section 5.1), a key point is the influence of the number of available tracks for SNR inversion (n_{sat}) on the final results. Figure 6a permits the analysis of statistical indicators for subsampled time series depending on n_{sat}, ranging from 2 to 8 in the Podensac case. Both R and ubRMSD showed an important deterioration for $n_{sat} < 4$ and stable results in other cases: R $\geq$ 0.96 and ubRMSD $\leq$ 0.32 m for $n_{sat} \geq 4$. We also note that when n_{sat} exceeded 6, the height comparison became excellent: ubRMSD = 0.12 m and maximum error = 0.39 m for n_{sat} = 7, and ubRMSD = 0.06 and maximum error = 0.14 m for n_{sat} = 8. Although definitive conclusions cannot be drawn due to the small sample size,

this result looks very promising for further work with better environmental conditions. An increasing number of satellites systematically improved all the statistical indicators we used in this study.

We are optimistic about the potential of the SNR height inversion as a survey method for rapid water height evolution, in tides or flash floods as an example, for two reasons. First, in most cases, a receiver configuration with less constraining azimuthal and elevation masks can be installed to increase the number of satellites visible and the accuracy of heights inverted with the dynamic SNR method. Second, a true multiconstellations approach including Galileo and Beidou satellites (limited interest in 2016) could now be performed, including a L5-only height retrieval based on GPS and Galileo.

5.3. The Dynamic SNR Method

Following the specificities of the study area and the velocity of the rising tides, we chose to employ the dynamic SNR method (see Sections 2.1 and 3.1). Would it be possible to retrieve equivalent or better results using a method that assumes a negligible $\dot{h}$ [23]? No dedicated study has been published, to our knowledge, on asymmetric phenomena with such a high amplitude as the tides in Podensac. Logical considerations make us think that these methods are not appropriate for a case study like Podensac. They are mostly employed on symmetric and predictable tides, and produce better results when amplitude is low [22]. The dominant frequencies are extracted with LSP, then heights are inverted assuming a quasi-static case ($\dot{h} < 1 \times 10^{-6}$ m $\cdot$ s^{-1}). Following this hypothesis only points in peaks and troughs of the tides can be correctly estimated. Various strategies help to reconstruct the complete time series: using tide models or functions (cubic splines) to fit the output SNR heights, and calculating a correction for the rate of change based on expected values. Furthermore, frequencies are extracted for each entire low-elevation subsampled satellite track, compromising the detection of rapid and unpredictable changes in the surface state with SNR time series.

On the contrary, the dynamic SNR method has arguments to deal with the problems enunciated above. First, the integration of the rate of change into a system of equations where h and $\dot{h}$ are jointly estimated with LSE being important: true values are directly extracted and no modelling or a priori knowledge of the phenomena is necessary. Second, the use of all-elevation SNR data (when not masked) as input, and the windowing of time series in the frequency estimations, produce finer detection of rapid changes in surface state. This is limited, as the influence of tidal bores, as an example, contaminated the results around 5:00 p.m. and had to be removed from the analysis (see Figure 4b,d and Figure 5a). However, this is an extreme case, being a highly nonlinear and extremely rapid ($\dot{h} \sim 1$ m $\cdot$ s^{-1}) phenomenon which occurs a few times a year. Third, the number of satellite tracks available can be a major issue while working on rivers, due to more constraining azimuthal and elevation masks, as in seas (see Section 4.3). The dynamic SNR method combines windowing with a multisatellite, multi-constellations and multi-frequency analysis, and provides good results with few satellites visible (Figure 5). This is of major interest as we look to extend the applications of GNSS-IR altimetry.

Following the idea in [25], we tried to add a new term to Equation (4) with the acceleration of the reflecting surface $\ddot{h}$. The LSE estimation then provides all h, $\dot{h}$ and $\ddot{h}$ values along with the dominant frequencies as inputs. The results for h and $\dot{h}$ looked like the without-acceleration case, with a small influence from acceleration. However, we noted an addition of speckle noise in the output h time series, so we finally rejected the idea. In [25], $\ddot{h}$ was calculated from true values or models. In our case, estimation was made from noisy frequencies based on the SNR time series (see Figure 4a). The derivative of h already adds noise in Equation (4). When deriving a second time for $\ddot{h}$ the noise level becomes prominent for a low gain, so including the acceleration has no advantage.

The noisy SNR time series in Podensac composed a challenging dataset. We were able to clean up the extracted dominant frequencies to retrieve h and $\dot{h}$ using two levels

of filtering noise: first removal of the multipeaks in the LSP output that materialize other reflection surfaces or high frequency perturbations, then filtering of frequencies up to 3σ (99%) with an iterative approach of LSE. These additions to the initial method [29] were mandatory to work with GNSS-IR on very constraining river environments, of which Podensac is a perfect example.

A key last point is the sampling of output SNR height time series: the results presented in this study had a 1-min sampling period with the size of the moving window for the LSE of h and $\dot{h}$ being 5 min. The choice of a tiny window was made in opposition to the literature, where the lowest window size found was 2 h [25], applied to detect abrupt water-level changes in a reservoir with a 10-min sampling time series of heights. We thus showed that the use of the dynamic SNR approach provided a very high temporal resolution when compared to other methods. We could smooth the results over a longer time span but our thoughts are that this approach offers a distinct and complementary solution to other SNR retrieval algorithms, and can be used to extend the potential applications of GNSS-IR altimetry. Combining the high temporal resolution, its independence with any tide model or fitting function, and the ability to detect surface changes with high $\dot{h}$ (at least 5×10^{-4} m $\cdot$ s^{-1}), it opens the field to detection and monitoring of extreme events such as flash floods, marine surges in estuaries or coastal areas, tsunamis and storms.

6. Conclusions

In this study, SNR time series acquired over a narrow river were analyzed to retrieve the evolution of water height during asymmetric tides. The characteristics of the study area in Podensac made it a complicate environment for GNSS-R studies: the river shape forms a constraining azimuthal mask and disrupt low elevation signals, the tidal range is high (~6 m) and other high-frequency phenomena (i.e., tidal bores) affect the water surface and, thereby, the SNR time series during the acquisition.

The dynamic SNR method was chosen because it considers the dynamics of the reflecting surface during the height inversion [29], which is a key point for studying asymmetric and nonpredictable water level variations. The rising tide vertical velocity reached 5×10^{-4} m $\cdot$ s^{-1}, which is far more than the limit set for a static case (1×10^{-6} m $\cdot$ s^{-1} [23]). As the dominant frequencies estimated with LSP suffered noise, introduced in SNR time series by the river environment, we added a two-step filtering method to the dynamic inversion. First, the multipeaks were eliminated in the LSP output, then we iterated over the least square estimation (LSE) of h and $\dot{h}$ to remove the frequency values showing the greatest error when compared to estimation. This improved the results considerably when compared to the no-filtering inversion.

The best results obtained versus reference pressure water heights were consistent whatever the GNSS band we used: ubRMSD = 0.33 m for L1, ubRMSD = 0.32 m for L2, ubRMSD = 0.31 m for L1, L2 and L5 combined, with a correlation value higher than 0.98. The raw output without filtering the dominant frequencies showed ubRMSD = 1.56 m for L1. When combining all GNSS bands, a more robust estimation was found when enough satellites were visible; therefore it increased the number of samples in the output time series. Other factors are discussed including the influence of the vertical velocity of the reflecting surface and the number of distinct satellite tracks used in the LSE inversion. The influence of $\dot{h}$ on both R and ubRMSD values was low in the analyzed $0 - 5 \times 10^{-4}$ m $\cdot$ s^{-1} span. The dynamic SNR method fitted the reference data well, whatever the elevation rate of the reflecting surface. On the contrary, the number of tracks available had a major impact on the results. R and ubRMSD were consistent when $n_{sat} \geq 4$ with values greater than 0.96 and lower than 0.33 m, respectively. A clear degradation was observed for $n_{sat} = 2$ (R = 0.89, ubRMSD = 0.97 m) or $n_{sat} = 3$ (R = 0.97, ubRMSD = 0.54 m). Finally, we obtained a great result for a small subsample when $n_{sat} = 7$ (ubRMSD = 0.12 m) and $n_{sat} = 8$ (ubRMSD = 0.06 m). This latest point would have to be confirmed with longer time series analysis, but it is very promising for further work. Studies on rivers wider than the Garonne (~150 m), with more

GNSS constellations (including Galileo, Beidou) and frequencies (recent developments of L5 band) would provide more available satellite tracks.

The windowing of the dynamic SNR method allows us to perform analyses with good temporal resolution when compared to previous methods. To take advantage of this point the sampling of our output time series of antenna heights is 1 min, and the size of the moving window is 5 min. The results exhibited good variations due to the rapid rising tide. It was also more subject to noise, because for tide modelling, at least a 2-h window is employed, which smooths the output time series. By contrast, this highlights the value of the dynamic SNR method to retrieve unpredictable water levels during extreme events with high elevation rates.

Author Contributions: Conceptualization, P.Z., F.F. and J.D.; methodology, P.Z.; software, P.Z. and N.R.; validation, P.Z., F.F., J.D. and P.B.; formal analysis, P.Z.; investigation, P.Z.; resources, P.Z., F.F. and J.D.; data curation, F.F., J.D., P.B., N.B. and G.D.; writing—original draft preparation, P.Z.; writing—review and editing, P.Z., F.F. and J.D.; visualization, P.Z.; supervision, F.F. and J.D.; project administration, F.F. and J.D.; funding acquisition, F.F. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by CNES/EUMETSAT OSTST FOAM, PNTS GRAEL and URGENS.

Data Availability Statement: All the data used for SNR processing and validation in this study are fully available on demand. Please contact the authors (P.Z., F.F. or J.D.) for any related question.

Acknowledgments: This study is dedicated to the memory of Richard Biancale who supported the development of GNSS-R activities at OMP and who funded this campaign through the CNES TOSCA GRGS project. We also thank the reviewers for their helpful comments and contributions to improve this work.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Bevis, M.; Businger, S.; Herring, T.A.; Rocken, C.; Anthes, R.A.; Ware, R.H. GPS Meteorology: Remote Sensing of Atmospheric Water Vapor Using the Global Positioning System. *J. Geophys. Res. Atmos.* **1992**, *97*, 15787–15801. [CrossRef]
2. Wickert, J.; Cardellach, E.; Martin-Neira, M.; Bandeiras, J.; Bertino, L.; Andersen, O.B.; Camps, A.; Catarino, N.; Chapron, B.; Fabra, F.; et al. GEROS-ISS: GNSS REflectometry, Radio Occultation, and Scatterometry Onboard the International Space Station. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2016**, *9*, 4552–4581. [CrossRef]
3. Lestarquit, L.; Peyrezabes, M.; Darrozes, J.; Motte, E.; Roussel, N.; Wautelet, G.; Frappart, F.; Ramillien, G.; Biancale, R.; Zribi, M. Reflectometry With an Open-Source Software GNSS Receiver: Use Case With Carrier Phase Altimetry. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2016**, *9*, 4843–4853. [CrossRef]
4. Hall, C.D.; Cordey, R.A. Multistatic Scatterometry. In Proceedings of the IEEE International Geoscience and Remote Sensing Symposium, Edinburgh, UK, 12–16 September 1988; pp. 561–562.
5. Martin-Neira, M. A Passive Reflectometry and Interferometry System (PARIS): Application to Ocean Altimetry. *ESA J.* **1993**, *17*, 331–335.
6. Kavak, A.; Vogel, W.J.; Xu, G. Using GPS to Measure Ground Complex Permittivity. *Electron. Lett.* **1998**, *34*, 254–255. [CrossRef]
7. Larson, K.M.; Small, E.E.; Gutmann, E.D.; Bilich, A.L.; Braun, J.J.; Zavorotny, V.U. Use of GPS Receivers as a Soil Moisture Network for Water Cycle Studies. *Geophys. Res. Lett.* **2008**, *35*, L24405. [CrossRef]
8. Rodriguez-Alvarez, N.; Camps, A.; Vall-Llossera, M.; Bosch-Lluis, X.; Monerris, A.; Ramos-Perez, I.; Valencia, E.; Martinez-Fernandez, J.; Baroncini-Turricchia, G.; Perez-Gutierrez, C.; et al. Land Geophysical Parameters Retrieval Using the Interference Pattern GNSS-R Technique. *IEEE Trans. Geosci. Remote Sens.* **2011**, *49*, 71–84. [CrossRef]
9. Chew, C.C.; Small, E.E.; Larson, K.M.; Zavorotny, V.U. Effects of Near-Surface Soil Moisture on GPS SNR Data: Development of a Retrieval Algorithm for Soil Moisture. *IEEE Trans. Geosci. Remote Sens.* **2014**, *52*, 537–543. [CrossRef]
10. Roussel, N.; Frappart, F.; Ramillien, G.; Darrozes, J.; Baup, F.; Lestarquit, L.; Ha, M.C. Detection of Soil Moisture Variations Using GPS and GLONASS SNR Data for Elevation Angles Ranging from 2° to 70°. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2016**, *9*, 4781–4794. [CrossRef]
11. Zhang, S.; Calvet, J.C.; Darrozes, J.; Roussel, N.; Frappart, F.; Bouhours, G. Deriving Surface Soil Moisture from Reflected GNSS Signal Observations from a Grassland Site in Southwestern France. *Hydrol. Earth Syst. Sci.* **2018**, *22*, 1931–1946. [CrossRef]
12. Larson, K.M.; Gutmann, E.D.; Zavorotny, V.U.; Braun, J.J.; Williams, M.W.; Nievinski, F.G. Can We Measure Snow Depth with GPS Receivers? *Geophys. Res. Lett.* **2009**, *36*, L17502. [CrossRef]
13. Rodriguez-Alvarez, N.; Aguasca, A.; Valencia, E.; Bosch-Lluis, X.; Camps, A.; Ramos-Perez, I.; Park, H.; Vall-llossera, M. Snow Thickness Monitoring Using GNSS Measurements. *IEEE Geosci. Remote Sens. Lett.* **2012**, *9*, 1109–1113. [CrossRef]

14. Small, E.E.; Larson, K.M.; Braun, J.J. Sensing Vegetation Growth with Reflected GPS Signals. *Geophys. Res. Lett.* **2010**, *37*, L12401. [CrossRef]
15. Zhang, S.; Roussel, N.; Boniface, K.; Cuong Ha, M.; Frappart, F.; Darrozes, J.; Baup, F.; Calvet, J.C. Use of Reflected GNSS SNR Data to Retrieve Either Soil Moisture or Vegetation Height from a Wheat Crop. *Hydrol. Earth Syst. Sci.* **2017**, *21*, 4767–4784. [CrossRef]
16. Anderson, K.D. Determination of Water Level and Tides Using Interferometric Observations of GPS Signals. *J. Atmos. Ocean. Technol.* **2000**, *17*, 1118–1127. [CrossRef]
17. Larson, K.M.; Ray, R.D.; Nievinski, F.G.; Freymueller, J.T. The Accidental Tide Gauge: A GPS Reflection Case Study From Kachemak Bay, Alaska. *IEEE Geosci. Remote Sens. Lett.* **2013**, *10*, 1200–1204. [CrossRef]
18. Löfgren, J.S.; Haas, R. Sea Level Measurements Using Multi-Frequency GPS and GLONASS Observations. *EURASIP J. Adv. Signal Process.* **2014**, *2014*, 50. [CrossRef]
19. Vu, P.L.; Ha, M.C.; Frappart, F.; Darrozes, J.; Ramillien, G.; Dufrechou, G.; Gegout, P.; Morichon, D.; Bonneton, P. Identifying 2010 Xynthia Storm Signature in GNSS-R-Based Tide Records. *Remote Sens.* **2019**, *11*, 782. [CrossRef]
20. Purnell, D.; Gomez, N.; Chan, N.H.; Strandberg, J.; Holland, D.M.; Hobiger, T. Quantifying the Uncertainty in Ground-Based GNSS-Reflectometry Sea Level Measurements. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2020**, *13*, 4419–4428. [CrossRef]
21. Tabibi, S.; Geremia-Nievinski, F.; Francis, O.; van Dam, T. Tidal Analysis of GNSS Reflectometry Applied for Coastal Sea Level Sensing in Antarctica and Greenland. *Remote Sens. Environ.* **2020**, *248*, 111959. [CrossRef]
22. Geremia-Nievinski, F.; Hobiger, T.; Haas, R.; Liu, W.; Strandberg, J.; Tabibi, S.; Vey, S.; Wickert, J.; Williams, S. SNR-Based GNSS Reflectometry for Coastal Sea-Level Altimetry: Results from the First IAG Inter-Comparison Campaign. *J. Geod.* **2020**, *94*, 70. [CrossRef]
23. Larson, K.M.; Löfgren, J.S.; Haas, R. Coastal Sea Level Measurements Using a Single Geodetic GPS Receiver. *Adv. Space Res.* **2013**, *51*, 1301–1310. [CrossRef]
24. Beckheinrich, J.; Hirrle, A.; Schon, S.; Beyerle, G.; Semmling, M.; Wickert, J. Water Level Monitoring of the Mekong Delta Using GNSS Reflectometry Technique. In Proceedings of the 2014 IEEE Geoscience and Remote Sensing Symposium, Quebec City, QC, Canada, 13–18 July 2014; pp. 3798–3801.
25. Tabibi, S.; Francis, O. Can GNSS-R Detect Abrupt Water Level Changes? *Remote Sens.* **2020**, *12*, 3614. [CrossRef]
26. Bonneton, P.; Bonneton, N.; Parisot, J.-P.; Castelle, B. Tidal Bore Dynamics in Funnel-Shaped Estuaries. *J. Geophys. Res. Ocean.* **2015**, *120*, 923–941. [CrossRef]
27. Martins, K.; Bonneton, P.; Frappart, F.; Detandt, G.; Bonneton, N.; Blenkinsopp, C.E. High Frequency Field Measurements of an Undular Bore Using a 2D LiDAR Scanner. *Remote Sens.* **2017**, *9*, 462. [CrossRef]
28. Frappart, F.; Roussel, N.; Darrozes, J.; Bonneton, P.; Bonneton, N.; Detandt, G.; Perosanz, F.; Loyer, S. High Rate GNSS Measurements for Detecting Non-Hydrostatic Surface Wave. Application to Tidal Borein the Garonne River. *Eur. J. Remote Sens.* **2016**, *49*, 917–932. [CrossRef]
29. Roussel, N.; Ramillien, G.; Frappart, F.; Darrozes, J.; Gay, A.; Biancale, R.; Striebig, N.; Hanquiez, V.; Bertin, X.; Allain, D. Sea Level Monitoring and Sea State Estimate Using a Single Geodetic Receiver. *Remote Sens. Environ.* **2015**, *171*, 261–277. [CrossRef]
30. Vu, P.-L.; Frappart, F.; Darrozes, J.; Ha, M.-C.; Dinh, T.-B.-H.; Ramillien, G. Comparison of Water Level Changes in the Mekong River Using Gnss Reflectometry, Satellite Altimetry and in-Situ Tide/River Gauges. In Proceedings of the IGARSS 2018—2018 IEEE International Geoscience and Remote Sensing Symposium, Valencia, Spain, 22–27 July 2018; pp. 8408–8411.
31. Bonneton, P.; Filippini, A.G.; Arpaia, L.; Bonneton, N.; Ricchiuto, M. Conditions for Tidal Bore Formation in Convergent Alluvial Estuaries. *Estuar. Coast. Shelf Sci.* **2016**, *172*, 121–127. [CrossRef]
32. Bishop, G.J.; Klobuchar, J.A.; Doherty, P.H. Multipath Effects on the Determination of Absolute Ionospheric Time Delay from GPS Signals. *Radio Sci.* **1985**, *20*, 388–396. [CrossRef]
33. Strandberg, J.; Hobiger, T.; Haas, R. Improving GNSS-R Sea Level Determination through Inverse Modeling of SNR Data: GNSS-R INVERSE MODELING. *Radio Sci.* **2016**, *51*, 1286–1296. [CrossRef]
34. Santamaría-Gómez, A.; Watson, C.; Gravelle, M.; King, M.; Wöppelmann, G. Levelling Co-Located GNSS and Tide Gauge Stations Using GNSS Reflectometry. *J. Geod.* **2015**, *89*, 241–258. [CrossRef]

3.3 Comments on the article published in *Remote Sensing*

Several points were noted by the reviewers that require further explanations. This will be treated here as the article has been published for almost two years. In particular, the short duration of the SNR time series (∼13 hours) was a huge constraint for us, as the acquisition was first designed to study tidal bores - *i.e.* a very high-frequency phenomena lasting only few minutes. We were lucky enough to have a full tidal cycle recorded to study water level variations during asymmetric tides in the Garonne River. However, we are aware of the limitations of our study. The method was developed to fit this limited dataset, but it should be evaluated on longer time series and possibly in other study areas. This was first intended in a coastal environment with the GNSS station from the french permanent geodetic network located in the fort of Socoa, in Saint-Jean-De-Luz. Early results have shown a slightly better performance from our adapted method than the former dynamic SNR method from Roussel et al. (2015), in the retrieval of water height variations with a rough sea during the storm season. For this reason, we installed a receiver in the port of Bilbao to further try to decompose the sea level and sea state components in the SNR time series (see Section 3.5. Over inland water, longer time series over rivers could validate the technique before diving into the detection of flash floods.

Further comments were required on the methodology, in particular on the iterative least square estimation (LSE) employed to filter noisy frequencies and estimate the correct antenna height (h). As for the initial dynamic SNR method (Roussel et al., 2015), the time series of h are inverted from the dominant frequencies using a moving window whose length was set to 5 minutes. The exact same moving window was then used to estimate f_{aprox} from the output h and $\dot{h}$. In a perfect case study with no noise, f_{aprox} should systematically be equal to $\tilde{f}$ and $\sigma_f = std(\tilde{f} - f_{aprox})$ should be equal to zero. As it was not the case, we filtered out the samples where $|\tilde{f} - f_{aprox}| \geq 3\sigma_f$. This was performed in a *while loop*, with a convergence criterion ϵ set on the variations of σ_f to stop the iterations. Several values were tested for this convergence criterion, and it was finally set to 10^{-2} Hz. This ensures a stable final solution and the iteration were not very time-consuming (regarding to our short time series). The Table 1 our published article (Section 3.2) highlights the very large errors obtained on the output water levels when the two-step filtering method was not implemented, and also when only the multi-peaks removal was performed but not the iterative LSE. On the contrary, results are correct using the two-step filtering method and with at least four satellites in visibility (maximum error ∼1.5 m and ubRMSD ∼30 cm). Note that the results should be smoother using a large window size, what was usually done in the literature. However, we aimed at proving that the SNR techniques can be adapted to the monitoring of abrupt and large amplitude water level changes, and high frequency phenomena such as flash floods.

Finally, we showed that the number of GNSS satellites in visibility mostly explains the variability of the results. The configuration of the acquisition in Podensac was not optimal at all: the river width is about 150 m and 20 m tall trees cover the riverbanks. Moreover, the platform where the antenna was set up is located on the southern side of the Garonne, with no GNSS satellite in visibility at North (between 350° and 10° azimuth). Hence, only

reflections on a ∼140° azimuth range and above 15° of elevation were monitored in this study. Moreover, we processed only the GPS and GLONASS data. Now, with the Galileo and BeiDou constellations fully operational, and with attention paid to the configuration of the acquisition, the number of satellites in visibility could be increased to improve the results.

3.4 Perturbation of the water surface during tidal bores

The early objective of this publication was to evaluate whether SNR time series are sensitive to the propagation of tidal bores, and whether they can be used to retrieve the simultaneous evolution of water height. For this, I had the ∼12 hrs multi-frequency and multi-constellation 20 Hz SNR time series, acquired the 17/10/2016 over the Garonne River in Podensac. Multiple analysis were conducted on this dataset, with no success for estimating water levels with sufficient accuracy during tidal bores. The SNR techniques are limited to the monitoring of surface water variations with a much lower vertical velocity. However, I found high-frequency perturbations of the SNR time series occurring exactly when the tidal bore propagates. Here, I show some unpublished results about this early work.

First, the Lomb-Scargle periodogram (LSP) used in both the static (Larson et al., 2013a; Alonso-Arroyo et al., 2014b) and dynamic (Roussel et al., 2015) SNR methods is inherently limited. As the frequency of SNR oscillations related to the antenna height is generally low, a wide frequency spectrum should be analyzed with the LSP so long time series are required. For an antenna height between 4 and 12 meters, as in Podensac, the shortest time series analyzed lasted about 30 minutes. However, the SNR acquisition was sampled at 20 Hz and the tidal bore wave trains lasted 1-3 minutes (Figure 3.1). Short-duration high frequencies are thus smoothed in the LSP. To overcome this limitation, I used the wavelet transforms and the cross-wavelet analysis techniques (Torrence and Compo, 1998).

The continuous wavelet transform of a time series x_n analyzes the convolution of x_n with the wavelet function, that is scaled in the frequency domain, and translated in the time series. This convolution results in the output wavelet power that is mapped in two dimensions: time and frequency/period. We used the conventional Morlet wavelet because it comprises both a real and an imaginary parts, and thus provides the phase along with the output wavelet power. The main advantage of this method is that it windows the time series with a small wavelet used for high-frequency/short period phenomena, and larger wavelets used for lower frequencies. So, it permits to accurately detect high amplitude and high frequency perturbations of the multipath SNR time series.

Figure 3.2a presents the wavelet transform power of the multipath SNR for satellite G03, with tidal bore perturbations occurring after 17h. The large peak power obtained for a period varying between ∼250 s and ∼100 s (depending on the satellite elevation, not shown) corresponds to the oscillations related to the receiver antenna height. This dominant frequency was also extracted using the usual LSP analysis (Larson et al., 2013a; Roussel et al., 2015). A 10-minute zoom presenting the wavelet power before, during and after the propagation of the tidal bore waves is shown in Figure 3.2b. Figure 3.2c presents the cross-wavelet analysis

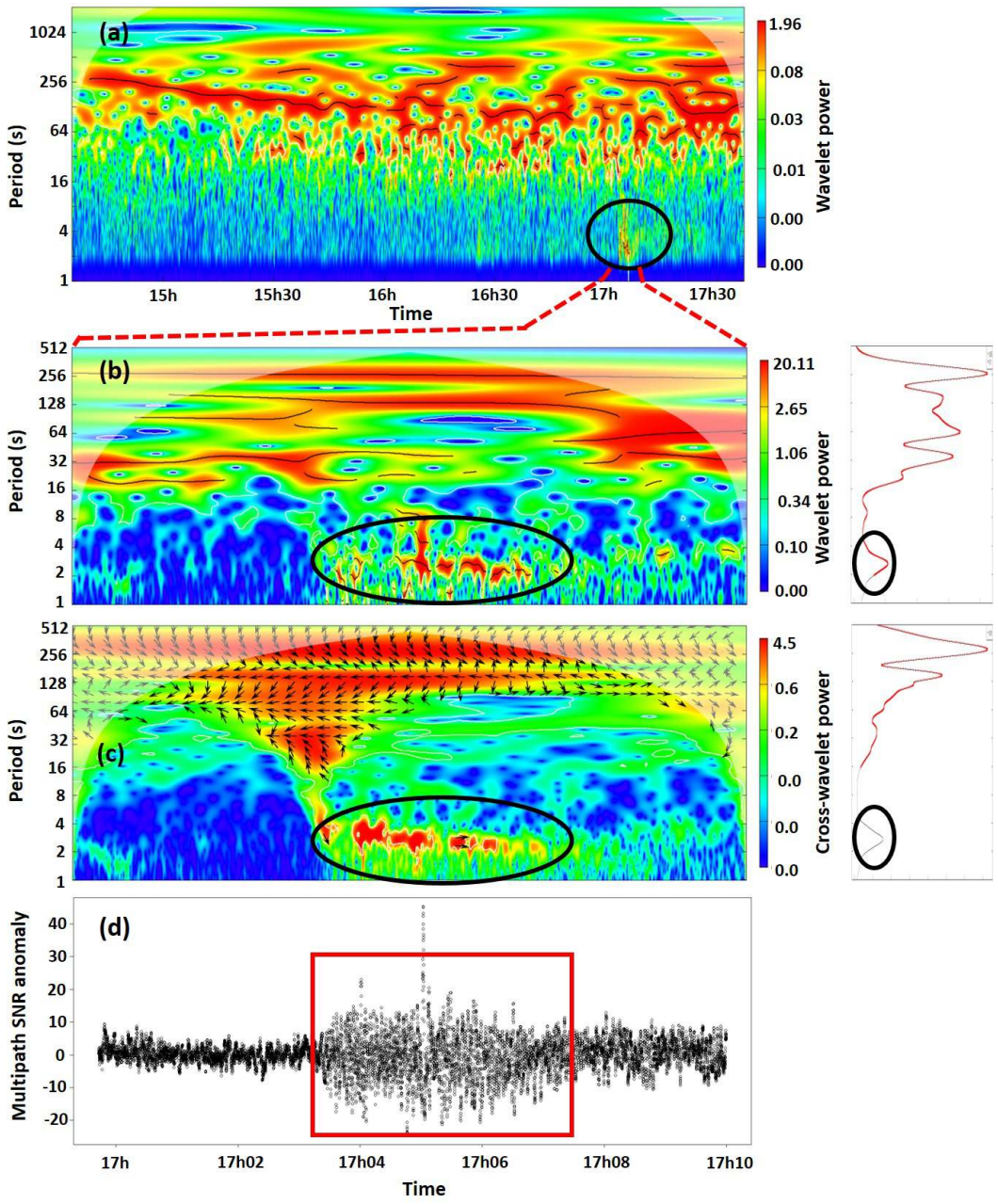

Figure 3.2: Use of the wavelet and cross-wavelet analysis to extract frequencies in SNR time series: example from the satellite G03 in Podensac, the 17/10/2016. **(a)** Wavelet transform power extracted from the multipath SNR, highlighting the dominant low-frequencies (high period) and the tidal bore starting around 17h. **(b)** Same plot between 17h and 17h10 to show the tidal bore high frequencies, and average power per wavelet period (right panel). **(c)** Cross-wavelet analysis with the time series of water heights during tidal bore passage (Figure 3.1), and average cross-wavelet power per period (right panel). **(d)** High frequency (period <8 s) anomalies in the multipath SNR, computed with the reconstructed time series using periods greater than 8 seconds in the wavelet transform. The circles and squares highlight the influence of tidal bores on the multipath signal.

of the same time series against DGNSS water heights computed from the buoy at the surface (Figure 3.1). The wavelet and cross-wavelet maps clearly show a high power with a period of ∼2.5 - 3 s just when the tidal bore propagates in the river (starting at ∼17h03m20s). Furthermore, this corresponds exactly to the period of the waves as illustrated in Figure 3.1. Finally, Figure 3.2d presents a reconstruction of the SNR multipath time series using only periods greater than 8 s. It shows a continuous noise before the occurrence of the waves, and high amplitude SNR anomalies related to the tidal bore propagation. To conclude with these results, here are several interesting points for future studies:

- The cross-wavelet analysis highlights a clear influence of the surface water height variations on the high frequency and high amplitude oscillations of multipath SNR time series, during the propagation of tidal bores.
- The wavelet tools can help in detecting high frequency signals in the SNR time series. It also shows how all frequencies change in time. Furthermore, the computation of a wavelet transform was found to be much faster than a usual LSP analysis.
- The high frequency perturbations of multipath SNR (abrupt increase in the SNR variability) correspond to the period of tidal bore waves. However, the multipath oscillations related to the antenna height over a stable water surface have a period of ∼100 s for this satellite around 17h. So, the water levels cannot be extracted directly following Equation (2.28) or Equation (2.27).

The last point represents an inherent limitation of the SNR method, due to the absence of complex I/Q information. I tried to extract the multipath phase with a least square estimation using Equation (2.29), in order to estimate the water level variations through phase unwrapping. In theory this could work, as the observations sampled at 20 Hz avoid a phase jump even with a vertical velocity of ∼1 m/s. However, this method was not successful mainly due to discontinuities in the tidal bore frequencies. Further investigations could also determine whether a high frequency waveform GNSS-R acquisition can be used to retrieve extreme variations of water level.

3.5 Contribution to this work and perspectives

The data acquisition was performed in 2016, way before the beginning of my PhD. It was thus a latent project just waiting for someone to process and analyze the data, as I finally did during my first year. This was my primary contribution to the work presented above. When diving deeper into the subject, I found that although the water height retrieval during the propagation of tidal bores was very complicated, it was possible to monitor the water heights during high amplitude asymmetric tides in the Garonne River. The direction of the study then shifted, and I modified the algorithm from Nicolas Roussel to accommodate for the monitoring of inland water level variations. Finally, I performed the analysis and the writing of the publication in *Remote Sensing*, as it is presented above.

Following this work, we also installed several conventional receivers for further studies. The modified dynamic SNR method, as presented in this work, can be very useful to analyze

the resulting datasets. The receivers are dedicated to study both inland water and coastal environments. For inland water, one is located in Salsigne (Aude, France) with consequent topography around. The GNSS-R configuration is thus particularly complicated, with important elevation and azimuth masks (see Figure 3.3). The methods I implemented in Section 3.2 to filter out noisy observations can be very useful in this case study. The receiver is dedicated to study abrupt water level rises in small rivers of the Orbiel catchment, as it occurred during devastating flash floods in 2018. The water surface suddenly rose from about 8 meters, causing numerous deaths, destruction of goods, and a large pollution in the river basin due to the washing of former arsenic mines.

We also installed an antenna in the port of Bilbao (Spain), complementing other receivers available in Saint-Jean-De-Luz, Biarritz and Cordouan (France) in the Bay of Biscay. These datasets were already used to monitor SSH, following the methods presented in Section 2.4.2.3. Further studies will be dedicated to sea state (Roussel et al., 2015), particularly during storms. The receiver in Bilbao is there particularly interesting. It was installed at the extremity of the pier constructed to protect the port from swell and large waves during storms (Figure 3.4). The pier is 2 km long and is located at about 7 km from the city of Bilbao. The receiver height is above 20 m, providing an optimal case study with very few azimuth and elevation masks (see Figure 3.4). It should allow to retrieve separately SSH inside and outside the area protected by the pier. The SSH inside the port should be reliable with smaller influence from the sea state, while we expect the reflections outside the port to provide information about the significant wave height (SWH) and the frequency of waves. Two data acquisitions have been configured: one with a 30 s time sampling to retrieve SSH, the second one with a higher 1 s sampling to study sea state.

The stations in Salsigne and Bilbao now record observations that will be further processed. I started this first part of my PhD with data from a former acquisition, and as the method went well, I participated in acquiring data for further studies. The observations collected by the receiver in Salsigne particularly need a very intense flooding event to be evaluated. Thus ended my work on the SNR techniques and applications, and I started the second part of my PhD on analyzing CYGNSS data at global scale to study the inundation dynamics.

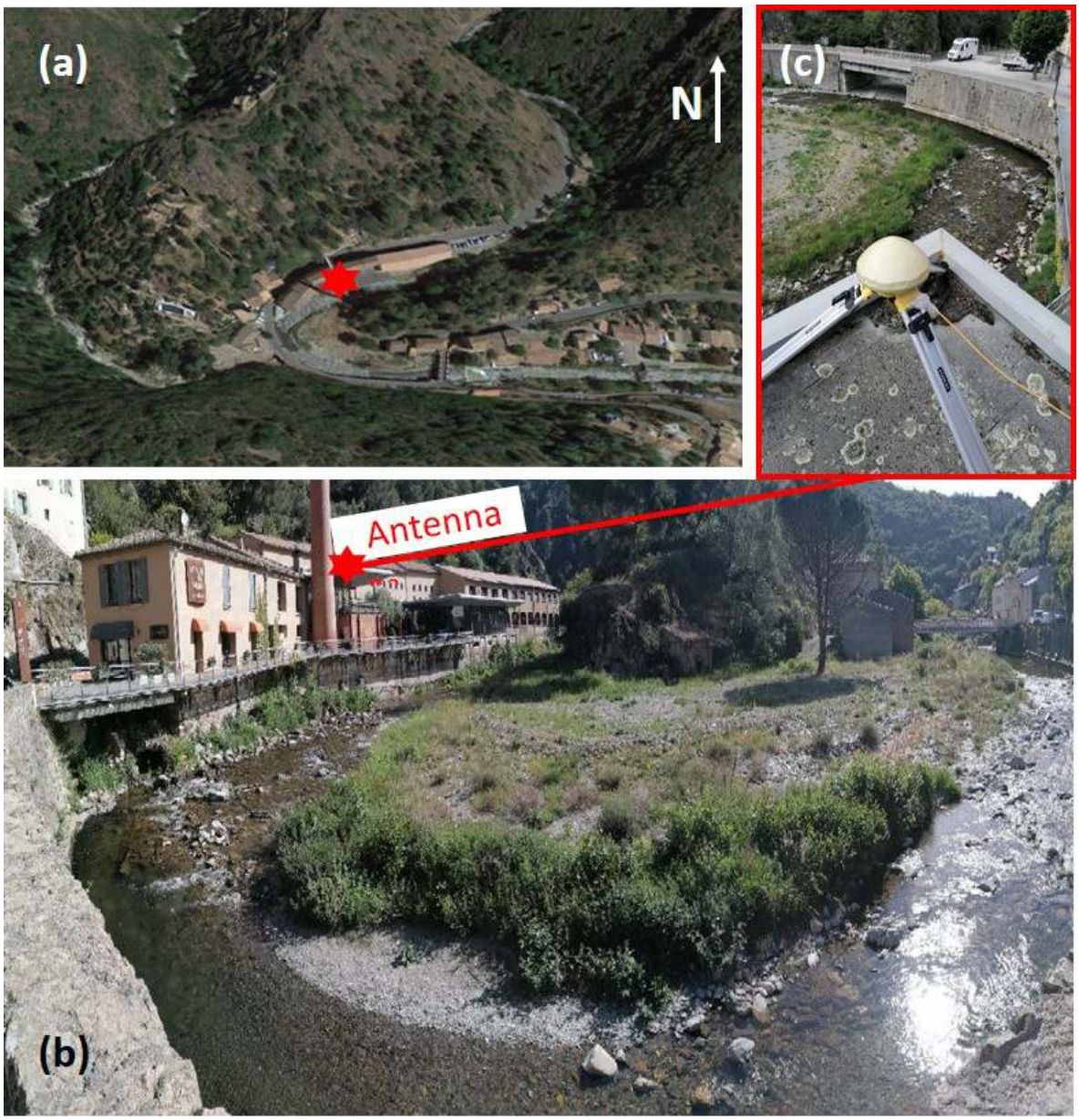

Figure 3.3: Configuration of the GNSS-R acquisition in Salsigne. **(a)** Google Earth view of the acquisition site at the confluence between the Orbiel and Grésillou streams, with the antenna location (red star) highlighted. **(b)** Image of the Orbiel stream with the antenna at the North, and **(c)** view of the antenna setup. At high floods, the water can pass over the road and the Orbiel meander is inundated.

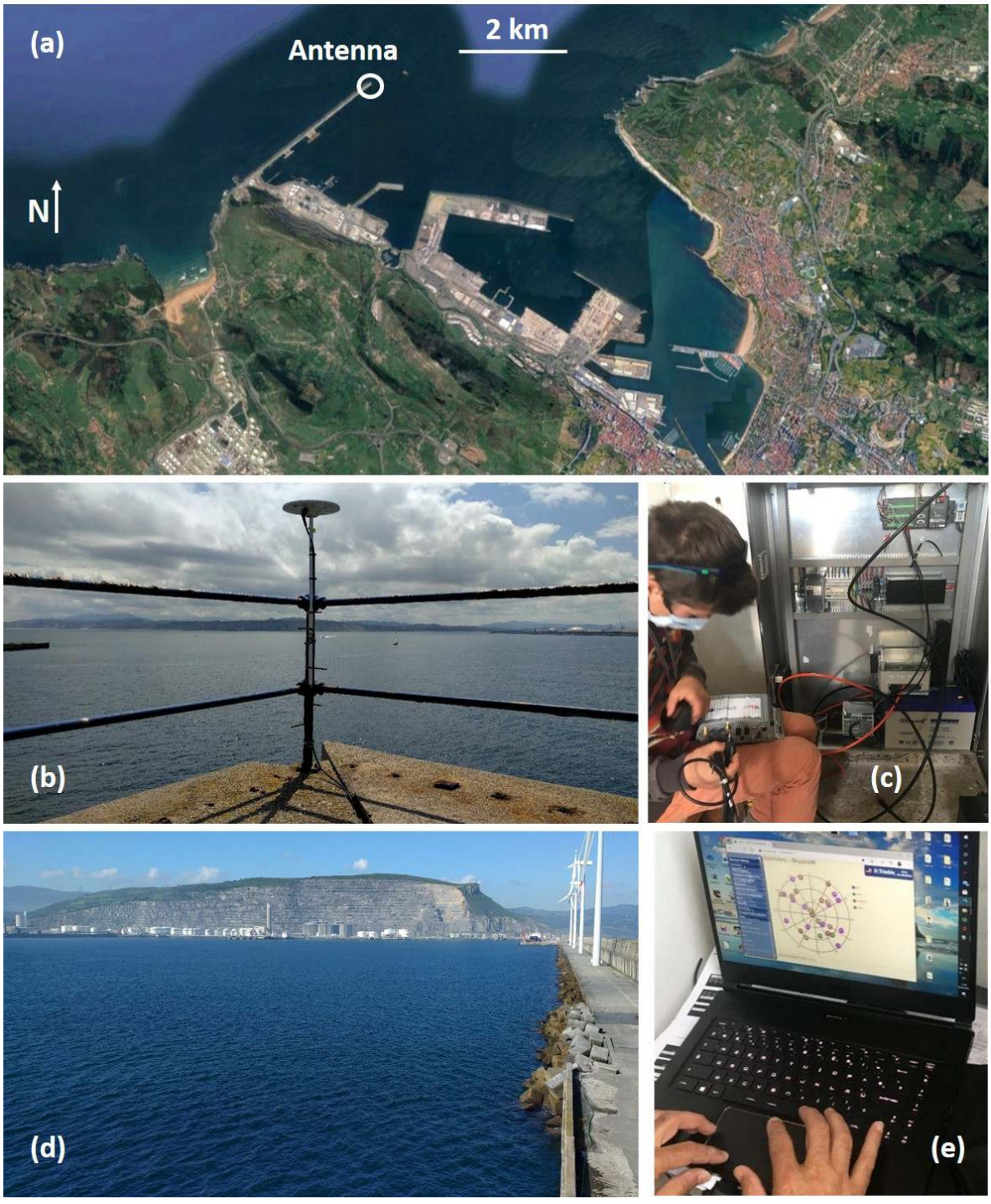

Figure 3.4: Configuration of the GNSS-R acquisition in Bilbao. **(a)** Google Earth view of the port of Bilbao with the receiver installed at the end of the pier, **(b)** antenna setup, **(c)** connection of the receiver to the antenna, **(d)** view of the large area protected by the pier, and **(e)** GNSS satellites tracked after the setup was completed.

Chapter 4

Analysis of spaceborne GNSS-R observations for flood detection

Contents

4.1 Introduction

This chapter presents the results obtained for the evaluation of the potential of CYGNSS observations for the determination of the global flood dynamics. It was the core of the PhD project on studying land geophysical parameters with GNSS-R. Multiple initial objectives were thought before starting this work. In particular, we thought about studying the dependence of the GNSS reflected signals on the SM, the presence of surface water and the vegetation cover. The idea was further to disentangle the effects of these variables to obtain joint estimations of SM and L-VOD. Another objective was to study the transition between soil moisture and surface water, for improving the SMAP and SMOS-based SM estimates over the tropics.

The development of a processing chain to compute CYGNSS reflectivity from the Level 1 DDM observations was time-consuming. When plotting my first reflectivity maps in the full coverage of CYGNSS, people working on hydrology were very excited about its potential, showing a strong signal over most of the floodplains. The surface water dynamics then became the principal object of study, for two main reasons. First, before a flurry of article were recently published on surface water detection using CYGNSS (Al-Khaldi et al., 2021b,a; Li et al., 2021, 2022; Collett et al., 2022; Chapman et al., 2022), this field was not much investigated. The initial contributions used very simple methods, mostly based on empirical thresholds, to map either spatial or temporal changes in the flooded areas at regional scale (Chew et al., 2018; Wan et al., 2019; Morris et al., 2019; Gerlein-Safdi and Ruf, 2019; Rajabi et al., 2020; Ghasemigoudarzi et al., 2020). Then, the state-of-the-art highlighted a lack of knowledge in tropical wetlands hydrology. In particular, the monitoring of inundation dynamics in tropical forests is very poor, and relies only on sparse SAR studies and global, 0.25° GIEMS estimates (see Chapter 1). This is specifically the contribution that should be expected from CYGNSS. It offers a higher temporal resolution than most of the SAR and a higher spatial resolution than radiometers, combined to an optimal penetration of the vegetation layers due to the large wavelength of GNSS L-band signals.

The state-of-the-art and the early results shifted this study toward the characterization of inundation dynamics with an interesting temporal resolution (weekly time sampling), and a quite low spatial resolution of 0.1°. The study was conducted on 1-year CYGNSS data, as the algorithms are very long to run for both the pre-processing and the processing. The methods and results are presented below, in a publication that has just been accepted in *Remote Sensing of Environment*. Section 4.3 then gives the conclusions and perspectives to this work.

4.2 Publication in *Remote Sensing of Environment*

Remote Sensing of Environment 282 (2022) 113278

Contents lists available at ScienceDirect

Remote Sensing of Environment

journal homepage: www.elsevier.com/locate/rse

Analysis of CYGNSS coherent reflectivity over land for the characterization of pan-tropical inundation dynamics

Pierre Zeiger [a,*], Frédéric Frappart [b], José Darrozes [c], Catherine Prigent [d,e], Carlos Jiménez [d,e]

[a] Laboratoire d'Etudes en Géophysique et Océanographie Spatiale (LEGOS), Université Paul Sabatier (UPS), Toulouse, France
[b] Interaction Sol Plante Atmosphère (ISPA), INRAE Bordeaux, Villenave-d'Ornon, France
[c] Géosciences Environnement Toulouse (GET), Université Paul Sabatier (UPS), Toulouse, France
[d] CNRS, LERMA, Observatoire de Paris, Sorbonne Université, Université PSL, Paris, France
[e] Estellus, Paris, France

ARTICLE INFO

Edited by Menghua Wang

Keywords:
GNSS-Reflectometry
CYGNSS
Coherent reflectivity
Flood dynamics
Surface water extent

ABSTRACT

The monitoring of flood and wetland dynamics at global scale is hampered by several limitations, including a reduced data availability in tropical areas due to the presence of clouds affecting visible and infrared imagery, or low spatial and/or temporal resolutions affecting passive and active microwave Earth Observation (EO) data. As a consequence, surface water extent estimates and their temporal variations remain challenging especially in equatorial river basins. Global Navigation Satellite System Reflectometry (GNSS-R) L-band signals recorded onboard Cyclone GNSS (CYGNSS) mission, composed of 8 Low Elevation Orbit (LEO) satellites, provide information on surface properties at high temporal resolution from 2017 up to now. CYGNSS bistatic observations were analyzed for detecting permanent water and seasonal floodplains over the full coverage of the mission, from 40°S to 40°N. We computed CYGNSS reflectivity associated to the coherent component of the received power, that was gridded at 0.1° spatial resolution with a 7-day time sampling afterwards. Several statistical metrics were derived from CYGNSS reflectivity, including the weighted mean and standard deviation, the median and the 90^{th} percentile (respectively Γ_{mean}, Γ_{std}, Γ_{median} and $\Gamma_{90\%}$) in each pixel. These parameters were clustered using the K-means algorithm with an implementation of the Dynamic Time Warping (DTW) similarity measure. They were compared to static inundation maps, and to dynamic estimations of surface water extent both at the global and regional scales, using the Global Inundation Extent from Multi-Satellites (GIEMS) and MODIS-based products. The difference between $\Gamma_{90\%}$ and Γ_{median} shows the best sensitivity to the presence of water. The river streams and lakes are correctly detected, and a strong seasonality is identified in CYGNSS reflectivity over the largest floodplains, with the exception of the Cuvette Centrale of Congo which is covered by dense vegetation. This seasonal reflectivity signal correlates well with inundation maps: Pearson's correlation coefficient between Γ_{median} and surface water extent from both GIEMS and MODIS is over 0.8 in the largest floodplains. The spatial patterns of reflectivity are consistent with static inundation maps: at the time of maximum flooding extent, a spatial correlation coefficient around 0.75 with Γ_{median} is obtained for several basins. We also evaluated the dependence of CYGNSS-derived clusters and reflectivity on the dominant land cover type and on the density of Above Groud Biomass (AGB) in the pixel. On the one hand, misclassifications of flooded pixels were observed over vegetated regions, probably due to uncertainties related to the attenuation by the vegetation in both CYGNSS and reference datasets. On the other hand, flooded pixels with a mean AGB up to ~300 Mg/ha were correctly detected with the clustering. High reflectivity values are also observed over rocky soils in arid regions and create false alarms. Finally, strong winds on large lakes cause surface roughness, and lower reflectivity values are observed in this case which weaken the detection of open water. While these constraints are to be taken in account and corrected in a future model, a pan-tropical mapping of surface water extent dynamics using CYGNSS can be envisaged.

* Corresponding author.
E-mail address: pierre.zeiger@legos.obs-mip.fr (P. Zeiger).

https://doi.org/10.1016/j.rse.2022.113278
Received 28 March 2022; Received in revised form 12 September 2022; Accepted 16 September 2022

1. Introduction

Wetland ecosystems and floodplains play a key role for the management of water and natural resources and for climate change, although they cover only 8% of the land surfaces (Davidson et al., 2018). They temper the water cycles through the regulation of river discharges and the mitigation of floods (Bullock and Acreman, 2003; Acreman and Holden, 2013). They are particularly sensitive to climate change, while contributing a lot to the world's greenhouse gases emissions through the global biogeochemical cycles. According to estimations in the literature, wetlands are the source of 20–25% of the world's methane emissions (Bartlett and Harriss, 1993; Whalen, 2005; Bergamaschi et al., 2007; Bloom et al., 2010; Ringeval et al., 2010; Melton et al., 2013; Nisbet et al., 2014; Saunois et al., 2020), and they store 16 to 33% of the soil carbon pool (Maltby and Immirzi, 1993; Page et al., 2011; Mitsch et al., 2013). Beyond the natural biogeochemical cycles, the anthropogenic activity strongly impacts wetlands, and exerts feedback loops along with climate change through the variations of the carbon stocks and methane emissions. Moreover, wetlands are known to be major reserves of biodiversity (Mitra et al., 2005; Junk et al., 2006; Webb et al., 2010). The temporal and spatial variations in inundation extent also affect the propagation of infectious diseases (Kouadio et al., 2012; Suk et al., 2020). All these effects affect the millions of people living worldwide in wetlands and relying on a healthy wetland ecosystem (Maltby and Acreman, 2011). Besides, unprecented floods and droughts increase the vulnerability in many regions of the world in spite of the improvement of risk management policies (Kreibich et al., 2022). For these aforementioned factors, a better monitoring from regional to global scales of the flood extent and dynamics is needed.

Inundation mapping usually depends on remote sensing due to the lack of *in situ* data and the difficulties to perform measurements in remote regions. Long time-series of inundation extent at 30 m spatial resolution were obtained with the processing of Landsat multispectral images (Pekel et al., 2016), but at the expense of a low temporal resolution. The Moderate Resolution Imaging Spectroradiometer (MODIS) instrument was also used to produce a delineation of flooded area with a 8-day temporal resolution, but with a lower spatial resolution from 250 m to 1 km (Chen et al., 2013; Di Vittorio and Georgakakos, 2018). Nevertheless, while these products provide reliable information on open water bodies, they suffer from limitations due to the presence of clouds in equatorial areas, and are unable to detect water under dense canopies such as the inundated forests present in the Amazon and the Congo basins.

Microwave remote sensing is less affected by these limitations in both passive and active domains. The presence of freshwater is responsible for: i) a decrease in the brightness temperature, in passive mode, affecting more sensibly the horizontal than the vertical polarization (Choudhury, 1991; Sippel et al., 1994; Prigent et al., 2001), ii) in active mode, low backscattering over open water and high backscattering in presence of water under vegetation owing to the double-bounce effect for side-looking radar such as Synthetic Aperture Radar (SAR) (Richards et al., 1987), and iii) also in active mode, high backscattering over open water and wetlands for nadir-looking radar altimeters (Frappart et al., 2021), except over large lakes where surface roughness (*e.g.* waves on windy days) can cause a strong scattering. L-band in particular is able to penetrate deeper the vegetation cover than the higher frequency microwave bands. It allows the detection of water under the vegetation even in equatorial basins, in both passive (Parrens et al., 2017) and active (Hess et al., 2003; Betbeder et al., 2014) domains. Nevertheless, major drawbacks limit the use of microwave observations for flood monitoring. First, passive microwave observations have a coarse spatial resolution, generally lower than 25 km, which limits their use for flood monitoring in spite of their quasi-daily temporal repeat. For example, the Soil Moisture and Ocean Salinity (SMOS, Kerr et al., 2001) and the Soil Moisture Active Passive (SMAP, Entekhabi et al., 2010) missions measure the brightness temperatures at L-band, with a nominal spatial resolution of several tens of kilometers and a revisit period of 2–3 days. Then, the use of Synthetic Aperture Radar (SAR) active microwave images was also limited by a low temporal sampling before the launch of Sentinel-1, the difficulty to identify the signature of water in complex environments, and the limited availability of L-band images at global scale. Finally, the use of radar altimetry for flood mapping is limited by its acquisition mode along the satellite tracks that does not offer a global coverage of land surfaces and by its low temporal resolutions ranging from 10 to 35 days. As a consequence, most of the studies were limited to regional mapping of the flood dynamics (Hamilton et al., 2004; Kuenzer et al., 2013; Parrens et al., 2017). The combination of multi-satellite information allows to overcome some of these limitations. The Global Inundation Extent Multi-Satellite (GIEMS) product combines information from passive microwave, radar scatterometers, and visible/ near-infrared images to account for vegetation effect in the flood detection. It has been providing a continuous monitoring of the wetland dynamics globally, at monthly time scale and ~0.25° spatial resolution since 1992 (Prigent et al., 2007; Prigent et al., 2020).

Global Navigation Satellite System (GNSS) Reflectometry (GNSS-R) onboard satellite platforms can be a great help to improve the spatial and temporal resolutions of wetlands dynamics mapping. It uses the L-band (f = 1.575 GHz for L1) GNSS signals scattered by the Earth's surface, and collected by a receiver as a multistatic measurement technique (Martin-Neira, 1993; Zavorotny et al., 2014). In-situ and airborne GNSS-R measurements have shown sensitivity to various oceanic and land geophysical parameters (Ruffini et al., 2004; Cardellach et al., 2011; Egido et al., 2014). Over the last years, satellite missions carrying GNSS-R receivers have been launched and emerge as new tools for global applications.

The UK TechDemoSat-1 (TDS-1) mission was a proof of concept of spaceborne GNSS-R measurements. It successfully allowed it to retrieve either oceanic parameters such as sea level and wind speed (Foti et al., 2015; Clarizia et al., 2016), or land geophysical parameters such as soil moisture and vegetation (Camps et al., 2016; Chew et al., 2016). It was followed by the launch of the NASA Cyclone GNSS (CYGNSS) mission, whose 8 Low Earth Orbit (LEO) micro-satellites are designed to monitor the formation of tropical cyclones through wind speed retrieval (Ruf et al., 2016). CYGNSS observations are also collected over land over the pan-tropical area (±38° latitude). The design of the mission suits the study of land parameters with a fine spatiotemporal resolution, as each of the 8 satellites records simultaneously 4 observations at a sampling rate of 1 Hz (2 Hz after July 2019). CYGNSS observations are sensitive to the properties of the reflecting surface, including surface roughness, soil moisture (SM), vegetation water content (VWC) and biomass density (Carreno-Luengo et al., 2019). One of the major applications over land of this dataset is soil moisture retrieval with a good accuracy (Root Mean Square Error ranging from 0.04 to 0.07 cm^3/cm^3) (Chew and Small, 2018; Kim and Lakshmi, 2018; Al-Khaldi et al., 2019; Clarizia et al., 2019; Eroglu et al., 2019; Senyurek et al., 2020). It has also been demonstrated that CYGNSS can be used to increase both the spatial and temporal sampling of existing SM products derived from passive microwave sensors (*e.g.*, SMAP, SMOS), that suffer from a lower spatio-temporal resolution (Yan et al., 2020).

CYGNSS reflectivity, based on the Delay Doppler Map (DDM) Peak power and metadata used for the calibration of the signal, was also used for identifying flood signatures in the spaceborne GNSS-R observations. In particular, the reflectivity maps show the changes in inundation extent following hurricanes, typhoons or extreme rain events (Chew et al., 2018; Morris et al., 2019; Wan et al., 2019; Ghasemigoudarzi et al., 2020; Rajabi et al., 2020). The methodologies mostly consist in an empirical thresholding of the CYGNSS reflectivity or Signal-to-Noise Ratio (SNR), except Ghasemigoudarzi et al. (2020) which used the Random Under-Sampling Boosted (RUSBoost) classification algorithm to identify flooded and non-flooded CYGNSS data. Gerlein-Safdi and Ruf (2019) used CYGNSS to produce annual water masks in the Congo basin,

with a 0.1° spatial resolution, using a spatial comparison of one pixel's value with its neighbors and the random walker segmentation method. Then, Gerlein-Safdi et al. (2021) used the same approach temporally, for comparing the monthly mean of a pixel with the average value of its time series. They produced monthly water masks at 0.01° spatial resolution over the Pantanal and Sudd wetlands. Finally, Jensen et al. (2018) studied the relationship between CYGNSS reflectivity, floods and vegetation in tropical wetlands in the hydrographic basin of the Amazon. Based on this analysis, Rodriguez-Alvarez et al. (2019) classified the open water and flooded vegetation over the same study area. Note that all these studies were performed at the regional scale. At larger scale, CYGNSS reflectivity and coherence over flooded areas vary temporally and spatially due to vegetation cover, due to the changes in roughness affecting soil and water surfaces (Chew and Small, 2020), and to the bistatic geometry of the acquisition Loria et al. (2020).

Recent improvements have been made in the detection of surface water using several indicators of the coherence of CYGNSS observations. In particular, the raw Intermediate Frequency (IF) data occasionally collected were analyzed in several studies (Li et al., 2021; Li et al., 2022a; Collett et al., 2022). They are unprocessed signal samples which contain phase information at a high sampling rate, before the incoherent averaging at 1–2 Hz used to derive Level-1 DDM products. Li et al. (2021) calculated the complex DDM and derived a coherent coefficient (CoC) for BeiDou-3 raw IF. They found that a high CoC is highly linked with the presence of water bodies. Li et al. (2022a) proposed a multi-GNSS complex waveform product with a high sampling rate (1000 Hz), that enables the detection of surface water at ∼100 m spatial resolution along the reflection track. And Collett et al. (2022) showed that the coherence of CYGNSS raw IF samples highly relies on the standing surface water, and if no water body is found in the footprint, coherence is found for saturated soils ($SM \geqslant 0.45 cm^3/cm^3$) with low large-scale roughness. An entropy-based metric was also proposed by Russo et al. (2022) to characterize the coherence of CYGNSS raw IF. It was very sensitive to open water and floods in forested areas along Yucatan Lake and the Mississippi River (Chapman et al., 2022). An operational and continuous generation of raw IF data at different frequencies and polarizations is planned for the future European Space Agency (ESA) GNSS-R mission, HydroGNSS, of which the detection of surface water is one of the main science objectives (Unwin et al., 2021). But currently, the availability of raw IF data is low compared to CYGNSS Level-1 data, and they are more an object of investigation rather than an operational solution for the monitoring of flood dynamics. For this reason, Al-Khaldi et al. (2021b) gave an estimation of the coherence based on the shape of CYGNSS Level-1 DDM. Some pan-tropical water masks were further proposed in Al-Khaldi et al. (2021a) using the aggregation of CYGNSS observations with varying spatio-temporal resolution: 1 km/1 year, 3 km/3 months, and 6 km/2 weeks. At 1-km spatial resolution, the mask is annual at best but the permanent water bodies are well delineated, including small tributaries of the Amazon and Congo rivers. On the contrary, lower spatial resolution allows a much higher temporal sampling and the detection of flood events.

In this study, we focus on CYGNSS reflectivity to map flood dynamics with a 0.1° spatial resolution and a 7-day temporal resolution. The CYGNSS reflectivity time series are classified using a K-means clustering technique with Dynamic Time Warping (DTW) similarity measurement (Müller, 2007), to delineate the flooded areas. Results are compared to reference static and dynamic maps of inundations at global and regional scales, among other ancillary datasets. Section 2 presents the datasets, Section 3 the methodology of this study, while Sections 4 and 5 show the results and a discussion about the main conclusions of this paper.

2. CYGNSS and reference datasets

2.1. CYGNSS

The CYGNSS mission is composed of 8 LEO micro-satellites covering the pan-tropical area (±38° latitude). Each satellite carries onboard a Delay Doppler Mapping Instrument (DDMI), which is composed of a GPS receiver, nadir-looking antennas for collecting GNSS signals, and a zenith-looking antenna for geolocation purpose. The DDMI records simultaneously 4 reflected signals integrated over a second (0.5 s after July, 2019), so the CYGNSS mission provides 32 (64, respectively) observations per second in the area of coverage.The temporal sampling is high, with a median and mean revisit time over a 25 km pixel of 3 h and 7 h, respectively (Ruf et al., 2016). With higher spatial resolution, the revisit time increases. Still, substantial improvements in terms of temporal resolution are expected using CYGNSS when compared to monostatic radars.

Due to the bistatic configuration of the measurements between an emitting Global Positioning System (GPS) satellite and a receiving CYGNSS observatory, the sampling of CYGNSS observations over the Earth's surface is pseudo-random. The L-band signals ($f = 1.575$ GHz for GPS L1) are emitted by GNSS satellites and received by each of the 8 CYGNSS micro-satellites, with an incidence angle ranging from 0° to ∼70° and following a Gaussian distribution centered on ∼30°. Over the ocean, the reflected GPS signals are dominated by the incoherent component due to an important surface roughness. The spatial resolution is there degraded and wind speed can be estimated with a 25 km spatial resolution and a sub-daily mean revisit time (Clarizia and Ruf, 2016). However, the reflected signals over land can be dominated by the coherent component, in case of the presence of water bodies or areas characterized by a high surface moisture content in the glistening zone, with low surface roughness at the scale of the GPS L1 wavelength (19.03 cm). The spatial resolution depends, in that case, on the size of the first Fresnel zone which is larger than ∼0.6 km × 0.6 km (the minimum area in the case of a 0° incidence angle) (Eroglu et al., 2019; Rodriguez-Alvarez et al., 2019). The finest theoretical spatial resolution for a CYGNSS observation over flooded regions is therefore ∼0.6 km × 6.6 km, with an elongation effect due to the integration of the reflected signals over 1 s (0.5 s, respectively) for observations before (after) July, 2019. Yet, CYGNSS observations are sensitive to the presence of a small fraction of water in the glistening zone, whose contribution dominates the total response of the reflecting surface. The maps of reflectivity over river basins, such as the Amazon, show a strong scattering over small tributaries, down to a few hundred meters.

The CYGNSS observables used in this study come from CYGNSS Level 1 science data record files (CYGNSS, 2020). They contain the 17x11 Delay Doppler Maps (DDMs) of raw bins and analog scattered power, along with all the parameters used for data processing, *i.e.* the geometry of the acquisition and data flags. We used the version 3.0 of L1 data over one year, from August 1st, 2018 to July 31st, 2019. The observations before July, 2019 are sampled at 1 Hz (2 Hz after), so we assume a ∼0.6 km × 6.6 km spatial resolution although the along-track resolution is twice higher for the last month of data. The daily files (one for each CYGNSS satellite) are distributed in the Physical Oceanography Distributed Active Archive Center (PODAAC: https://podaac-opendap.jpl.nasa.gov/opendap/hyrax/allData/cygnss/L1/v3.0/). The CYGNSS DDM of scattered analog power are used in this study, along with quality flags and metadata from the L1 files. The description of the preprocessing and analysis of CYGNSS dataset can be found in Section 3 (methodology).

2.2. Other sources of data

Other data sources were used to analyze and validate the results presented below. They consist in static and dynamic inundation maps available at regional and global scales, land cover thematic map and Above Ground Biomass (AGB).

2.2.1. Static inundation maps

Two different datasets were considered for a delineation of flooded areas: the Level 3 Global Lakes and Wetlands Database (GLWD-3) at 30" resolution (Lehner and Döll, 2004), and the wetlands maps at 15" (500 m at the Equator) from Tootchi et al. (2019). Note that these two datasets do not provide any temporal variations and are thus static. After some analyzes, the latter was selected because it provides maps of both the Regularly Flooded Wetlands (RFWs, overlapping open water with several inundation datasets) and the Composite Water (CWs) which associates RFWs with Groundwater modelling (GDWs). It is of interest to compare RFWs to CWs maps, because CYGNSS may be sensitive to the groundwater-driven wetlands that represent a considerable source of floods. The CW-WTD product is used as it provides both RFWs areas, and a groundwater wetlands map derived from a direct water table depth (WTD) modeling (Fan et al., 2013).

The RFWs in Tootchi et al. (2019) were derived merging 3 inundation datasets, namely the open water and flooded areas extracted from the ESA-CCI Land Cover (see Section 2.2.3), the mean annual maximum water extent from GIEMS-D15 (Fluet-Chouinard et al., 2015), and the maximum surface water from the Joint Research Center (JRC) product (Pekel et al., 2016). It assumes that every single product has observation gaps and should be complemented by the use of other ones. It is particularly the case between the high-resolution JRC maps based on Landsat, and GIEMS-D15 which downscales the coarse 0.25° estimations from GIEMS at 15" resolution. The JRC product is unable to detect water under dense vegetation cover, while GIEMS-D15 is able to detect the floods in tropical or boreal forests. However, GIEMS-D15 is dominant when compared to other products. It represents 55% of the total RFWs areas, and only 10% of its extent is confirmed by either CCI land cover or JRC surface water. As a consequence, the uncertainties of GIEMS are directly propagated in the RFWs dataset. In particular, a recurrent overestimation of Surface Water Extent (SWE) in GIEMS has been found in the Ganges irrigated paddy fields, in the Sahel wetlands, and in coastal areas, due to a confusion between the signals from surface water and wet soils. Also, as the original spatial resolution of GIEMS is 0.25°, the sensitivity to small river streams in several regions, including the Amazon and Congo basins, is found to be low in both GIEMS-D15 and Tootchi's RFWs dataset.

2.2.2. Dynamic inundation maps

Few global products exist that provide temporal variations of inundation extent at the global scale. To compare with CYGNSS data, we used the Global Inundation Extent from Multi-Satellite version 2 (GIEMS-2) dataset (Prigent et al., 2020). It contains a global estimation of monthly flooded area in a 0.25° grid since 1992. In this study, we have used the new global maps for years 2018 and 2019, matching the times of our CYGNSS dataset.

A more precise comparison is performed regionally between CYGNSS observables and multispectral-based flood maps derived from the gridded atmospherically corrected surface reflectances from MODIS, acquired in seven spectral bands from visible/ near infrared (NIR) to shortwave infrared (SWIR). The surface reflectances contained in the MOD09A1 product (8-day binned level 3, version 6) have a spatial resolution of 500 m and a temporal resolution of 8 days. Flood extent maps were generated using the threshold method proposed in Frappart et al. (2018) and Normandin et al. (2018). It is a simplified version of the multi-threshold approach developed by Sakamoto et al. (2007). In this approach, a MODIS pixel is considered fully or partially covered with water if: i) the value of its Enhanced Vegetation Index (EVI) (Huete et al., 1997) is lower or equal to 0.05 and the value of its land surface water index (LSWI) (Xiao et al., 2005) is negative or equal to zero, or ii) its EVI value is lower than 0.3 and the difference between EVI and LSWI is lower than 0.05. This method was used to produce inundation maps in the Lower Mekong Basin (LMB) in Asia, the Inner Niger Delta (IND) in Africa, and La Plata basin in South America during the study period.

2.2.3. CCI Land Cover

The characteristics of CYGNSS reflectivity highly rely on the type of Land Cover (LC) present in the glistening zone (Carreno-Luengo et al., 2019; Chew and Small, 2020). To analyze the capability of CYGNSS to provide reliable information on the presence of water over land depending on the LC type, we used global LC maps from the European Space Agency's (ESA) Climate Change Initiative (CCI) at 300 m resolution (ESA, 2017, available at https://www.esa-landcover-cci.org/). These maps are obtained from 1992 to 2015 using various imagery data, including the Medium Resolution Imaging Spectrometer (MERIS) and Project for On-Board Autonomy – Vegetation (PROBA-V), and applying the GlobCover unsupervised classifier (Defourny et al., 2007). More recent maps have been produced and validated by the Copernicus Climate Change Service (C3S), so we used the 2019 map which corresponds the best to our CYGNSS time series. Additionally, we aggregated the initial 38 LC types defined following the United Nations Land Cover Classification System (UN-LCCS) (Di Gregorio, 2016) into 10 general land cover classes, which is summarized in Table 1. Croplands, forests and herbaceous are separated into dry and flooded regions to study the changes in CYGNSS reflectivity associated to the presence of water.

2.2.4. Above Ground Biomass

Dense vegetation layers can attenuate and even extinct the coherent component of GNSS-R signals (Loria et al., 2020), and the corresponding decrease of CYGNSS reflectivity was also modeled with respect to the Above Ground Biomass (AGB) of the reflecting surface (Carreno-Luengo et al., 2020). We therefore analysed the distribution of areas detected as flooded and non-flooded using CYGNSS reflectivity, depending on the mean Above Ground Biomass (AGB) in the pixel. We used the global GlobBiomass AGB maps (Santoro, 2018) which are given at 1 ha (approximately 3.2 arcsec) spatial resolution for the reference year 2010. It was derived from a set of observations including the L-band backscatter from ALOS-PALSAR and multiple C-band observations from Advanced Synthetic Aperture Radar (ASAR) onboard Envisat, with lidar-based metrics from Ice Cloud and Land Elevation Satellite (ICESat), and Landsat reflectances at several bands (Santoro et al., 2021). The study independently estimated the Growing Stock Volume (GSV) for the backscatters of ALOS-PALSAR and ASAR, which are combined to obtain a more robust dataset and then invert AGB. The AGB maps reproduced well the known spatial patterns with high level of detail. Over the tropics, the comparison with a database of field measurements highlighted

Table 1
List of the 10 aggregated land cover classes, and the associated class numbers in the CCI LC classification system.

Class numbering	Class name	CCI Land Cover classes associated
LC1	Open water	210
LC2	Cropland non irrigated	10, 11, 12, 30
LC3	Cropland irrigated	20
LC4	Non flooded forest	50, 60, 61, 62, 70, 71, 72, 80, 81, 82, 90
LC5	Flooded forest	160, 170
LC6	Dry shrubs, herbaceous	40, 100, 110, 120, 121, 122, 130
LC7	Flooded shrubs, herbaceous	180
LC8	Bare soils	140, 150, 151, 152, 153, 200, 201, 202
LC9	Snow, ice	220
LC10	Urban	190

a systematic underestimation of AGB in the densest forests. This is due to the low dynamic range of backscatter observations over regions with AGB higher than 250 Mg/ha, and to the unavailability of ASAR GSV estimates which causes the AGB inversion to only rely on ALOS-PALSAR backscatter.

2.2.5. Resampling of the ancillary datasets

All the ancillary datasets have higher spatial resolutions than our 0.1° CYGNSS grid, ranging from ~5100 m to 1 km at the equator. To compare with CYGNSS observables, these data were upscaled into a matching 0.1° grid. For CCI LC, the percentage of each LC class in every pixel was computed. The mean AGB was extracted from GlobBiomass 1 ha spatial resolution dataset. For static and dynamic inundation maps, we calculated the open water and flood extent as a percentage of the pixel surface. A notable exception is GIEMS-2, whose spatial resolution is 0.25°. To compare at the global scale GIEMS vs CYGNSS, we also computed CYGNSS-derived parameters at 0.25° spatial resolution.

2.2.6. Water levels in 2018–2019

We have used water levels derived from radar altimetry as a proxy to estimate the severity of floods in the different river basins further analyzed, during the time span of our study (August 2018 - July 2019). We downloaded the time series of water levels from the Hydroweb database (Crétaux et al., 2011; Hydroweb, 2022) at several specific locations which are described in the caption of Table 2. We extracted the yearly maximal water heights since 2016 (launch of Sentinel-3A). We then computed the anomaly of 2018–2019 peak water height compared to the average annual maximum, which are listed in Table 2. The relative anomaly represents the anomaly normalized by the average amplitude of seasonal variations, and is also shown. We notice high water levels in the Mekong River (2.96 m over the average, with a relative anomaly of 36.5%), and in South America: 0.76–0.77 m for the Orinoco and Madeira rivers, with a relative anomaly of 11.1–11.5%. This could be due to a weak El Niño in 2018–2019 that may also have affected the Yangtze basin (not shown here). The Parana has a relative anomaly of 42.2% mainly due to a strong decrease of water heights and inundation extent after the severe droughts of the summer 2019, posterior to the time span of our study. For the other basins, the Congo, Ganges and Brahmaputra rivers show low negative anomalies of 0.19 m to 0.39 m (-2.3% to −14.3%), while the Niger shows a positive anomaly of 0.54 m (+7.2%), likely linked to high floods in the Inner Niger Delta. We can also mention the very long-lasting 2019 spring floods along the Mississippi River, that have already been analyzed in several CYGNSS-related studies (Li et al., 2021; Chapman et al., 2022).

3. Methods

3.1. Preprocessing of the CYGNSS dataset

An overview of the processing chain of CYGNSS data developed in this study is presented in Fig. 1. First, the peak of each CYGNSS DDM in the daily L1 files (see Section 2.1) is extracted and the entire 17x11 L1 DDM is removed, as a further processing would be very time-consuming. The CYGNSS overland flag is used to remove all reflections over the oceans, and some other quality flags are applied: *S-band powered up*, *Large spacecraft attitude error*, *Black-body DDM*, *DDM is test pattern*, *Low confidence GPS EIRP estimate*. This set of flags was used in previous studies to ensure a correct filtering with the removal of the least samples possible (Chew and Small, 2018; Eroglu et al., 2019). We also tried to apply more constraining masks using all the flags combined in the CYGNSS overall quality flag with the logical exception of the overland flag (Clarizia et al., 2019; Rodriguez-Alvarez et al., 2019). However, this latter approach removed a major part of CYGNSS observations over land (Eroglu et al., 2019), and decreased the performance of land geophysical parameter retrieval with a high spatiotemporal resolution.

The CYGNSS L1 algorithms performed an estimation of the specular point location based on a mean sea surface model, which is close to the geoid (Gleason et al., 2019). This can affect the DDMs over land as topography is not taken into account. Before the end of 2017, data collected over 600 m were of poor quality, which lead several authors to apply an elevation cutoff (Eroglu et al., 2019; Rodriguez-Alvarez et al., 2019; Yan et al., 2020). Further versions of the L1 calibration procedures for land reflections include the topography to estimate more accurately the specular point location (Gleason et al., 2020). For the time span of our study, CYGNSS DDM are of good quality over a large variety of terrains and elevation ranges, including almost all the wetlands and water bodies in CYGNSS coverage. Notable exceptions are the Tibetan plateau in the Himalayas, and the Altiplano in South America (Gleason et al., 2020). Still, the position of the peak power in CYGNSS DDMs is sometimes shifted due to several factors, including the topography around the specular point. In this study, we filtered the observations when the peak of the 17x11 DDM is located in the three first and last delay rows as previously reported (Yan et al., 2020). This corresponds to a quality control affecting a small part (~4%) of the full dataset.

3.2. CYGNSS reflectivity

The computation of the reflectivity is mostly preferred to the use of the Signal-to-Noise Ratio (SNR) when analyzing CYGNSS data, as it combines information on the DDM peak power along with other CYGNSS metadata useful for the calibration of the signal. The GNSS signals scattered by open water, floodplains and in general smooth land and ice surfaces are considered to be dominated by the coherent component (Carreno-Luengo et al., 2019; Li et al., 2017; Rodriguez-Alvarez et al., 2019). The power ratio (PR) defined in Al-Khaldi et al. (2021a); Al-Khaldi et al., 2021b also highlights high coherent returns with a low diffuse scattering over inland water bodies. In our study, CYGNSS reflectivity is hence determined using the coherent scattering equation (De Roo and Ulaby, 1994; Gleason et al., 2020):

$$\Gamma(\theta) = \left(\frac{4\pi}{\lambda}\right)^2 \frac{P_{DDM}(R_r + R_t)^2}{G_r G_t P_t} \quad (1)$$

where P_{DDM} is the peak of the DDM analog power, R_r and R_t are the receiver and the transmitter ranges (distance from the receiver and the transmitter to the specular point), G_r is the receiver antenna gain, $P_t G_t$ is the GPS Equivalent Isotopically Radiated Power (EIRP), and λ is the GPS L1 signal wavelength (i.e. $\lambda = \lambda_{L1} = 0.1903$ m). P_{DDM} is computed using the 17 delay × 11 Doppler DDMs, while the other variables come from CYGNSS Level 1 metadata. In Eq. (1), the reflectivity depends on the incidence angle (θ) of CYGNSS observations. In Fig. 2, the curves of the mean reflectivity vs. θ are shown without correction (red curve), with a

Table 2
Anomalies of the 2018–2019 peak water height compare to the average yearly maximum, at the locations listed hereafter. The **Orinoco** River, downstream the Llanos de Orinoco; the **Amazon** River close to its estuary, and the **Madeira** River (one of the biggest tributaries of the Amazon), downstream the confluence between the Beni, Madre de Dios and Mamore rivers; the **Parana** River, after its confluence with Rio Salado; the **Congo** River, after its confluence with the Ubangui; the **Niger** River, downstream the Inner Niger Delta (IND); the **Ganges** and **Brahmaputra** rivers just before their confluence; and finally, the **Mekong** River at the outlet of Tonle Sap.

River	Orinoco	Amazon	Madeira	Parana	Congo	Niger	Ganges	Brahm.	Mekong
Peak anomaly (m)	0.77	−0.03	0.76	1.24	−0.39	0.54	−0.19	−0.3	2.96
Rel. anomaly (%)	11.5	−0.4	11.1	42.2	−14.3	7.2	−2.3	−5.6	36.5

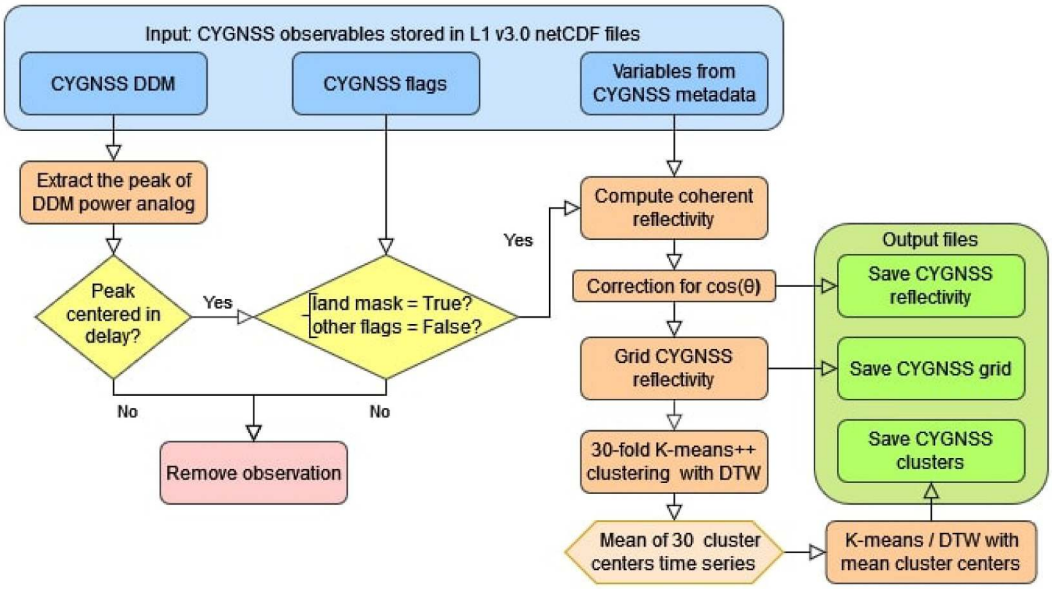

Fig. 1. Overview of the processing chain for the analysis and the clustering of CYGNSS observations.

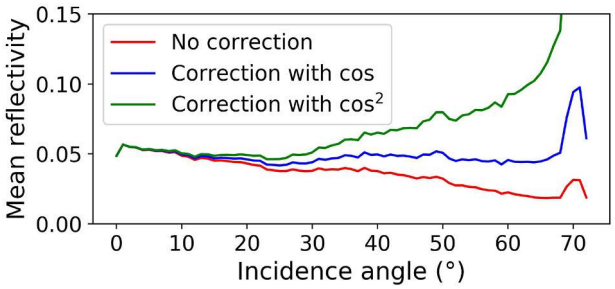

Fig. 2. Mean reflectivity vs. the incidence angle for CYGNSS observations without correction for θ, with a correction based on $cos(\theta)$, and with a correction based on $cos^2(\theta)$.

correction based on $cos(\theta)$ (in blue), and with a correction based on $cos^2(\theta)$ (in green). With no correction, the reflectivity decreases while increasing the incidence angle. With the $cos(\theta)$ correction, the reflectivity is ∼constant in all the incidence range, except for $\theta > 68°$ (these values are finally filtered out). With the latter correction in $cos^2(\theta)$, the reflectivity increases exponentially with an increasing θ. These corrections based on $cos(\theta)$ or $cos^2(\theta)$ are commonly used in correcting the influence of the incidence angle on SAR backscatter (Ulaby et al., 1982; Kellndorfer et al., 1998). As a consequence, the reflectivity corrected from the incidence angle Γ is:

$$\Gamma = \Gamma(\theta = 0°) = \frac{\Gamma(\theta)}{cos(\theta)} \quad (2)$$

Once CYGNSS reflectivity was computed, these unevenly distributed values were gridded at 0.1° spatial resolution (∼11 km at the equator). Several tests were performed to evaluate the best compromise between spatial and temporal resolutions. For the mapping of variables exhibiting a high temporal variability, such as SM, a daily time step is preferable and the best spatial resolution associated is 0.25°. For studying flood dynamics, we prefered to ensure a higher spatial resolution to reduce the confusion between flooded and non-flooded areas. With a 0.1° grid, a 7-day time sampling can be reached. This observation is consistent with Al-Khaldi et al. (2021a), who mapped CYGNSS PR at 2 weeks/6 km, 3 months/3 km or 1 year/1 km. Moreover, our grid is based on a maximum of 32 observations per second (1 Hz sampling of each CYGNSS track). For data after July 2019 the sampling is 2 Hz, so the spatial resolution at weekly time scale could be improved. Finally, an alternative 0.25° grid was also computed to compare CYGNSS reflectivity with the flooded extent from GIEMS-2.

Due to the pseudo-random configuration of CYGNSS bistatic observations, some missing values were present in most of the pixels at specific time steps. A 30-day moving window was applied during the gridding step to fill these gaps. Fig. 3 shows that this process increased the average number of samples per pixel, while the orbits of CYGNSS satellites ensure more observations at the extreme latitudes, and fewer ones in the equatorial regions. A 30-day Gaussian window with a standard deviation $\sigma = 7$ days was used to weight the observations. The weighted mean and standard deviation values of reflectivity per pixel were calculated, which are further expressed as Γ_{mean} and Γ_{std} in linear units. Other statistical parameters were used to describe the distribution of CYGNSS reflectivity values in each pixel: the median (Γ_{median}), the 90^{th} percentile ($\Gamma_{90\%}$), the median absolute deviation (MAD, Γ_{MAD}), and the difference between $\Gamma_{90\%}$ and Γ_{median} ($\Gamma_{90\%-50\%} = \Gamma_{90\%} - \Gamma_{median}$). They were extracted at each time step of the 0.1° grid.

A subset of these variables is shown in Fig. 4 over western and central Africa, including a major part of the Congo, Chad and Niger basins, for the first week of September when the Sahel region is affected by floods. The reflectivity derived from CYGNSS observations is usually high over

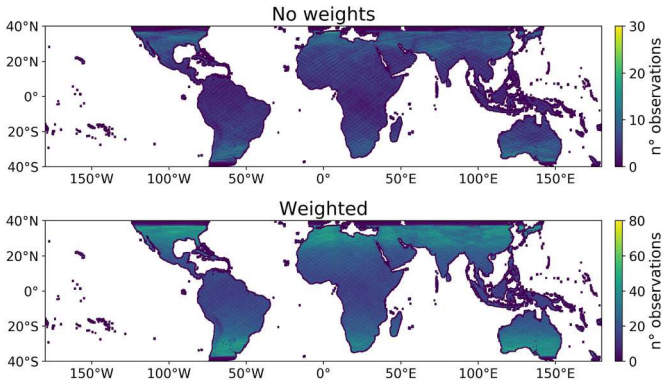

Fig. 3. Number of observations per pixel in 7 days, from 2019–01-02 to 2019–01-08. **(a)** Count without the moving window, **(b)** count with a moving window and 1-month Gaussian weighting.

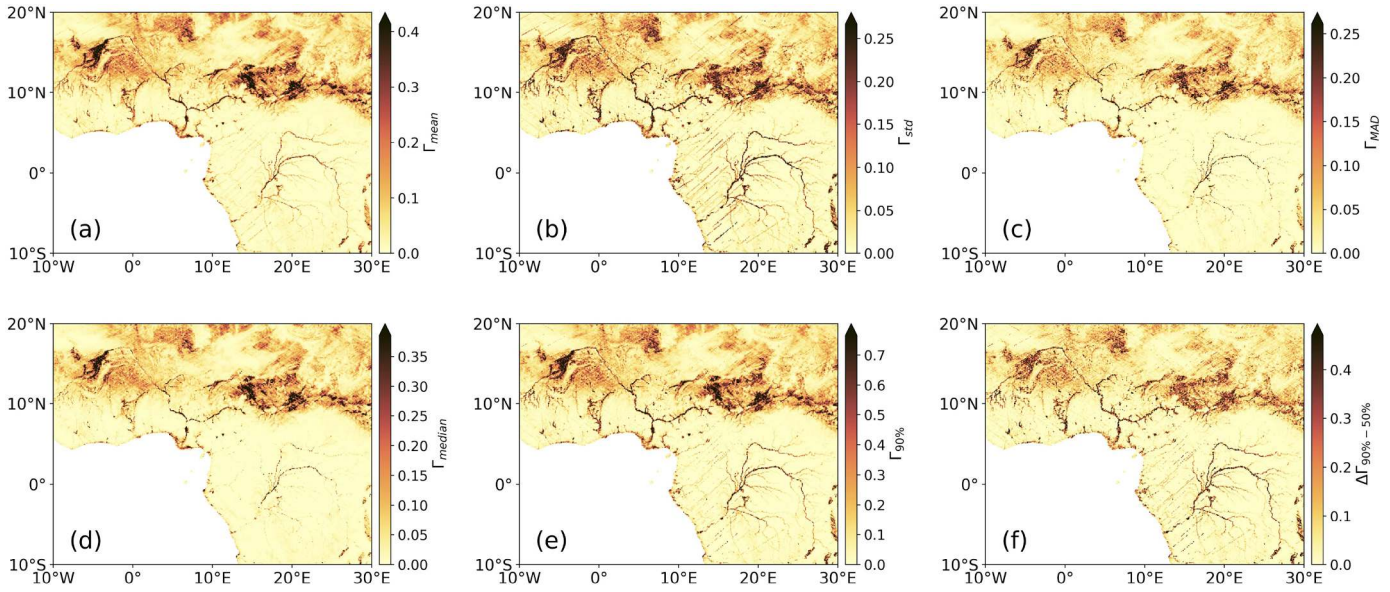

Fig. 4. Parameters derived from CYGNSS reflectivity over western and central Africa, with the Congo, Niger and Chad hydrographic basins, at time step 2018–09-05. **(a)**Γ_{mean}, **(b)**Γ_{std}, **(c)**Γ_{MAD}, **(d)**Γ_{median}, **(e)**$\Gamma_{90\%}$, **(f)**$\Gamma_{90\%-50\%}$.

flooded regions and rivers, which are smooth reflecting surfaces. All the observables show high values over large wetlands, such as the Inner Niger Delta (IND) and the floodplains around Lake Chad. Over smaller water bodies, the 0.1° pixels are composed of a fraction of water with non-inundated areas dominating around. As a consequence, Γ_{median} is low as it comes from an observation over soils non-covered with water, while $\Gamma_{90\%}$ corresponds to the highest reflectivity values associated to water bodies. For this reason, the detection of rivers is lower using Γ_{median} and Γ_{MAD} than with other variables, as it is observed along the streams of the Congo and Niger rivers and their tributaries. The interest of Γ_{median} and Γ_{MAD} is their robustness to noise, when compared to Γ_{mean} and Γ_{std}, respectively. Finally, the $\Gamma_{90\%-50\%}$ parameter is relevant, because it discriminates well the river streams (high $\Gamma_{90\%}$ and low Γ_{median}), the large water bodies and floodplains (high $\Gamma_{90\%}$ and high Γ_{median}), and the non-flooded areas (low $\Gamma_{90\%}$ and low Γ_{median}).

The maps of $\Gamma_{90\%}$ and $\Gamma_{90\%-50\%}$ in the full coverage of CYGNSS

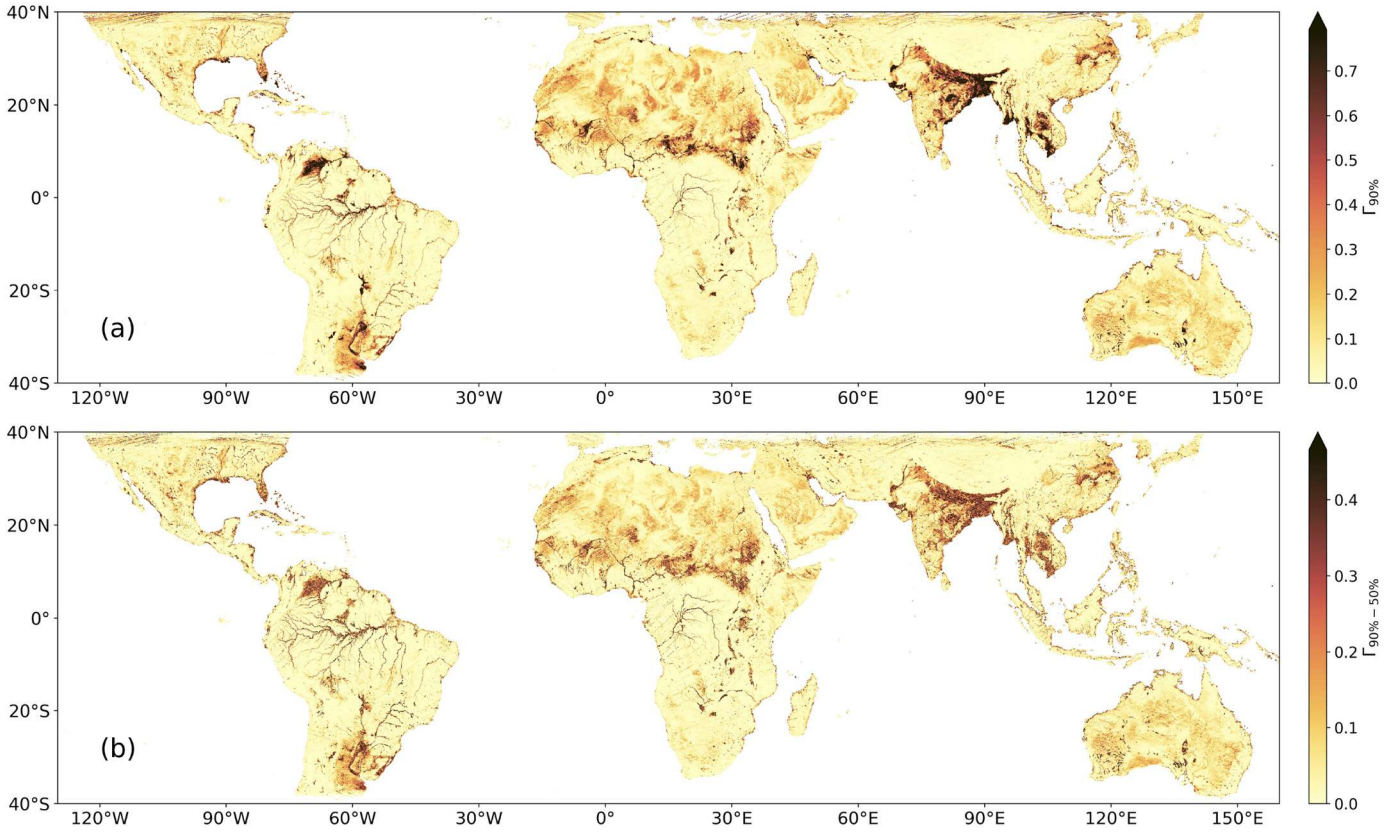

Fig. 5. Values at time step 2018–09-05 of CYGNSS-derived parameters: **(a)**$\Gamma_{90\%}$, **(b)**$\Gamma_{90\%-50\%}$.

mission are shown in Fig. 5. The Amazon, Parana, Congo and Niger Rivers and their tributaries, as well as other smaller streams in the pantropical area, are clearly delineated for both parameters. High reflectivity values are also obtained over the major floodplains in South America, Africa or Southeast Asia. Finally, a strong specular scattering is observed in the arid regions over bare soils, such as in the Sahara, the Arabian Peninsula, and Australia (see Section 5.4, discussion). Due to the homogeneity of the land cover in these areas, high values are obtained for both $\Gamma_{90\%}$ and Γ_{median}. As a consequence, the difference $\Gamma_{90\%-50\%}$ is low, which improves the separation between water bodies, floodplains and bare soils. For all the aforementioned factors, $\Gamma_{90\%-50\%}$ and $\Gamma_{90\%}$ were used along with Γ_{median} for further analysis in this paper.

3.3. Clustering

We clustered the CYGNSS-derived parameters $\Gamma_{mean}, \Gamma_{std}, \Gamma_{median}, \Gamma_{MAD}, \Gamma_{90\%}$ and $\Gamma_{90\%-50\%}$ using the unsupervised K-means technique (Macqueen, 1967). Starting from an initial set of centroids, it assigns all the points to the closest cluster center and then calculates the mean value of each cluster. This defines a new set of centroids, and the algorithm iterates until convergence. The K-means++ method is used to select the initial cluster centers, as it performs better and provides a good stability of the final solutions (Arthur and Vassilvitskii, 2007). The first centroïd is chosen randomly at this step. Then, the other centroids are iteratively determined between all the dataset points, with a probability equal to the inverse distance between a point and the closest centroïd, the weights being adjusted every iteration. This ensures a stable distribution of the cluster centers that is representative of the dataset.

In this study, we aim to detect inundated areas whatever their seasonality, with a time-series based approach. This is not possible using the K-means clustering with Euclidean distance, which makes the algorithm extremely sensitive to shifting and distortion in time. To overcome this limitation of the default K-means algorithm, we used the Dynamic Time Warping (DTW) similarity measure (Berndt and Clifford, 1994; Müller, 2007) which is implemented in the python package *tslearn* (Tavenard et al., 2020).

Given two time series $X = (x_1, x_2, \ldots, x_N), N \in \mathbb{N}$ and $Y = (y_1, y_2, \ldots, y_M), M \in \mathbb{N}$, the optimization problem associated with DTW is formulated as follow:

$$DTW(X,Y) = min_{\pi} \sqrt{\sum_{(i,j)\in\pi} d(x_i, y_j)} \tag{3}$$

where $\pi = [\pi_0, \pi_1, \ldots, \pi_K]$ with $\pi_k = (i_k, j_k)$ is called a warping path, and $d(\cdot,\cdot)$ is a distance metric (by default the Euclidean distance is used). The warping path must satisfy the following conditions:

(i)Boundary condition: $\pi_0 = (1,1)$ and $\pi_K = (N,M)$
(ii)Monotonicity condition: $i_k \leqslant i_{k+1}$ and $j_k \leqslant j_{k+1}, \forall k \in [1, K-1]$
(iii)Step size condition: $i_{k+1} - i_k \leqslant 1$ and $j_{k+1} - j_k \leqslant 1, \forall k \in [1, K-1]$

The minimization of all the potential warping paths according to the distance metric $d(\cdot,\cdot)$ results in the optimal warping path, which is here expressed as $DTW(X,Y)$. The optimal warping path is computed between the K-means cluster centers and every pixel at each iteration of the algorithm, and permits a shift and a distortion in time between the time series. This makes the DTW similarity measure well suited for our study. When applied to the detection of inundations, it allows to identify similarities in flood patterns from distinct regions of the world, with differences in phase and intensity. One limitation in our case is the boundary condition (i), because we only use one year of CYGNSS data to detect floods with a yearly seasonal cycle. Thus, we implemented in our methodology a padding of the CYGNSS dataset over 3 consecutive years to limit the boundary effects. The evaluation of a 3 to 5-year dataset of CYGNSS observations would be optimal. However, the choice of a padding was driven by constraints on our computation capacities. It assumes a constant seasonality, which seems reasonable. The methodology could still be extended to the full 5-year dataset of CYGNSS in the future.

The choice of an optimal number of centroids in the K-means algorithm can be discussed. Due to the important computational resources needed by the DTW similarity measure, we were not able to calculate any metric as the Calinski-Harabasz score (Calinski and Harabasz, 1974) and the Silhouette score (Rousseeuw, 1987) with this version of the algorithm. We finally adopted an empirical approach for the determination of the optimal number of clusters, and we tested our methodology with a number of classes ranging from 2 to 8. The results are evaluated graphically with the maps and time series of labelled pixels, and numerically with confusion matrices between CYGNSS clusters and flood reference classes based on Tootchi et al. (2019). We also evaluated the capability of each of the 6 CYGNSS-based parameters presented in Fig. 4 to detect floods and water bodies, either individually or associated to each-other. The most interesting results are obtained using $\Gamma_{90\%-50\%}$, which associates the information of Γ_{median} and $\Gamma_{90\%}$ into a unique variable. We therefore only present the results based on this parameter, although all the configurations were evaluated.

3.4. Sensitivity to the choice of initial centroids

Although the use of K-means++ algorithm produced stabler results than a pure random selection of the initial centroids, some differences were observed between successive versions of the clustering. We computed 30 times the K-means++/ DTW algorithm with 4 clusters and the $\Gamma_{90\%-50\%}$ parameter (the best configuration, see Section 4), to perform a sensitivity analysis. We extracted the inertia (sum of distances of samples to their closest cluster centers) at each iteration. We also calculated the percentage of correspondence between the labelled pixels from every pair of distinct clustering results. If the correspondence reaches 100%, the two versions are exactly the same. Table 3 presents some statistical parameters for both the inertia and the correspondence, which are the mean, median, min, max and std values. As can be seen, the inertia covers a very low dynamic range from 0.1005 to 0.1055 with a standard deviation of 10^{-3}, indicating that the algorithm converges toward similar solutions, whatever the choice of the initial centroids by the K-means++ method. The dynamic range of the correspondence is higher ($\sim 80-99\%$), with a median value of 94.57% and a standard deviation of 4.07%, respectively. The results are usually quite stable, in the range of 92%-99%, with the exception of few ones. Graphically, the different solutions found in 30 iterations also look very consistent.

We finally achieved a second round of the K-means/ DTW clustering, which was based on the results of the sensitivity analysis. There, the 4 initial centroids were obtained by averaging the cluster centers of the 30 outputs obtained with K-means++ initialization. This step permits to obtain a robust classification that does not vary through the random choice of the initial centroïds, but is still based on unsupervised techniques. All the clustering results presented in the following section are computed that way.

Table 3
Sensitivity of the clustering result to the choice of initial centroids by the K-means++ algorithm. We calculated the inertia of each clustering, and the correspondence between every pair of distinct clustering results. We present here the mean, median, minimum, maximum and standard deviation values of both parameters.

Result	mean	median	min	max	std
Inertia	0.1027	0.1026	0.1005	0.1055	0.0010
Correspondence (%)	93.76	94.64	80.26	98.78	4.01

4. Results

4.1. Clustering of CYGNSS reflectivity

We clustered the CYGNSS $\Gamma_{90\%-50\%}$ parameter with a number of classes k ranging from 2 to 8, in order to empirically determine the optimal number of clusters. For all the configurations, we plotted the maps of labelled pixels and we calculated a confusion matrix between the CYGNSS clusters and a reference set of water and inundation classes derived from RFWs maps (Tootchi et al., 2019). The definition of the reference classes is detailed in Table 4. Open water is grouped into a single class with a threshold at 10% corresponding to the empirical breaking point of permanent water (class A) vs. dry land (class E). Other classes, *i.e.* high floods (class B, maximum surface water extent SWE $\geqslant$ 80%), medium floods (class C, 40% $\leqslant$ SW $<$ 80%) and low floods (class D, 5% $\leqslant$ SW $<$ 40%), correspond to different levels of maximum inundation extent per pixel.

For a low number of clusters (*i.e.*, $k = 2$ or $k = 3$), the clustering problem was highly simplified as the separation of flooded vs. dry areas, according to the average values of $\Gamma_{90\%-50\%}$ related to the intensity of inundation. The results with $k = 4$ are more interesting, as the additional cluster represents the pixels in floodplains with a strong seasonal cycle during the year. Finally, the results with $k \geqslant 5$ show an increasing confusion between one or several clusters associated to both seasonal flood patterns and permanent water. The choice of $k = 4$ as the optimal configuration was based on both statistical outputs (confusion matrices), and graphical interpretation of the clustering results.

We present the confusion matrix and the results of the clustering with $\Gamma_{90\%-50\%}$ and $k = 4$ in Table 5 and Fig. 6, respectively. The output clusters are referred as C1, C2, C3 and C4 from the lowest to the highest average $\Gamma_{90\%-50\%}$ values. The map of labelled pixels is shown in Fig. 6.a, while Fig. 6.b–j represent the average time series of $\Gamma_{90\%-50\%}$ and $\Gamma_{90\%}$ for the pixels classified in C3 and C4, over 9 large river basins. The cluster C4 (in blue in Fig. 6) exhibits high values for both $\Gamma_{90\%-50\%}$ and $\Gamma_{90\%}$ over the entire time series, associated to a seasonality in phase with variations of inundation extent. The pixels in C4 are located on lakes, river streams such as the Amazon, Parana, Niger and Congo Rivers, and also some large floodplains with permanent water or high SM throughout the year. This cluster represents only 4.4% of the total labelled pixels, but includes 58.8% of the reference class A (open water), 32.7% of class B (high floods) and 19.0% of class C (medium floods), according to the results in Table 5 (percentage values in red).

Then, the cluster C3 in dark green shows variations from low to high values of $\Gamma_{90\%-50\%}$ and $\Gamma_{90\%}$, with a seasonal trend particularly visible in $\Gamma_{90\%}$. The pixels in C3 are located on the major floodplains in the pan-tropical area: the Llanos de Orinoco, Llanos de Mojos, Rio Branco, Pantanal and Parana floodplains in South America, the Inner Niger Delta (IND), Lake Chad and along the Nile in Africa, the Ganges–Brahmaputra, Indus, Irrawaddy, Yangtze and Mekong basins in Asia. The seasonality in C3 exhibits maximum values of $\Gamma_{90\%-50\%}$ and $\Gamma_{90\%}$ during summer in the Orinoco (Fig. 6.b), Niger (Fig. 6.e), Lake Chad (Fig. 6.f) and Ganges–Brahmaputra (Fig. 6.h) basins, all located in the northern hemisphere. On the contrary, in the Amazon (Fig. 6.c, across the equator) and Parana (Fig. 6.d, in the southern hemisphere) basins, the maximum values in C3 for $\Gamma_{90\%-50\%}$ and $\Gamma_{90\%}$ are obtained during spring. This validates the use of the DTW similarity measurement to cluster $\Gamma_{90\%-50\%}$ time series, for the extraction of flooded areas with distinct temporal cycles. For statistical results, C3 represents 8.3% of the total labelled pixels, and includes 23.8% of the reference class A, 40.6% of class B, 30.3% of class C, and 17.8% of class D (low floods) according to Table 5 (values in red). If we merge the two clusters C3 and C4, their combination represent only 12.7% of the world pixels, but 82.6% and 73.3% of the reference classes A and B, respectively. Thus, the identification of open water and flood signatures in CYGNSS reflectivity is highly reliable.

The cluster C2 has medium values of $\Gamma_{90\%-50\%}$ (not shown) corresponding to either water bodies and pixels with a fraction of water but not totally flooded (28.9% and 30.2% of classes C and D, respectively), or non-flooded regions from the reference class E. In Fig. 6.a, it is possible to locate C2 mainly over bare soils in arid regions such as the Sahara, the Arabian Peninsula and Australia. Finally, the cluster C1 is dominant (66.6% of the labelled pixels), and is interpreted as non-flooded areas although it includes about 9.5% of the pixels from both reference classes A and B.

In fact, some confusions are identified in the clustering despite a strong sensitivity to the presence of water. First, the pixels from reference class E which are theoretically non-inundated, represent 13.2% and 35.0% respectively (values in blue in Table 5) of the clusters C4 and C3, exhibiting the highest $\Gamma_{90\%-50\%}$ values. Then, the cluster C2 is a mix of pixels located over flooded and non-flooded regions. This is mainly due to a strong specular scattering over rocky bare soils (see the Section 5.4, discussion for more details). And lastly, while the open water and high floods (classes A and B) are well detected, it is not the case for the medium and low floods (classes C and D). They are associated to a mixing of wetlands and dry areas, and are quite equally divided into the 4 CYGNSS clusters. All these misclassifications can be related to: i) the influence of other factors such as the vegetation, SM, elevation, topography and small-scale roughness in the GNSS-R signals over land, ii) the heterogeneity of land cover and flood occurrence inside a CYGNSS pixel at 0.1° spatial resolution, and iii) errors in the definition of the reference classes, either due to the choice of our thresholds or to the high uncertainties associated to the RFWs product from Tootchi et al. (2019) itself (see Section 2.2.1 for more details).

4.2. Comparison with ancillary datasets

We have used CCI Land Cover maps and Tootchi's RFWs dataset described in Section 2 to investigate how the land cover changes and the occurrence of water influence CYGNSS-derived clusters. In Fig. 7, these parameters are analyzed over three complementary study areas: the northern part of South America, between 20°S and 10°N, encompassing the Amazon, Orinoco, Tocantins and São Francisco river basins, as well as the Pantanal floodplains, the Titicaca and Poopó lakes (Fig. 7.a1–5); Western and Central Africa around the gulf of Guinea, between 5°S and 20°N, including the Niger River Basin, Lake Chad and the Cuvette Centrale of Congo (Fig. 7.b1–5); the Indian subcontinent including the Ganges–Brahmaputra and the Irrawaddy basins (Fig. 7.c1–5). These regions include all types of land cover present in the pan-tropical area, *e.g.* equatorial forests, floodplains under canopy layers or with herbaceous covers, croplands irrigated or not, savannas, bare soils, high mountain ranges.

There is a good correspondence between the DTW clusters C3 and C4 in Fig. 7.a1–c1 corresponding to areas with high values of $\Gamma_{90\%}$ and $\Gamma_{90\%-50\%}$, and the reference map of water extent derived from Tootchi et al. (2019) in Fig. 7.a2–c2. The delineation of permanent water bodies and in particular the streams of the Amazon, Congo and Niger Rivers as well as their tributaries is clear, mainly in cluster C4. The most extensive floodplains are also detected in C3 and C4. In South America, we identify the Llanos in the Northeast Orinoco Basin (around 7°N and 70°W), Llanos de Mojos in the Southwest Amazon Basin (15°S and 65°W), the Rio Branco floodplain in the northern Amazon Basin (0° and 63°W), and

Table 4
Definition of the reference classes of open water and floods with the RFWs dataset (Tootchi et al., 2019).

Reference class	Name/ interpretation	Open water %	Seasonal water %
A	Permanent water	⩾10%	–
B	High floods	< 10%	⩾80%
C	Medium floods	< 10%	⩾40% and < 80%
D	Low floods	< 10%	⩾5% and < 40%
E	Non flooded areas	< 10%	< 5%

Table 5
Confusion matrix between reference classes A to E defined in Table 4 and CYGNSS clusters C1-4 shown in Fig. 6. For each cell in the table, the proportion of pixels in this cell vs. the total number of pixels from the reference class (column Total) is shown in red, and the proportion vs. the total number of pixels from the CYGNSS cluster (line All) is shown in blue. As an example, the box A/ C1 contains 9.6% of the total pixels in reference class A, and 0.2% of all the pixels in CYGNSS cluster C1.

	C1	C2	C3	C4	Total
A	804 (9.6%, 0.2%)	650 (7.8%, 0.4%)	1993 (23.8%, 3.3%)	4911 (58.8%, 15.5%)	**8358** (1.2%)
B	2150 (9.5%, 0.4%)	3861 (17.1%, 2.6%)	9150 (40.6%, 15.2%)	7353 (32.7%, 23.3%)	**22514** (3.1%)
C	8037 (21.9%, 1.7%)	10620 (28.9%, 7.1%)	11121 (30.3%, 18.4%)	6980 (19.0%, 22.1%)	**36758** (5.1%)
D	41226 (43.3%, 8.5%)	28782 (30.2%, 19.2%)	16963 (17.8%, 28.1%)	8211 (8.6%, 26.0%)	**95182** (13.2%)
E	430066 (76.7%, 89.2%)	105628 (18.8%, 70.6%)	21092 (3.8%, 35.0%)	4160 (0.7%, 13.2%)	**560946** (77.5%)
All	**482283** (66.6%)	**149541** (20.7%)	**60319** (8.3%)	**31615** (4.4%)	**723758**

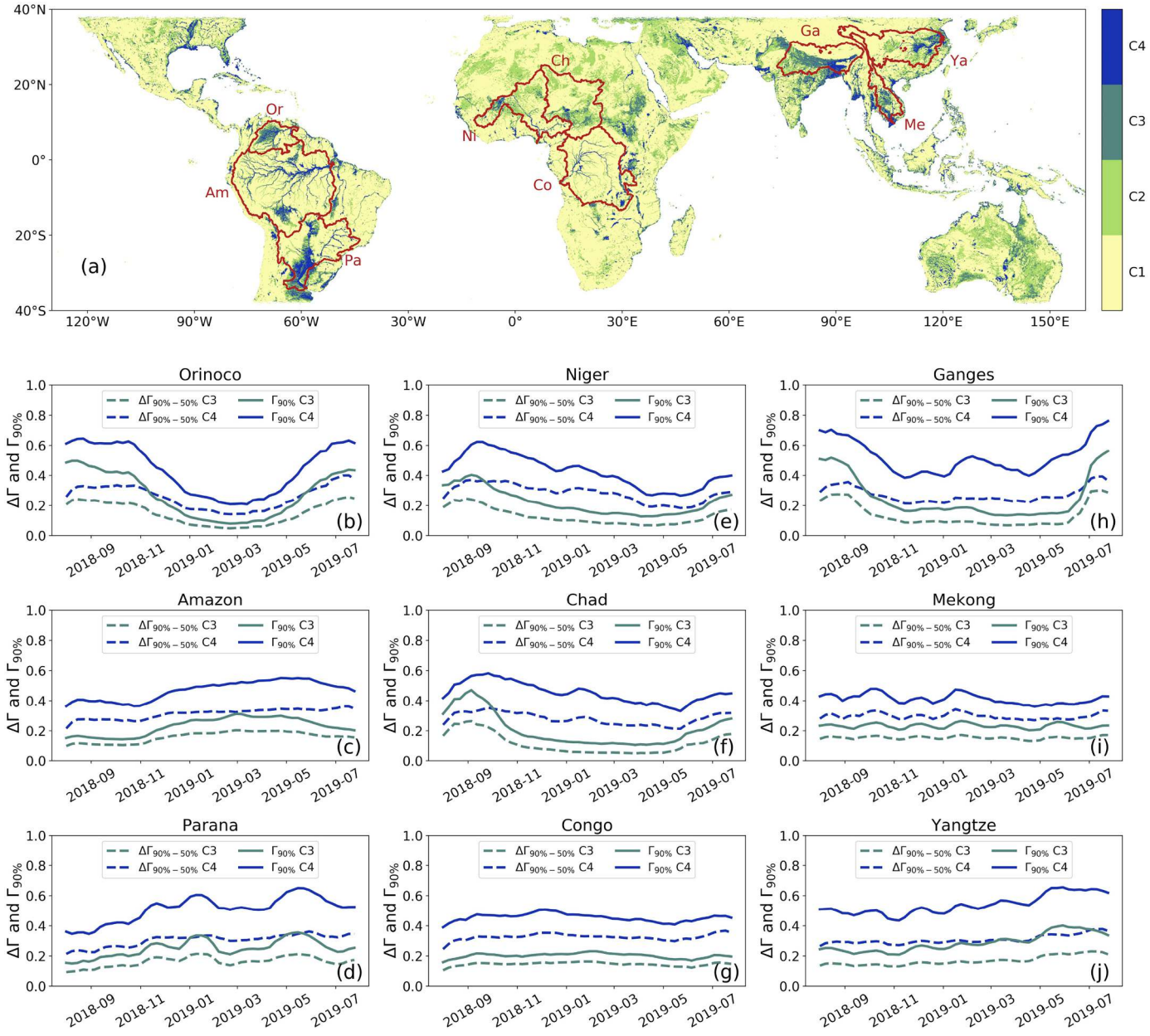

Fig. 6. Clustering of $\Delta\Gamma = \Gamma_{90\%-50\%}$ with 4 clusters. **(a)** Map of the labelled CYGNSS pixels using the K-means/ DTW algorithm with $k = 4$. **(b-j)** Average time series of $\Gamma_{90\%-50\%}$ and $\Gamma_{90\%}$ for clusters C3 and C4 in 9 large river basins: **(b)** Orinoco, **(c)** Amazon, **(d)** Parana/ La Plata, **(e)** Niger, **(f)** Lake Chad, **(g)** Congo, **(h)** Ganges–Brahmaputra, **(i)** Mekong, and **(j)** Yangtze. The boundaries of all the 9 basins are superposed with the map of labelled clusters in **(a)**, with the two first letters of the basin name allowing its identification.

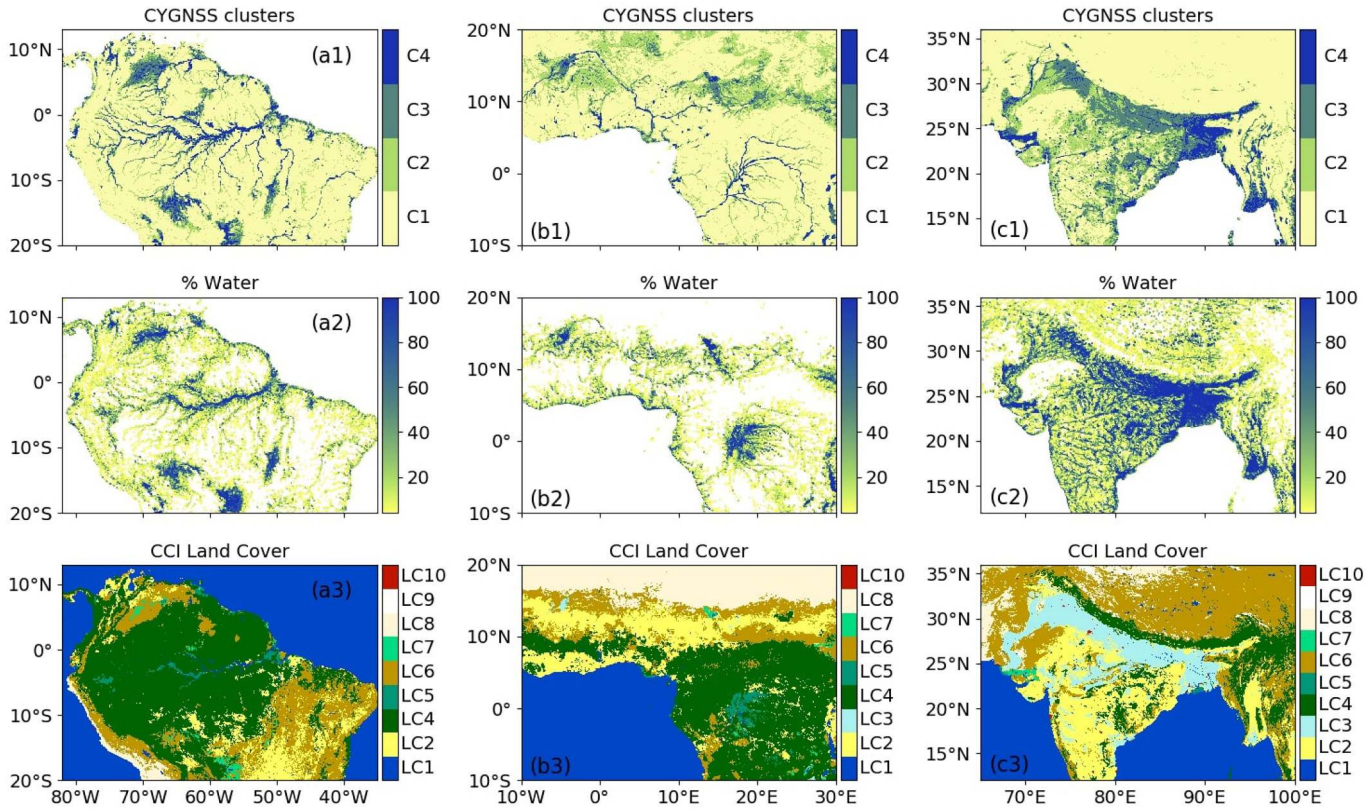

Fig. 7. Comparison of CYGNSS-derived parameters with ancillary datasets over the northern part of South America **(a1-3)**, Western and Central Africa with the Gulf of Guinea **(b1-3)**, and the Indian subcontinent **(c1–3)**. **(a1, b1, c1)** CYGNSS clusters, **(a2, b2, c2)** Percentage of water from Tootchi's RFWs maps (Tootchi et al., 2019), **(a3, b3, c3)** CCI Land Cover with the 10 aggregated classes defined in Table 1.

the Pantanal floodplains (18°S and 58°W). In Africa, the IND (15°N and 5°W) and the floodplains around Lake Chad (10°N and 15°E) are well detected, but the Cuvette Centrale of Congo (0° and 17°E) is not. In the Indian subcontinent, a strong seasonal signal is obtained and mainly associated to C3 over the irrigated croplands (land cover class LC3 in cyan) along the Ganges plain. The Ganges–Brahmaputra delta, the Irrawaddy River and their tributaries are also well delineated. The interesting point is that CYGNSS reflectivity maps show a good ability for monitoring heavy seasonal floods, almost regardless of the LC types. Yet, the densest canopies cause a strong attenuation or a diffusion of the L-band signal. Over the Cuvette Centrale of Congo (bottom right in Fig. 7.b1–3), we notice that CYGNSS likely underestimates the presence of floods nearby the streams of the Congo and Ubangi Rivers, if we refer to the RFWs dataset. Also, several lakes and wet areas in the Tibetan plateau in the upper right corner of Fig. 7.c1–3 are not even detected using CYGNSS. The estimation of the specular point location is of poor confidence over the Tibetan plateau, which is about 5000 m high (Gleason et al., 2020), thus CYGNSS DDMs are not centered on the specular point and likely contain only thermal noise and diffuse scattering from surrounding areas.

The same datasets are analyzed in Fig. 8 but on smaller regions, with a focus on some significant water bodies. Fig. 8.a1–5 is centered on Lake Titicaca (3812 m high), Fig. 8.b1–5 on Lake Victoria (1133 m high), Fig. 8.c1–5 on the Lower Mekong Basin including Tonle Sap, the largest lake of south east Asia, and Fig. 8.d1–5 on Lake Chad (both at low elevation). Several indications must be highlighted in Fig. 8. First, all water bodies are detected using CYGNSS, and mostly classified in C3 or C4 by the DTW clustering. These classes include a part of Lake Titicaca, Lake Poopó and the Salar de Uyuni located on the Altiplano at high elevation. Although the estimation of specular point location in this region has low confidence, a strong returned power is often observed with coherent reflection conditions. Then, it appears that the Tonle Sap and Lake Chad are classified in C4, while the surrounding floodplains are mostly classified in C3. This shows an example of the separation of permanent water vs. seasonal floods using CYGNSS. The Delta of Mekong mainly belongs to C4, likely due to continuous flooding or wet conditions in the paddy fields (Kuenzer et al., 2013). Thirdly, in Lake Victoria and Lake Titicaca, a spatial heterogeneity is observed with higher reflectivity (associated to C3 and C4) observed closer to the banks. This is attributed to the effect of the winds on large lakes and was previously reported in the literature, especially over Lake Victoria (Al-Khaldi et al., 2021b). The roughness of water surfaces due to the wind waves can nullify the coherent scattering assumption, and lead to a decrease of surface reflectivity (Chew and Small, 2020).

The CYGNSS reflectivity also depends on the dominant type of land cover inside a pixel. Fig. 9 presents the distribution of CYGNSS clusters into the aggregated LC classes defined in Table 1, either as the total number of pixels (Fig. 9.a) or the percentage (Fig. 9.b) in each class. Only pixels with a dominant type of LC covering at least 80% of the pixel are considered to avoid an influence of LC mixing. First, we notice a strong dominance at the global scale of non-flooded LC types (bare soils, dry shrubs and herbaceous, non-flooded forests, croplands non irrigated), with fewer pixels affected by permanent or seasonal water. These dry classes are composed quasi-exclusively of the CYGNSS clusters C1 and C2, with the lowest reflectivity values, and include also a part of C3, with high seasonal variations of $\Gamma_{90\%}$ and other CYGNSS-derived parameters. Then, we make several assessments over flooded areas. The flooded forests are mostly composed of C1 and C2, while the flooded shrubs and herbaceous are mainly composed of C3 and C4, with higher $\Gamma_{90\%}$ and $\Gamma_{90\%-50\%}$ values and seasonality (Fig. 6). In densely forested areas, the L-band reflected signals recorded by CYGNSS can be either strongly attenuated on their way to or from the reflecting surface, or scattered by the top of the canopy. In both cases, the ability of CYGNSS to detect inundations below the canopy is affected. The irrigated

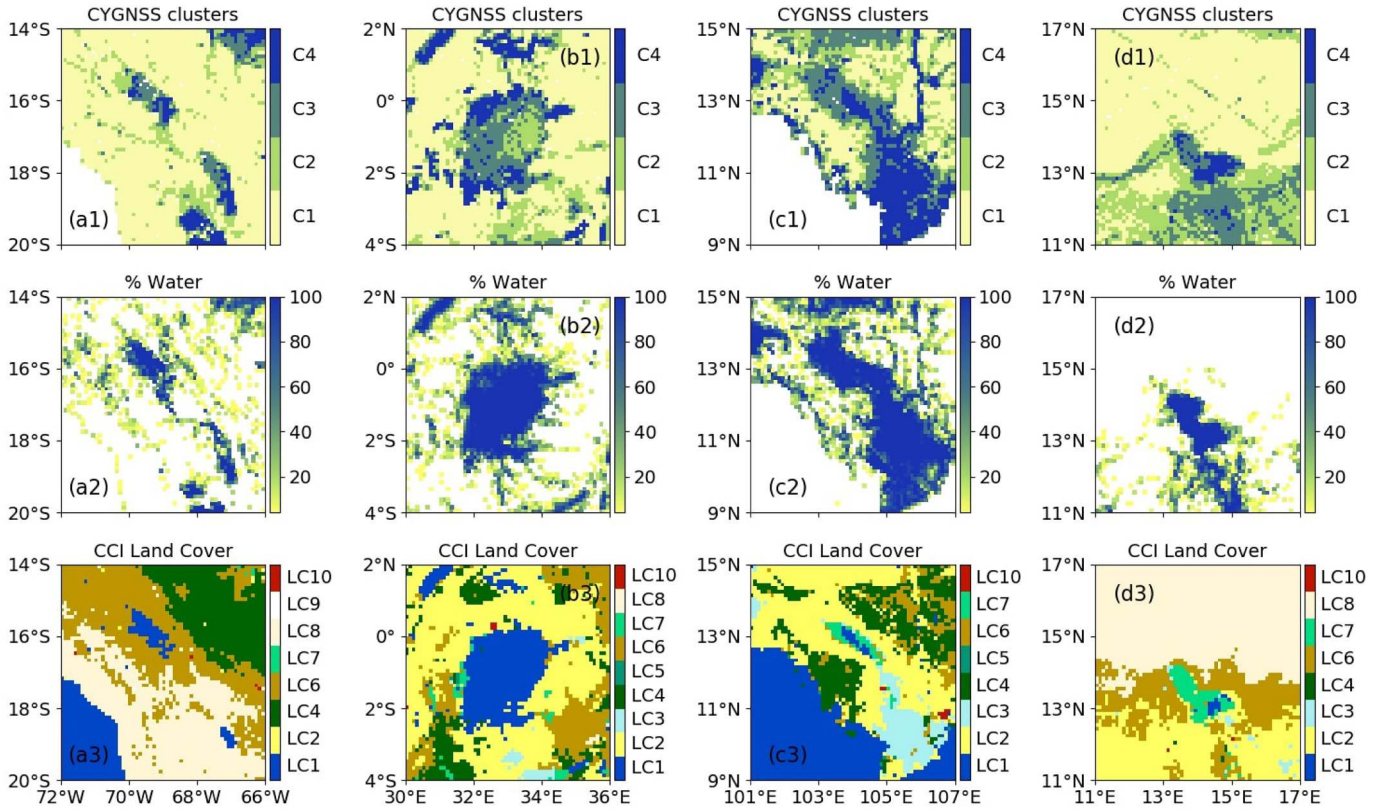

Fig. 8. Comparison of CYGNSS-derived parameters with ancillary datasets over small subsets centered on water bodies. **(a1-3)** Portion of the Andes centered on Lake Titicaca, **(b1-3)** Lake Victoria, **(c1–3)** Delta of Mekong with the Tonle Sap, and **(d1–3)** Lake Chad and a portion of Sahel. The same datasets are used than in Fig. 7.

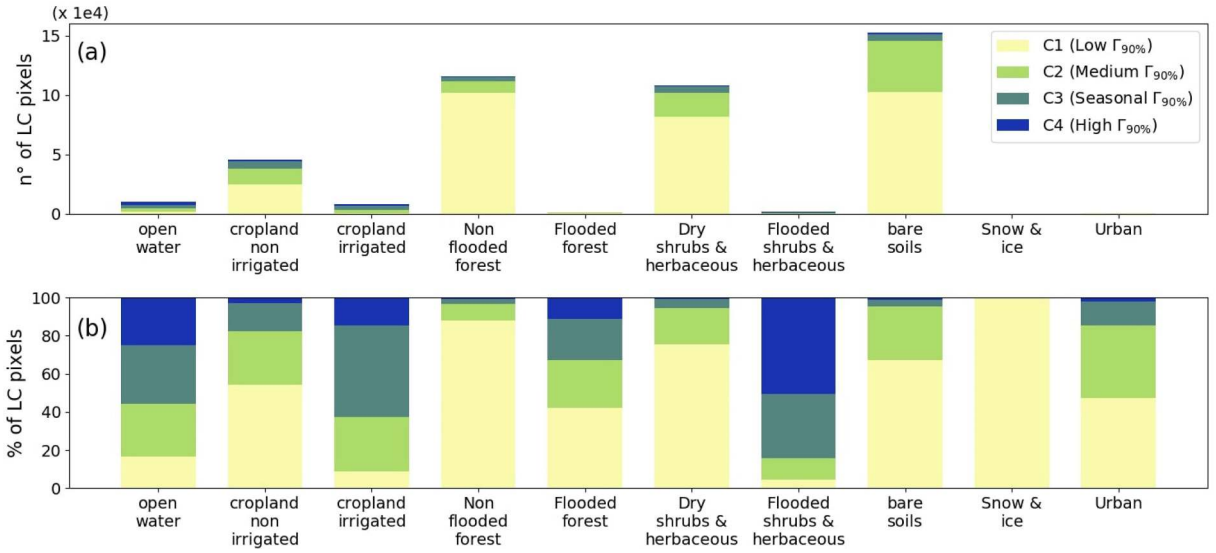

Fig. 9. Repartition of the CYGNSS clusters into the main Land Cover types (see Table 1 for the aggregation of CCI LC classes). **(a)** Number of pixels from each cluster, **(b)** Percentage of pixels from each cluster.

croplands, mainly located in the paddy fields in Southeast Asia, are dominated by the pixels with a strong seasonal signal in C3. On the other hand, the croplands non-irrigated are dominated by the dry pixels in C1. Both also contribute in the cluster C2, highlighting a mixing between these two classes and some misclassifications in Table 5. Non-flooded croplands and herbaceous covers affected by large seasonal rainfall present high SM values, responsible for a high CYGNSS reflectivity. Finally, open water bodies are divided between low and high reflectivity clusters. This should be linked to: i) the attenuation or scattering of the signals in equatorial forests and the non-detection of small streams, and ii) lower returned signal power on large windy lakes as shown in Fig. 8.

4.3. Comparison against flood products at regional and global scales

In this subsection, we analyze the correlations between CYGNSS-derived parameters and flood reference products. Because the spatial

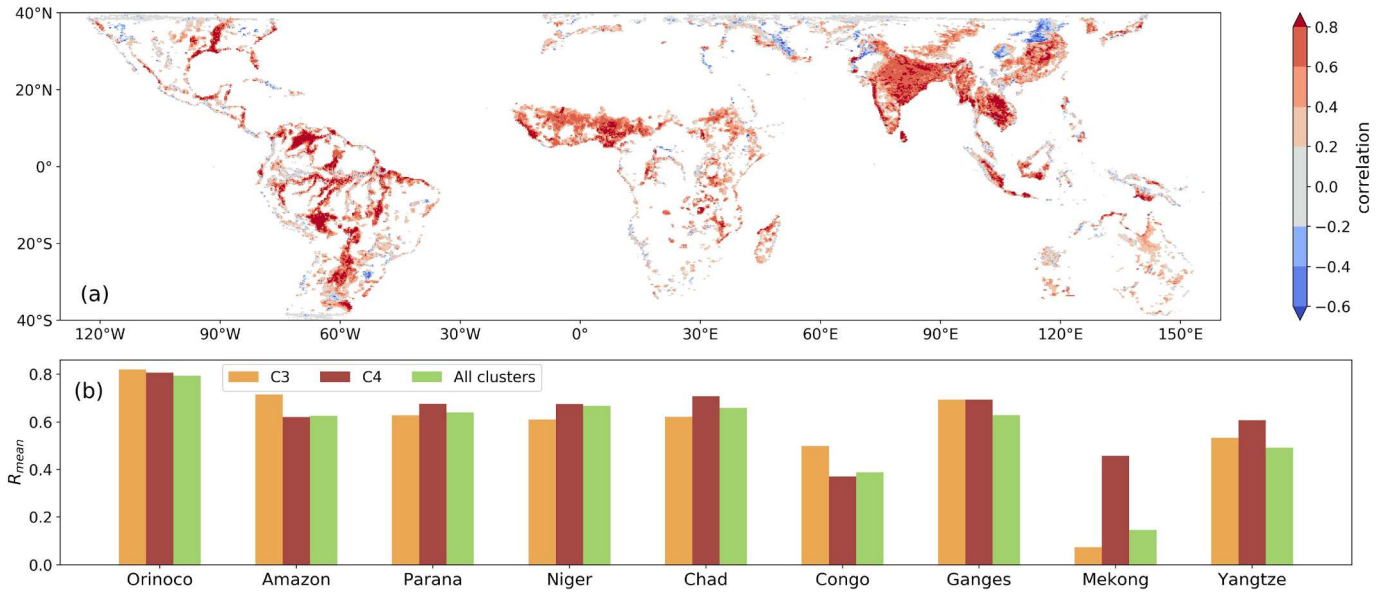

Fig. 10. Spearman correlations between CYGNSS Γ_{median} at 0.25° spatial resolution and GIEMS-2 from August 2018 to July 2019. **(a)** Map of correlations over the flooded areas in GIEMS (maximal SWE $>$ 20%), **(b)** Barplot of average correlations for CYGNSS clusters C3, C4 and all clusters in 9 river basins.

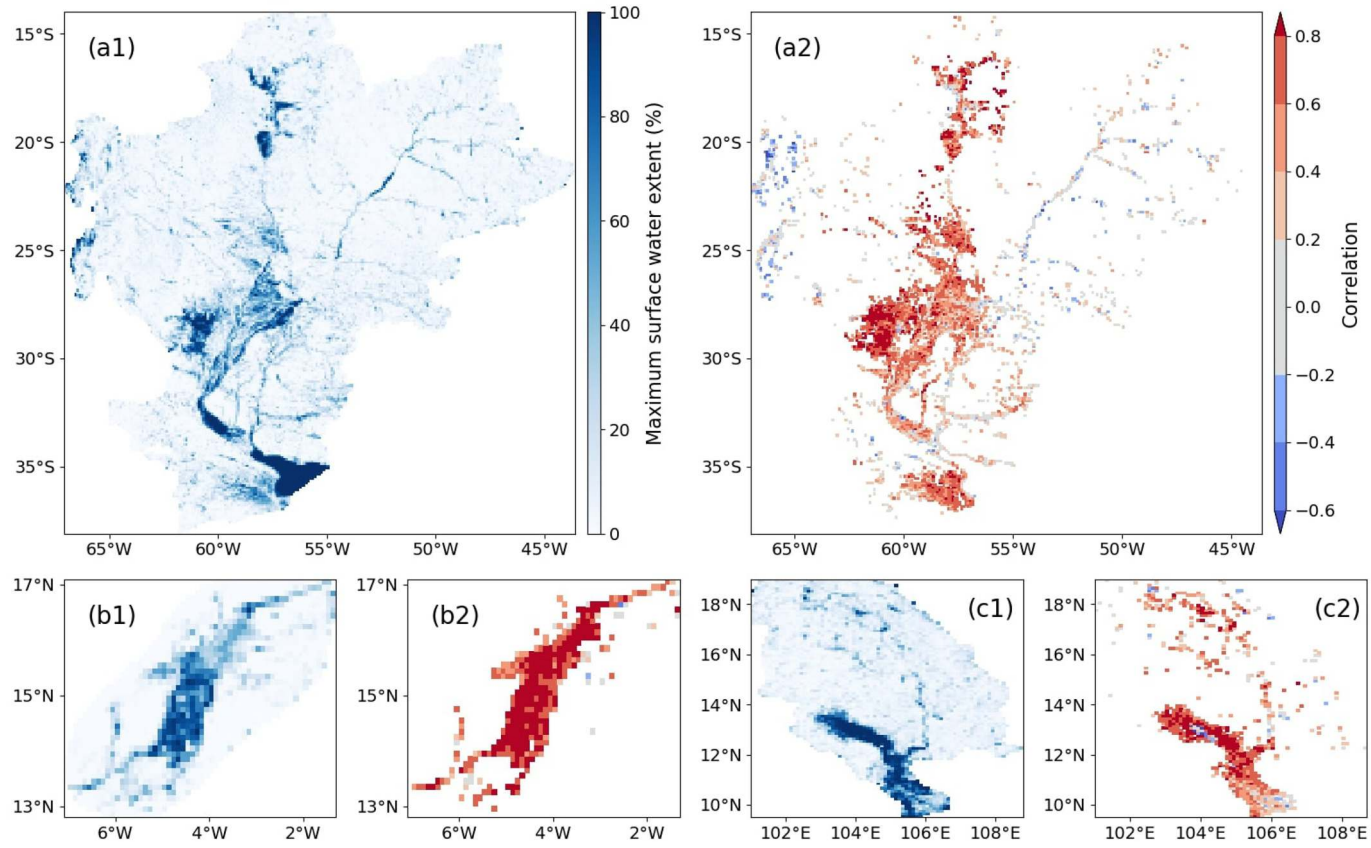

Fig. 11. Correlation between CYGNSS Γ_{median} and SWE derived from MODIS over 3 flooded regions. **(a1–2)** show the maximum water extent in the year and the pixel-by-pixel correlations between water extent and Γ_{median} over La Plata basin, **(b1–2)** are the equivalent plots over the Inner Niger Delta (IND), and **(c1–2)**are the equivalent plots over the southern Mekong basin (including the Delta of Mekong and Tonle Sap).

resolution of GIEMS is 0.25°, CYGNSS has also been gridded and clustered into an equivalent 0.25° grid for further analysis. We only present the results for Γ_{median} which show the best correlations, as $\Gamma_{90\%}$ is sometimes saturated with a fraction of water inside the pixel. The relationship between CYGNSS reflectivity and the fraction of water in the footprint is not linear (Chew and Small, 2020). Thus, we evaluated both Pearson's linear and Spearman's rank correlations. The latter evaluates whether the two variables are linked with a monotonic function, and should perform better for non-linear relationships. In fact, the results are slightly better using Spearman's R, in particular for the

regional comparisons against MODIS dataset (Section 4.3.2). Spearman's temporal correlations between Γ_{median} and the dynamic inundation maps at global and regional scales are therefore presented in Sections 4.3.1 and 4.3.2, and in Figs. 10 and 11. Finally, the spatial correlations against RFWs static inundation maps are presented in Section 4.3.3.

4.3.1. Global comparison of CYGNSS and GIEMS

Fig. 10 presents the Spearman's correlation coefficients in every pixel between CYGNSS Γ_{median} and GIEMS Surface Water Extent (SWE) at 0.25° (Fig. 10.a). A strong correlation ($R > 0.8$) is obtained in most of the major floodplains located in CYGNSS's spatial coverage, where the seasonality drives both signals due to the high temporal variations of SWE. It is the case in Southeast Asia and India, characterized by the annual monsoon events and irrigated paddy fields in the Ganges–Brahmaputra, Irrawaddy, Mekong and Yangtze basins. In South America also, the seasonal rainfalls cause large floods in Llanos de Orinoco, in the Amazon basin (including Llanos de Mojos and the Branco River), or in La Plata basin (including the Pantanal wetlands). On the contrary, poor results are obtained along some of the Amazon and Congo tributaries, in the upstream parts of these basins. This can be attributed to either the vegetation effect on Γ_{median}, or to a lower seasonality in both signals due to a continuous, high SWE, leading mechanically to a lower correlation. Low or even sometimes negative correlations are also observed in some coastal areas. The surface water estimations from GIEMS can be contaminated by the ocean contribution, although it is expected to be filtered out. Finally, negative correlations are also observed over some land areas (Fig. 10). We had a closer look at two specific targets in East Asia. These regions mostly comprise non irrigated croplands and herbaceous land cover, and they showed both a seasonal SWE cycle with low amplitude in GIEMS, and a low signal in CYGNSS with episodic variations. GIEMS likely overestimates SWE in these areas during the wet season, while CYGNSS peaks may be a response to SM or short duration floods following intense precipitation.

Fig. 10.b presents the average of CYGNSS vs. GIEMS Spearman correlations over the same river basins than in Fig. 6, for C3, C4 and all CYGNSS clusters. Cluster C3 in orange and C4 in red show an average correlation over 0.6 for most of the basins. The results in the Orinoco, Ganges, Niger, Lake Chad, Amazon and Parana basins are globally consistent because the water cycle is driven by seasonal floods and irrigation. Two main exceptions are the Congo and Mekong basins where the seasonal variations of reflectivity are very weak (below 0.1). The low correlations observed can be due to either a constant or random signal in both GIEMS and CYGNSS observables. In the Congo basin, the vegetation attenuates CYGNSS signals and provokes a lower detection of inundations below the canopy, and alternating rainfalls in the northern and southern hemispheres produce bimodal variations of inundation extent. In the Mekong basin, a continuous flooding or irrigation in paddy fields, including several parts of the Delta of Mekong (Kuenzer et al., 2013), certainly reduces the seasonality of the signals.

4.3.2. Regional comparison

The comparison with GIEMS SWE gives an indication of the correlation between CYGNSS reflectivity and flood dynamics at the global scale, but at the regional scale GIEMS is less adequate. First, it is gridded at 0.25° of spatial resolution, thus the information can be degraded when compared to CYGNSS 0.1° grid. Also, we gridded CYGNSS into at 0.25° for comparing against GIEMS, which is not the optimal spatial resolution of this product. Moreover, CYGNSS at 0.1° provides information about the flood dynamics up to a 7-day temporal resolution, and up to a daily time-scale at 0.25°, while GIEMS has a monthly time sampling. It also has uncertainties, especially in densely forested regions, where it is unclear whether flood signatures can be fully identified in the brightness temperatures at K-band from SSM/I. Finally, the large estimations of seasonal SWE in GIEMS, particularly in the Sahel region and the Indian subcontinent, may be overestimated due to the confusion between standing water and saturated soils during rainfall periods. For all these reasons, we also performed a regional comparison between CYGNSS Γ_{median} and the SWE derived from MODIS-based regional dynamic flood maps at 500 m spatial and 8-day temporal resolutions (see Section 2.2).

Fig. 11 presents the results of this comparison over 3 regions: La Plata basin including the Parana, Uruguay and Paraguay rivers (Fig. 11.a1–2), the Inner Niger Delta (IND) (Fig. 11.b1–2), and the Lower Mekong Basin (LMB) including Tonle Sap and the Delta of Mekong (Fig. 11.c1–2). Table 2 show that these regions have experienced above-average water level peaks during the study period, which are usually correlated with larger inundated areas. For every region, both maps of maximum flooded extent (as the percentage of water inside the pixel) and Spearman's correlation coefficients between the time series of SWE and Γ_{median} are shown. Only the correlations on pixels with a maximum flooded extent above 20% are presented. Over La Plata basin, we obtain medium ($R > 0.4$) to high ($R > 0.8$) values in the major flooded areas along the streams of the Parana and Paraguay rivers. In particular, the correlations are high in the Pantanal wetlands (~58°W, 20°S), in the Paraguay River and its confluence with the Parana River, and in wetlands close the Salado River in the region of Santa Fé (~61°W, 29°S). The values are lower ($R \sim 0.5$) in the major part of the Parana flooded savannas. This could be due to continuous high reflectivity in this region affected by permanent floods or very high SM content. Low correlation values are obtained along the streams of the Uruguay and the upper Parana rivers, as well as in mountainous regions along the Andes.

The correlations over the LMB are heterogeneous, with high values over the floodplains and irrigated croplands near the Tonle Sap and along the stream of the Mekong River. On the contrary, low and even negative correlations are obtained over the Tonle Sap itself, in some parts of the Delta and in isolated pixels. Fig. 12 present the time series of SWE and Γ_{median} in 4 neighboring pixels located near the Tonle Sap. The two left panel pixels (Fig. 12.a and.c) show a very high seasonality for both parameters, with a water extent varying from 0 to 100% during the year. It likely corresponds to seasonal floodplains, and the correlations calculated between the time series are around 0.9. On the contrary, the pixels plotted in Fig. 12.b and.d show limited variations in water extent throughout the year, with values systematically over 90% and 50%, respectively. They are mainly covered by permanent water from Tonle Sap. There, the variations of Γ_{median} are not correlated with variations in the inundation extent. Moreover, in all the pixels, we observe high

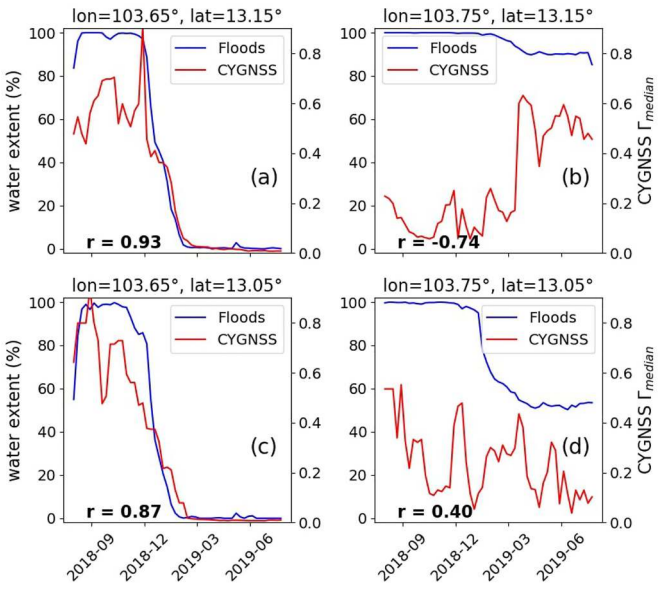

Fig. 12. Time series of CYGNSS Γ_{median} (in red) and SWE from MODIS (in blue) with their correlation, for four pixels near the Tonle Sap in the Mekong basin.

Table 6
Spatial correlations between CYGNSS Γ_{median} and the percentage of water derived from RFWs static inundation maps (Tootchi et al., 2019), in 9 river basins inside the area of coverage. The values of Γ_{median} are extracted at the time of the maximum flooded extent.

River basin	Orinoco	Amazon	Parana	Niger	Chad	Congo	Ganges	Mekong	Yangtze
Spatial R	0.76	0.77	0.75	0.52	0.51	0.58	0.75	0.66	0.74

fluctuations of Γ_{median} with an amplitude reaching ~ 0.4 and a period of one to several months. These fluctuations mostly occur when a large fraction of water is present inside the pixel. It could be linked to variations of water roughness over Tonle Sap, the greatest lake in southeast Asia. Moreover, similar time series of Γ_{median} with high amplitude fluctuations were found to be located close to the banks of Lake Victoria, where the direction and speed of winds can cause an alternation between coherent and incoherent scattering regimes. Unfortunately, we were not able to find wind speed data close enough to Tonle Sap to compare with Γ_{median} time series. The understanding of this phenomena could help further work on retrieving SWE using CYGNSS reflectivity or coherency. Other pixels in the Delta of Mekong show low correlation values, but large parts of this area are flooded or wet throughout the year (Kuenzer et al., 2013).

Finally, the results over the IND show a very good consistency between the time series of SWE and Γ_{median}. The correlations are greater than 0.8 in most of the pixels. It indicates a strong correspondence between the dynamics of floods and CYGNSS reflectivity in this region. It is certainly due to two reasons: i) the low vegetation cover cause lower uncertainties in both CYGNSS observations and the reference inundation maps, and ii) a strong seasonality is observed in the precipitation and flooding events, so Γ_{median} is low during the dry season unlike other floodplains in the LMB and the Parana.

4.3.3. Spatial correlations at the regional scale

Spatial correlation coefficients were also calculated between Γ_{median} and the SWE derived from static RFWs maps (Tootchi et al., 2019), at the time of the maximum flooded extent during the year. They are estimated for the 9 river basins presented in Figs. 6 and 10, and values are reported in Table 6. The spatial correlations in the Orinoco, Amazon, Parana, Ganges–Brahmaputra and Yangtze basins are high, ranging between 0.74 and 0.77, showing a good correspondence between the reference maps and CYGNSS reflectivity at the flood maximum extent. On the contrary, spatial correlations are lower in the other basins and especially in the Sahel region (0.52 for Niger, 0.51 for Lake Chad). This highlights the contribution of multiple factors in the scattering of GNSS signals over these areas, and maybe uncertainties in the RFWs dataset. While the temporal correlations indicate whether the variations of SWE are linked with an increase in CYGNSS reflectivity, the spatial correlations are particularly sensitive to a saturation of both the reference and CYGNSS signals during the wet season. High SM can provoke high CYGNSS reflectivity, and also an overestimation of the fraction of water estimated in GIEMS, whose uncertainty is further propagated in the RFWs dataset (see Section 2.2.1 for more details).

5. Discussion

The objective of this study was to assess the potential of CYGNSS reflectivity for a pan-tropical mapping of flood dynamics. Based on the results presented in Section 4, we discuss several points to pave the way to a CYGNSS-based inundation product. To begin with, we take a broader view on the interests and limitations of our methodology, as well as the uncertainties in the reference datasets considered. We then discuss the attenuation of GNSS-R signals by the vegetation, especially over tropical forests, and the feasibility of flood detection under dense canopies using CYGNSS reflectivity. We also point out the misleading role of high forward scattering in the specular direction in arid areas, due to changes in morphology, *i.e.* roughness for CYGNSS, but also lithology (sand vs. rock, see Section 5.4).

5.1. Results of the K-means clustering

We have used the K-means clustering algorithm with a Dynamic Time Warping similarity measure to perform the clustering of CYGNSS reflectivity time series. Other common unsupervised classification algorithms were tested such as Agglomerative Clustering or Birch, but the best results were obtained with K-means. The implementation of a DTW similarity measure instead of a simple Euclidean distance makes it more robust to shifts in phase and distortion in time between time series of reflectivity. Promising results were obtained and analyzed in this study, but the empirical approach led us to discuss our choices in the implementation of the methodology.

First, the optimal number of clusters needed to be determined. We were not able to calculate any criteria like the Calinski-Harabasz and the Silhouette scores with the implementation of the DTW similarity measure. We also tested our methodology with a common Euclidean distance in the K-means approach, which gave us an optimal number of 2 clusters with both the Calinski-Harabasz and the Silhouette scores. This is a highly simplified version of the phenomena as dry land vs. flooded areas, and does not give information on the seasonality of floodplains. Thus, we empirically determined the optimal number of clusters, which was found to be $k = 4$ (Fig. 6 and Table 5). The two clusters with highest reflectivity are mostly associated to the occurrence of water, one with permanent or long-lasting floods, the second with high seasonal variations. A third cluster is composed of a mix of wetlands and bare soils with medium reflectivity, and can be especially sensitive to an increase of CYGNSS reflectivity over non-flooded areas due to high SM content. The last cluster is mainly associated to dry soils, never flooded throughout the year. We also tried the clustering with several combinations of parameters from Fig. 4: 1) Γ_{mean} only, 2) Γ_{median} only, 3) $\Gamma_{90\%}$ only, 4) $\Gamma_{90\%-50\%}$ only, 5) Γ_{median} and $\Gamma_{90\%}$, 6) Γ_{mean} and Γ_{std}, 7) Γ_{mean}, Γ_{std}, $\Gamma_{90\%}$ and Γ_{median}. All the confusion matrices and the plots (similar to Table 5 and Fig. 6, respectively) were evaluated for each version of the clustering, leading to the determination of the best scenario.

To ensure the stability of our clustering, we performed a sensitivity analysis on 30 iterations of the K-means++/ DTW clustering. It was shown that the 30 results are quite stables, especially in terms of inertia (see Table 3). However, the random (although weighted) choice of the initial 4 centroids by the K-means++ algorithm leads to differences in the final results. In 25 results out of 30, very small differences are observed in the output labels, especially for clusters C3 and C4 associated to water bodies (see Fig. 6). The last 5 versions show lower agreement with each-other. The correspondence (*i.e.* percentage of pixels similarly labelled between two versions of the clustering) has a minimum of 80.26%, and a median of 94.64% (Table 3). As a consequence, we adapted our methodology to guarantee the reproducibility of the clustering. The mean time series of cluster centers C1 to C4 were computed from the 30 random DTW outputs, and selected as user-defined centroids for a new K-means/ DTW clustering. This final version was analyzed in Section 4.

Finally, the DTW was implemented to manage shifts in time between the flooding events in different regions of the world. However, it has a limitation due to the boundary condition. The DTW algorithm starts with the first index and finishes with the last index of each time series that are compared. As a consequence, the values of $\Gamma_{90\%-50\%}$ and other CYGNSS parameters at the boundaries of the time series play a determinant role for the clustering. We use only one year of CYGNSS observations because of the high computational capacities required to process and analyze this dataset, and due to the experimental nature of our

methodology whose interest needed to be proven before a global application to the 5-year CYGNSS dataset. To avoid the boundary effect, we implemented a padding of the 1-year time series of our parameters to create artificially a 3-year dataset. The results prove that this technique does not affect the ability of the K-means/ DTW clustering to extract a seasonal flood signal from both the wetlands in the northern (*e.g.* Orinoco, Rio Branco, Ganges, Yangtze, Mekong) and southern (*e.g.* Llanos de Mojos, Pantanal, Parana) hemispheres.

5.2. Reference inundation maps

The reference datasets used in this study are also sources of uncertainties. The static inundation maps from Tootchi et al. (2019) were used as a delineation of permanent water bodies and regularly flooded areas. We first tried to use the composite wetlands (CWs) maps as it contains both the regularly flooded wetlands (RFWs) and the groundwater-driven wetlands (GDWs). It seemed logical to us that GDWs also play a role in the scattering of GNSS signals, as they are an important source of SM in saturated soils. However, we noticed a strong saturation of the CWs maps in several river basins. As an example, the entire Amazon basin and a great part of Sahel were listed as partially or totally affected by floods, which is largely exaggerated. Also, spatial correlation coefficients calculated with CYGNSS Γ_{median} show higher values with RFWs maps both at the global and regional scales. For the Amazon basin, the spatial correlation is 0.77 between Γ_{mean} and the SWE from RFWs maps (Table 6), and is 0.48 with CWs maps. As a consequence, we decided to use the RFWs maps for comparison with CYGNSS datasets. Additionally, the thresholding of RFWs variables into the 5 reference classes defined in (Table 4) is debatable. The open water class (A) corresponds to pixels covered with, at least, 10% of open water. When plotting Γ_{mean} as a function of open water extent, the Γ_{mean} values are saturated above this threshold. On the contrary, the function of Γ_{mean} vs. inundation extent is linear, with no saturation of the signal when increasing the fraction of flooded areas in the pixel. The separation of reference classes B, C and D is then arbitrary, with thresholds corresponding to high, medium or low inundation extent in the pixel.

The uncertainties of GIEMS are mainly associated to its low spatial resolution (0.25°), leading to a low sensitivity to small water bodies. It is also influenced by dense vegetation covers, and high seasonal SM content confounded with surface water in several regions (Sahel, Indian subcontinent). These are sources of uncertainty for our analysis because even Tootchi's RFWs dataset is based on GIEMS-D15 (Fluet-Chouinard et al., 2015), derived from GIEMS itself. Still, the comparison of CYGNSS vs. GIEMS and MODIS estimated SWE looks consistent at the regional scale, although the spatial resolutions are different (0.25° for the first, 0.1° for the latter).

5.3. Reflectivity over vegetated areas

The detection of floods under equatorial forests remains difficult even with the use of microwave signals. It is a limitation of GIEMS (Prigent et al., 2020) for example. The use of CYGNSS data is promising because: i) L-band GNSS signals penetrate deeper the canopy than shorter wavelengths/higher frequencies in the microwave domain, ii) it provides information at a higher spatiotemporal resolution than the passive microwave sensors. Several studies took advantage of these characteristics. Rodriguez-Alvarez et al. (2019) classified CYGNSS corrected Signal-to-Noise Ratio (SNR) as flooded vegetation (FV), open water and dry land. This study was performed in a small subset of the Amazon basin and showed a detection of FV around 70%. Then, Carreno-Luengo et al. (2020) studied the relationship between CYGNSS observables, including the reflectivity Γ, and AGB over subsets of the Congo and Amazon basins. A polynomial fit was used to characterize the vegetation attenuation at several incidence angles, and to further derive CYGNSS-based maps of AGB over equatorial forests. These maps were consistent with reference datasets up to ∼ 350 Mg/ha. Also, Li et al. (2021) derived the coherency of Beidou-3 raw IF tracks over flooded areas along the Mississippi River, in densely vegetated areas. High coherency, which is linked to the presence of water below the canopy, was obtained for AGB up to 200–300 Mg/ha. And finally, the PR defined in Al-Khaldi et al. (2021a); Al-Khaldi et al., 2021b as a coherency proxy has shown low sensitivity to vegetation, allowing the detection of small river streams obscured by trees.

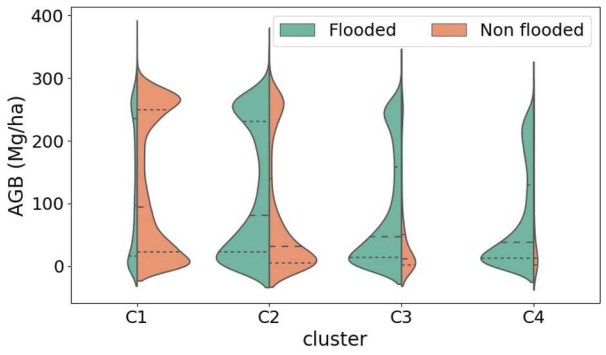

Fig. 13. Distribution of the AGB for flooded and non-flooded pixels over South America in the 4 CYGNSS clusters. The dashed lines represent the quartiles of the distributions. Flooded pixels are defined with a fraction of water greater than 0% in the 0.1° regridded RFWs dataset.

In Figs. 6.a and 7, the streams of the Congo and Amazon rivers and their tributaries are well delineated. We can still question whether CYGNSS is able to monitor nearby floodplains under vegetation cover. Fig. 13 presents the distribution of AGB in CYGNSS clusters C1 to C4, for both flooded and non-flooded pixels. All pixels with an occurrence of water greater than 0% in the 0.1° regridded RFWs dataset are considered as flooded, which likely overestimates the inundated areas. Only pixels in South America are considered in this figure, to obtain a balance between flooded and dry pixels, and between high and low AGB values. The dynamic range of AGB is approximately 0–300 Mg/ha and shows two peaks for all clusters. The first one, between 0 and 50 Mg/ha, corresponds both to dry pixels with few vegetation, and to large floodplains with herbaceous-type land cover, as in the Llanos de Orinoco, Llanos de Mojos, and the Pantanal wetlands (see Fig. 7 for dominant land cover types in South America). The second one, between 200 and 300 Mg/ha, corresponds to dense forests mainly located in the Amazon basin. These forests are either dry (cluster C1), wet or flooded (clusters C2 to C4). In particular, flooded pixels in C3 and C4 mostly represent the large floodplains and open water (see Figs. 6, 7 and Table 5), with low AGB. However, the fourth decile (*i.e.* the top 25% of AGB values) in flooded pixels from C3 and C4 ranges from ∼150 Mg/ha to ∼300 Mg/ha. The results in Fig. 13 show that the K-means/ DTW clustering based on CYGNSS reflectivity is able to detect either flooded areas below the canopy or narrow river streams surrounded by vegetation, in pixels with dense forests and average AGB as high as 250–300 Mg/ha. Some sources of uncertainties can be related to: i) signals from wet, saturated soils without standing water, ii) errors in the definition of flooded areas using RFWs dataset, and iii) the attenuation of GNSS-R signals by the vegetation, especially at high incidence angle, leading to a potential extinction of the coherent component on which we base our analysis (Al-Khaldi et al., 2021a; Loria et al., 2020). The results presented in Fig. 9 show a lower detection of the inundations under forest land cover types, when compared to an herbaceous cover.

5.4. Reflectivity over bare soils

In Figs. 5 and 6.a, we notice the heterogeneity of $\Gamma_{90\%}$, $\Gamma_{90\%-50\%}$ and CYGNSS clusters over arid regions, where bare soils is the dominant land cover type. In particular, reflections over deserts in the Sahara and the Arabica Peninsula have a high dynamic range, with $\Gamma_{90\%}$ values ranging

from 0 to ~0.4. As the time series of $\Gamma_{90\%}$ and $\Gamma_{90\%-50\%}$ are almost constant in the year due to a very low moisture, these areas can be interpreted as open water bodies, in particular when the reflected signals over the latter are attenuated as for wind conditions over lakes (see Fig. 8). An important forward scattering of GNSS-R signals has already been observed in the literature, and mainly linked to areas of low small-scale roughness. In particular, the water mask given by Al-Khaldi et al. (2021a) also had false alarms over desert regions characterized by (almost) flat surfaces, based on their PR coherency metric. They handled the problem with a filtering of the locations showing a recurrent coherence, and a 0% occurrence of water in the mask from Pekel et al. (2016). This type of filtering could be easily applied to our reflectivity dataset, and is a requirement for a further inversion of a fractional water extent product.

The state of the art beyond the scope of CYGNSS-related studies shows that anomalies in the radar signals over arid regions are well known for several active and passive sensors, and can be related to the geological and lithologic properties. Prigent et al. (2015) analyzed the multiangle backscattering at Ku-band from Tropical Rainfall Measurement Mission/ Precipitation Radar (TRMM/ PR) and QuikSCAT in the inter-tropical zone covered by CYGNSS. They showed that sand dunes have lower backscattering coefficients when compared to rock deserts, because of an important volume scattering in dry sand and a relatively low large-scale roughness. Similar effects were also observed at C-band using ASCAT (Fatras et al., 2015). This phenomenon increases with an increasing incidence angle. Similarly, the microwave emissivities over arid regions show higher values for sand dunes due to the contribution of both the surface and deeper soils to microwave emissions (Jiménez et al., 2010). GNSS-R forward scattering is likely subjected to the same phenomena, as the spatial agreement between $\Gamma_{90\%}$ and $\Gamma_{90\%-50\%}$ maps and TRMM/ PR-based maps of backscattering coefficients in Prigent et al. (2015) is good.

5.5. Guidelines for a CYGNSS-based dynamic product of SWE

This study shows that both permanent water and seasonal floodplains can be monitored using CYGNSS reflectivity. In particular, the 0.1° spatial resolution and 7-day temporal resolution grid merges by time and location the bistatic observations, and offers the possibility to compute time series of several statistical parameters (see Section 3). In particular, $\Gamma_{90\%}$ is very sensitive to the presence of a fraction of water, and $\Gamma_{90\%-50\%}$ can be used to extract pixels with recurrence or occurrence of water. Also, the comparison of Γ_{mean} with the fraction of water in the pixel, either from static or dynamic estimations, shows an interesting linear trend. This linearity could be further exploited to produce a CYGNSS-based SWE product, covering the pan-tropical area.

For this, the confounding effects of several geophysical parameters need to be removed. False alarms over deserts can be avoided with a simple filtering as in Al-Khaldi et al. (2021a). Then, a correction of reflectivity for the vegetation attenuation must be implemented, which is usually performed using the incidence angle and ancillary L-band Vegetation Optical Depth (L-VOD) data provided by radiometers such as SMOS (Wigneron et al., 2021) or SMAP (Konings et al., 2017; Li et al., 2022b). Also, GIEMS-2 (Prigent et al., 2020), assumes the linearity of SWE retrieval under given vegetation conditions, and the dataset is therefore binned depending on the values of vegetation parameters. Based on this idea, CYGNSS dataset could be binned with either NDVI, AGB or VOD to improve the linearity of Γ_{mean} vs. SWE relations. The effect of small-scale and large scale roughness on the performances of future SWE product could be important to study. Water roughness over large lakes is a limitation due to a dominant incoherent scattering regime, so it must also be taken in account. As we focus on dynamic estimations of SWE in floodplains, recurrent open water bodies could be filtered out as they are well delineated using other remote sensing sensors. Finally, the confounding effect of SM can lead to an overestimation of flooded areas in regions affected by large seasonal rainfall. Further investigations could use the coherence proxy defined in Al-Khaldi et al. (2021a); Al-Khaldi et al., 2021b, in helping to filter out the pixels with strong reflectivity that is associated to very wet soils with no surface water.

6. Conclusion

We have analyzed CYGNSS land surface reflectivity to evaluate its potential for a global mapping of flood dynamics. A 0.1° spatial resolution and weekly time sampling of CYGNSS reflectivity was found to be the most suitable compromise between high spatial and temporal resolutions. We have used a K-means clustering technique with DTW similarity measure to separate: i) the low, constant reflectivity signals from dry land, ii) the high signals from open water and long-lasting inundations, and iii) the seasonal signals associated to large floodplains. Static and dynamic inundation maps along with other ancillary datasets were used to analyze the clustering results. The largest water bodies, floodplains and irrigated croplands areas are detected. Various sources of misclassification are identified. The flooded areas are mostly detected in absence of vegetation or under herbaceous cover, but less accurately under forests. The concordance between CYGNSS parameters and the static inundation maps is not good in the Cuvette Centrale of Congo in particular. The detection of water bodies is also weakened under windy conditions over large lakes, as it was already reported by several studies. In the deserts, a low reflectivity is observed on sand dunes where the penetration of microwave signals is high, and a high reflectivity is observed on flat, rocky regions where a specular scattering was found to create false alarms. However, CYGNSS clusters and parameters make possible to identify the main floodplains and open water areas, including samples in areas with AGB as high as ~300 Mg/ha along the streams of Amazon and Congo rivers. The spatial correlations between Γ_{median} and static inundation maps were calculated at the time of the maximum SWE. They show high values (R ranging from 0.74 to 0.77) in 5 large river basins and lower values in Congo, Lake Chad and Niger, highlighting the contribution of multiple known factors (vegetation, soil moisture, small and large scales roughness, type of soil, SWE) in the GNSS-R forward-scattered signals. Temporal correlations were also calculated between Γ_{median} and SWE from either GIEMS or regional MODIS-based inundation maps. Regional comparisons over the IND, Parana and Mekong basins perform well and are consistent with global comparison using GIEMS. High values ($R > 0.8$) are obtained on the principal floodplains in the CYGNSS coverage, with once again the exception of the Cuvette Centrale of Congo. Lower correlations are obtained in the Amazon and Congo tributaries, either due to an attenuation of GNSS-R signals by vegetation canopies, or to the uncertainties of reference datasets in these areas.

Finally, our results show that a global mapping of inundation dynamics using CYGNSS reflectivity is possible, as some open water and flood patterns were clearly identified in the K-means/ DTW clustering results. We tried to list the several sources of misclassification to identify contributing factors in CYGNSS reflectivity either at the global or regional scale. We look forward to producing a dynamic mapping of floods using CYGNSS, with ancillary and reference datasets helping to characterize the relation between CYGNSS reflectivity and SWE.

CRediT authorship contribution statement

Pierre Zeiger: Conceptualization, Methodology, Software, Validation, Formal-analysis, Investigation, Visualization, Writing - original draft, Writing - review & editing. **Frédéric Frappart:** Conceptualization, Validation, Resources, Supervision, Funding-acquisition. **José Darrozes:** Conceptualization, Validation, Supervision. **Catherine Prigent:** Validation, Data-curation. **Carlos Jiménez:** Validation, Data-curation.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgements

P. Zeiger is funded by a PhD grant from Ministère de l'Enseignement Supérieur, de la Recherche et de l'Innovation (MESRI). This work was supported by Centre National de la Recherche Scientifique (CNRS) and Centre National d'Etudes Spatiales (CNES) through INSU PNTS grant URGENS and TOSCA grants SWHYM and SCOMAG, respectively.

We thank all the reviewers for their helpful and detailed suggestions, to help us improve the revised version of this manuscript.

References

Acreman, M., Holden, J., 2013. How Wetlands Affect Floods. Wetlands 33, 773–786. https://doi.org/10.1007/s13157-013-0473-2.

Al-Khaldi, M.M., Johnson, J.T., Gleason, S., Chew, C.C., Gerlein-Safdi, C., Shah, R., Zuffada, C., 2021a. Inland water body mapping using cygnss coherence detection. IEEE Trans. Geosci. Remote Sens. 59, 7385–7394. https://doi.org/10.1109/TGRS.2020.3047075.

Al-Khaldi, M.M., Johnson, J.T., Gleason, S., Loria, E., O'Brien, A.J., Yi, Y., 2021b. An algorithm for detecting coherence in cyclone global navigation satellite system mission level-1 delay-doppler maps. IEEE Trans. Geosci. Remote Sens. 59, 4454–4463. https://doi.org/10.1109/TGRS.2020.3009784.

Al-Khaldi, M.M., Johnson, J.T., O'Brien, A.J., Balenzano, A., Mattia, F., 2019. Time-Series Retrieval of Soil Moisture Using CYGNSS. IEEE Trans. Geosci. Remote Sens. 57, 10.

Arthur, D., Vassilvitskii, S., 2007. k-means++: The Advantages of Careful Seeding. In: Proceedings of the Eighteenth Annual ACM-SIAM Symposium on Discrete Algorithms. Philadelphia, USA, pp. 1027–1035.

Bartlett, K.B., Harriss, R.C., 1993. Review and assessment of methane emissions from wetlands. Chemosphere 26, 261–320. https://doi.org/10.1016/0045-6535(93)90427-7.

Bergamaschi, P., Frankenberg, C., Meirink, J.F., Krol, M., Dentener, F., Wagner, T., Platt, U., Kaplan, J.O., Körner, S., Heimann, M., Dlugokencky, E.J., Goede, A., 2007. Satellite chartography of atmospheric methane from SCIAMACHY on board ENVISAT: 2. Evaluation based on inverse model simulations. J. Geophys. Res. 112, D02304. https://doi.org/10.1029/2006JD007268.

Berndt, D.J., Clifford, J., 1994. Using Dynamic Time Warping to Find Patterns in Time Series. In: AAAI-94 Workshop on Knowledge Discovery in Databases, vol. 10, pp. 359–370.

Betbeder, J., Gond, V., Frappart, F., Baghdadi, N.N., Briant, G., Bartholome, E., 2014. Mapping of Central Africa Forested Wetlands Using Remote Sensing. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 7, 531–542. https://doi.org/10.1109/JSTARS.2013.2269733.

Bloom, A.A., Palmer, P.I., Fraser, A., Reay, D.S., Frankenberg, C., 2010. Large-Scale Controls of Methanogenesis Inferred from Methane and Gravity Spaceborne Data. Science 327, 322–325. https://doi.org/10.1126/science.1175176.

Bullock, A., Acreman, M., 2003. The role of wetlands in the hydrological cycle. Hydrol. Earth Syst. Sci. 7, 358–389. https://doi.org/10.5194/hess-7-358-2003.

Calinski, T., Harabasz, J., 1974. A dendrite method for cluster analysis. Commun. Stats. - Theory Methods 3, 1–27. https://doi.org/10.1080/03610927408827101.

Camps, A., Park, H., Pablos, M., Foti, G., Gommenginger, C.P., Liu, P.W., Judge, J., 2016. Sensitivity of GNSS-R Spaceborne Observations to Soil Moisture and Vegetation. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 9, 4730–4742. https://doi.org/10.1109/JSTARS.2016.2588467.

Cardellach, E., Fabra, F., Nogués-Correig, O., Oliveras, S., Ribó, S., Rius, A., 2011. GNSS-R ground-based and airborne campaigns for ocean, land, ice, and snow techniques: application to the GOLD-RTR data sets. Radio Sci. 46, 1–16. https://doi.org/10.1029/2011RS004683.

Carreno-Luengo, H., Luzi, G., Crosetto, M., 2019. Sensitivity of CyGNSS Bistatic Reflectivity and SMAP Microwave Radiometry Brightness Temperature to Geophysical Parameters over Land Surfaces. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 12, 107–122. https://doi.org/10.1109/JSTARS.2018.2856588. Publisher: IEEE.

Carreno-Luengo, H., Luzi, G., Crosetto, M., 2020. Above-ground biomass retrieval over tropical forests: a novel gnss-r approach with cygnss. Remote Sens. 12 https://doi.org/10.3390/rs12091368.

Chapman, B.D., Russo, I.M., Galdi, C., Morris, M., di Bisceglie, M., Zuffada, C., Downs, B., Lavalle, M., Loria, E., O'Brien, A.J., 2022. Comparison of sar and cygnss surface water extent metrics. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 15, 3235–3245. https://doi.org/10.1109/JSTARS.2022.3162764.

Chen, Y., Huang, C., Ticehurst, C., Merrin, L., Thew, P., 2013. An Evaluation of MODIS Daily and 8-day Composite Products for Floodplain and Wetland Inundation Mapping. Wetlands 33, 823–835. https://doi.org/10.1007/s13157-013-0439-4.

Chew, C., Reager, J.T., Small, E., 2018. CYGNSS data map flood inundation during the 2017 Atlantic hurricane season. Sci. Rep. 8, 9336. https://doi.org/10.1038/s41598-018-27673-x.

Chew, C., Shah, R., Zuffada, C., Hajj, G., Masters, D., Mannucci, A.J., 2016. Demonstrating soil moisture remote sensing with observations from the UK TechDemoSat-1 satellite mission. Geophys. Res. Lett. 43, 3317–3324. https://doi.org/10.1002/2016GL068189.

Chew, C., Small, E., 2020. Estimating inundation extent using CYGNSS data: a conceptual modeling study. Remote Sens. Environ. 246, 111869 https://doi.org/10.1016/j.rse.2020.111869.

Chew, C.C., Small, E.E., 2018. Soil Moisture Sensing Using Spaceborne GNSS Reflections: Comparison of CYGNSS Reflectivity to SMAP Soil Moisture. Geophys. Res. Lett. 45, 4049–4057. https://doi.org/10.1029/2018GL077905.

Choudhury, B.J., 1991. Passive microwave remote sensing contribution to hydrological variables. Surv. Geophys. 12, 63–84.

Clarizia, M.P., Pierdicca, N., Costantini, F., Floury, N., 2019. Analysis of CYGNSS Data for Soil Moisture Retrieval. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 12, 9.

Clarizia, M.P., Ruf, C.S., 2016. Wind speed retrieval algorithm for the cyclone global navigation satellite system (cygnss) mission. IEEE Trans. Geosci. Remote Sens. 54, 4419–4432. https://doi.org/10.1109/TGRS.2016.2541343.

Clarizia, M.P., Ruf, C.S., Cipollini, P., Zuffada, C., 2016. First spaceborne observation of sea surface height using GPS-Reflectometry. Geophys. Res. Lett. 43, 767–774. https://doi.org/10.1002/2015GL066624.

Collett, I., Wang, Y., Shah, R., Morton, Y.J., 2022. Phase coherence of gps signal land reflections and its dependence on surface characteristics. IEEE Geosci. Remote Sens. Lett. 19, 1–5. https://doi.org/10.1109/LGRS.2021.3094407.

Crétaux, J.-F., Arsen, A., Calmant, S., Kouraev, A., Vuglinski, V., Bergé-Nguyen, M., Gennero, M.-C., Nino, F., Rio, R.A.D., Cazenave, A., Maisongrande, P., 2011. SOLS: A lake database to monitor in the Near Real Time water level and storage variations from remote sensing data. Adv. Space Res. 47, 1497–1507. https://doi.org/10.1016/j.asr.2011.01.004.

CYGNSS, 2020. CYGNSS level 1 science data record version 3.0. ver. 3.0. PO.DAAC, CA, USA. URL:https://doi.org/10.5067/CYGNS-L1X30 accessed: 2021-10-01.

Davidson, N.C., Fluet-Chouinard, E., Finlayson, C.M., 2018. Global extent and distribution of wetlands: trends and issues. Mar. Freshwater Res. 69, 620. https://doi.org/10.1071/MF17019.

De Roo, R.D., Ulaby, F.T., 1994. Bistatic specular scattering from rough dielectric surfaces. IEEE Trans. Antennas Propag. 42, 220–231.

Defourny, P., Vancutsem, C., Bicheron, P., Brockmann, C., Nino, F., Schouten, L., Leroy, M., 2007. GlobCover: A 300M Global Land Cover Product for 2005 Using ENVISAT MERIS Time Series. In: Proceedings of the ISPRS Commission VII Symposium Remote Sensing: From Pixels to Processes. Enschede, The Netherlands, pp. 4.

Di Gregorio, A., 2016. Land Cover Classification System - Classification concepts Software version 3. Technical Report Food and Agriculture Organisation Roma, Italy.

Di Vittorio, C.A., Georgakakos, A.P., 2018. Land cover classification and wetland inundation mapping using MODIS. Remote Sens. Environ. 204, 1–17. https://doi.org/10.1016/j.rse.2017.11.001.

Egido, A., Paloscia, S., Motte, E., Guerriero, L., Pierdicca, N., Caparrini, M., Santi, E., Fontanelli, G., Floury, N., 2014. Airborne GNSS-R polarimetric measurements for soil moisture and above-ground biomass estimation. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 7, 1522–1532. https://doi.org/10.1109/JSTARS.2014.2322854. Publisher: IEEE.

Entekhabi, D., Njoku, E.G., O'Neill, P.E., Kellogg, K.H., Crow, W.T., Edelstein, W.N., Entin, J.K., Goodman, S.D., Jackson, T.J., Johnson, J., et al., 2010. The Soil Moisture Active Passive (SMAP) Mission. Proc. IEEE 98, 704–716. https://doi.org/10.1109/JPROC.2010.2043918.

Eroglu, O., Kurum, M., Boyd, D., Gurbuz, A.C., 2019. High spatio-temporal resolution cygnss soil moisture estimates using artificial neural networks. Remote Sens. 11 https://doi.org/10.3390/rs11192272.

ESA, 2017. Land cover CCI product user guide version 2. tech. rep. (2017). URL:https://maps.elie.ucl.ac.be/CCI/viewer/download/ESACCILC-Ph2-PUGv2_2.0.pdf.

Fan, Y., Li, H., Miguez-Macho, G., 2013. Global Patterns of Groundwater Table Depth. Science 339, 940–943. https://doi.org/10.1126/science.1229881.

Fatras, C., Frappart, F., Mougin, E., Frison, P.-L., Faye, G., Borderies, P., Jarlan, L., 2015. Spaceborne altimetry and scatterometry backscattering signatures at C- and Ku-bands over West Africa. Remote Sens. Environ. 159, 117–133. https://doi.org/10.1016/j.rse.2014.12.005.

Fluet-Chouinard, E., Lehner, B., Rebelo, L.-M., Papa, F., Hamilton, S.K., 2015. Development of a global inundation map at high spatial resolution from topographic downscaling of coarse-scale remote sensing data. Remote Sens. Environ. 158, 348–361. https://doi.org/10.1016/j.rse.2014.10.015.

Foti, G., Gommenginger, C., Jales, P., Unwin, M., Shaw, A., Robertson, C., Roselló, J., 2015. Spaceborne GNSS reflectometry for ocean winds: first results from the UK TechDemoSat-1 mission. Geophys. Res. Lett. 42, 5435–5441. https://doi.org/10.1002/2015GL064204.

Frappart, F., Biancamaria, S., Normandin, C., Blarel, F., Bourrel, L., Aumont, M., Azemar, P., Vu, P.-L., Le Toan, T., Lubac, B., Darrozes, J., 2018. Influence of recent climatic events on the surface water storage of the Tonle Sap Lake. Sci. Total Environ. 636, 1520–1533. https://doi.org/10.1016/j.scitotenv.2018.04.326.

Frappart, F., Zeiger, P., Betbeder, J., Gond, V., Bellot, R., Baghdadi, N., Blarel, F., Darrozes, J., Bourrel, L., Seyler, F., 2021. Automatic Detection of Inland Water Bodies along Altimetry Tracks for Estimating Surface Water Storage Variations in the Congo Basin. Remote Sens. 13, 22.

Gerlein-Safdi, C., Bloom, A.A., Plant, G., Kort, E.A., Ruf, C.S., 2021. Improving representation of tropical wetland methane emissions with cygnss inundation maps. Global Biogeochem. Cycles 35, e2020GB006890. https://doi.org/10.1029/2020GB006890.

Gerlein-Safdi, C., Ruf, C.S., 2019. A CYGNSS-Based Algorithm for the Detection of Inland Waterbodies. Geophys. Res. Lett. 46, 12065–12072. https://doi.org/10.1029/2019GL085134.

Ghasemigoudarzi, P., Huang, W., De Silva, O., Yan, Q., Power, D.T., 2020. Flash Flood Detection From CYGNSS Data Using the RUSBoost Algorithm. IEEE Access 8, 171864–171881. https://doi.org/10.1109/ACCESS.2020.3025302.

Gleason, S., O'Brien, A., Russel, A., Al-Khaldi, M.M., Johnson, J.T., 2020. Geolocation, calibration and surface resolution of cygnss gnss-r land observations. Remote Sens. 12 https://doi.org/10.3390/rs12081317.

Gleason, S., Ruf, C.S., O'Brien, A.J., McKague, D.S., 2019. The CYGNSS Level 1 Calibration Algorithm and Error Analysis Based on On-Orbit Measurements. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 12, 37–49. https://doi.org/10.1109/JSTARS.2018.2832981.

Hamilton, S.K., Sippel, S.J., Melack, J.M., 2004. Seasonal inundation patterns in two large savanna floodplains of South America: the Llanos de Moxos(Bolivia) and the Llanos del Orinoco(Venezuela and Colombia). Hydrol. Process. 18, 2103–2116. https://doi.org/10.1002/hyp.5559.

Hess, L.L., Melack, J.M., Novo, E.M., Barbosa, C.C., Gastil, M., 2003. Dual-season mapping of wetland inundation and vegetation for the central Amazon basin. Remote Sens. Environ. 87, 404–428. https://doi.org/10.1016/j.rse.2003.04.001.

Huete, A., Liu, H., Batchily, K., van Leeuwen, W., 1997. A comparison of vegetation indices over a global set of TM images for EOS-MODIS. Remote Sens. Environ. 59, 440–451. https://doi.org/10.1016/S0034-4257(96)00112-5.

Hydroweb, 2022. Water height time series available online: URL:http://hydroweb.theia-land.fr (accessed on 07 July 2022).

Jensen, K., McDonald, K., Podest, E., Rodriguez-Alvarez, N., Horna, V., Steiner, N., 2018. Assessing L-Band GNSS-reflectometry and imaging radar for detecting sub-canopy inundation dynamics in a tropicalwetlands complex. Remote Sens. 10 https://doi.org/10.3390/rs10091431.

Jiménez, C., Catherinot, J., Prigent, C., Roger, J., 2010. Relations between geological characteristics and satellite-derived infrared and microwave emissivities over deserts in northern Africa and the Arabian Peninsula. J. Geophys. Res. 115, D20311. https://doi.org/10.1029/2010JD013959.

Junk, W.J., Brown, M., Campbell, I.C., Finlayson, M., Gopal, B., Ramberg, L., Warner, B. G., 2006. The comparative biodiversity of seven globally important wetlands: a synthesis. Aquat. Sci. 68, 400–414. https://doi.org/10.1007/s00027-006-0856-z.

Kellndorfer, J., Pierce, L., Dobson, M., Ulaby, F., 1998. Toward consistent regional-to-global-scale vegetation characterization using orbital sar systems. IEEE Trans. Geosci. Remote Sens. 36, 1396–1411. https://doi.org/10.1109/36.718844.

Kerr, Y., Waldteufel, P., Wigneron, J.-P., Martinuzzi, J., Font, J., Berger, M., 2001. Soil moisture retrieval from space: the Soil Moisture and Ocean Salinity (SMOS) mission. IEEE Trans. Geosci. Remote Sens. 39, 1729–1735. https://doi.org/10.1109/36.942551.

Kim, H., Lakshmi, V., 2018. Use of Cyclone Global Navigation Satellite System (CyGNSS) Observations for Estimation of Soil Moisture. Geophys. Res. Lett. 45, 11.

Konings, A.G., Piles, M., Das, N., Entekhabi, D., 2017. L-band vegetation optical depth and effective scattering albedo estimation from SMAP. Remote Sens. Environ. 198, 460–470. https://doi.org/10.1016/j.rse.2017.06.037.

Kouadio, I.K., Aljunid, S., Kamigaki, T., Hammad, K., Oshitani, H., 2012. Infectious diseases following natural disasters: prevention and control measures. Expert Rev. Anti-infect. Ther. 10, 95–104. https://doi.org/10.1586/eri.11.155.

Kreibich, H., Van Loon, A.F., Schröter, K., Ward, P.J., Mazzoleni, M., Sairam, N., Abeshu, G.W., Agafonova, S., AghaKouchak, A., Aksoy, H., et al., 2022. The challenge of unprecedented floods and droughts in risk management. Nature 1–7. https://doi.org/10.1038/s41586-022-04917-5.

Kuenzer, C., Guo, H., Huth, J., Leinenkugel, P., Li, X., Dech, S., 2013. Flood Mapping and Flood Dynamics of the Mekong Delta: ENVISAT-ASAR-WSM Based Time Series Analyses. Remote Sens. 5, 687–715. https://doi.org/10.3390/rs5020687.

Lehner, B., Döll, P., 2004. Development and validation of a global database of lakes, reservoirs and wetlands. J. Hydrol. 296, 1–22. https://doi.org/10.1016/j.jhydrol.2004.03.028.

Li, W., Cardellach, E., Fabra, F., Rius, A., Ribó, S., Martín-Neira, M., 2017. First spaceborne phase altimetry over sea ice using TechDemoSat-1 GNSS-R signals. Geophys. Res. Lett. 44, 8369–8376. https://doi.org/10.1002/2017GL074513.

Li, W., Cardellach, E., Ribó, S., Oliveras, S., Rius, A., 2022a. Exploration of multi-mission spaceborne gnss-r raw if data sets: Processing, data products and potential applications. Remote Sens. 14 https://doi.org/10.3390/rs14061344.

Li, W., CARDELLACH, E., RIBó, S., RIUS, A., ZHOU, B., 2021. First spaceborne demonstration of beidou-3 signals for gnss reflectometry from cygnss constellation. Chin. J. Aeronaut. 34, 1–10. https://doi.org/10.1016/j.cja.2020.11.016.

Li, X., Wigneron, J.-P., Fan, L., Frappart, F., Yueh, S.H., Colliander, A., Ebtehaj, A., Gao, L., Fernandez-Moran, R., Liu, X., et al., 2022b. A new SMAP soil moisture and vegetation optical depth product (SMAP-IB): algorithm, assessment and inter-comparison. Remote Sens. Environ. 271, 112921 https://doi.org/10.1016/j.rse.2022.112921.

Loria, E., O'Brien, A., Zavorotny, V., Downs, B., Zuffada, C., 2020. Analysis of scattering characteristics from inland bodies of water observed by cygnss. Remote Sens. Environ. 245, 111825 https://doi.org/10.1016/j.rse.2020.111825.

Macqueen, J., 1967. Some Methods for Classification and Analysis of Multivariate Observations. Fifth Berkeley Symposium, pp. 17.

Maltby, E., Acreman, M.C., 2011. Ecosystem services of wetlands: pathfinder for a new paradigm. Hydrol. Sci. J. 56, 1341–1359. https://doi.org/10.1080/02626667.2011.631014.

Maltby, E., Immirzi, P., 1993. Carbon dynamics in peatlands and other wetland soils regional and global perspectives. Chemosphere 27, 999–1023. https://doi.org/10.1016/0045-6535(93)90065-D.

Martin-Neira, M., 1993. A passive reflectometry and interferometry system (PARIS): application to ocean altimetry, vol. 17. ISSN: 03792285 Publication Title: ESA Journal.

Melton, J.R., Wania, R., Hodson, E.L., Poulter, B., Ringeval, B., Spahni, R., Bohn, T., Avis, C.A., Beerling, D.J., Chen, G., et al., 2013. Present state of global wetland extent and wetland methane modelling: conclusions from a model inter-comparison project (WETCHIMP). Biogeosciences 10, 753–788. https://doi.org/10.5194/bg-10-753-2013.

Mitra, S., Wassmann, R., Vlek, P.L.G., 2005. An appraisal of global wetland area and its organic carbon stock. Curr. Sci. 88, 12.

Mitsch, W.J., Bernal, B., Nahlik, A.M., Mander, l., Zhang, L., Anderson, C.J., Jørgensen, S. E., Brix, H., 2013. Wetlands, carbon, and climate change. Landsc. Ecol. 28, 583–597. https://doi.org/10.1007/s10980-012-9758-8.

Müller, M., 2007. Dynamic Time Warping. In: Müller, M. (Ed.), Information Retrieval for Music and Motion. Springer, Berlin Heidelberg, Berlin, Heidelberg, pp. 69–84. https://doi.org/10.1007/978-3-540-74048-3_.

Morris, M., Chew, C., Reager, J.T., Shah, R., Zuffada, C., 2019. A novel approach to monitoring wetland dynamics using CYGNSS: everglades case study. Remote Sens. Environ. 233, 111417 https://doi.org/10.1016/j.rse.2019.111417.

Nisbet, E.G., Dlugokencky, E.J., Bousquet, P., 2014. Methane on the Rise—Again. Science 343, 493–495. https://doi.org/10.1126/science.1247828.

Normandin, C., Frappart, F., Lubac, B., Bélanger, S., Marieu, V., Blarel, F., Robinet, A., Guiastrennec-Faugas, L., 2018. Quantification of surface water volume changes in the Mackenzie Delta using satellite multi-mission data. Hydrol. Earth Syst. Sci. 22, 1543–1561. https://doi.org/10.5194/hess-22-1543-2018.

Page, S.E., Rieley, J.O., Banks, C.J., 2011. Global and regional importance of the tropical peatland carbon pool. Glob. Change Biol. 17, 798–818. https://doi.org/10.1111/j.1365-2486.2010.02279.x.

Parrens, M., Al Bitar, A., Frappart, F., Papa, F., Calmant, S., Crétaux, J.-F., Wigneron, J.-P., Kerr, Y., 2017. Mapping Dynamic Water Fraction under the Tropical Rain Forests of the Amazonian Basin from SMOS Brightness Temperatures. Water 9, 350. https://doi.org/10.3390/w9050350.

Pekel, J.-F., Cottam, A., Gorelick, N., Belward, A.S., 2016. High-resolution mapping of global surface water and its long-term changes. Nature 540, 418–422. https://doi.org/10.1038/nature20584.

Prigent, C., Aires, F., Jimenez, C., Papa, F., Roger, J., 2015. Multiangle Backscattering Observations of Continental Surfaces in Ku-Band (13 GHz) From Satellites: Understanding the Signals, Particularly in Arid Regions. IEEE Trans. Geosci. Remote Sens. 53, 1364–1373. https://doi.org/10.1109/TGRS.2014.2338913.

Prigent, C., Jimenez, C., Bousquet, P., 2020. Satellite-Derived Global Surface Water Extent and Dynamics Over the Last 25 Years (GIEMS-2). J. Geophys. Res. Atmos. 125 https://doi.org/10.1029/2019JD030711.

Prigent, C., Matthews, E., Aires, F., Rossow, W.B., 2001. Remote sensing of global wetland dynamics with multiple satellite data sets. Geophys. Res. Lett. 28, 4631–4634. https://doi.org/10.1029/2001GL013263.

Prigent, C., Papa, F., Aires, F., Rossow, W.B., Matthews, E., 2007. Global inundation dynamics inferred from multiple satellite observations, 1993–2000. J. Geophys. Res. 112, D12107. https://doi.org/10.1029/2006JD007847.

Rajabi, M., Nahavandchi, H., Hoseini, M., 2020. Evaluation of CYGNSS Observations for Flood Detection and Mapping during Sistan and Baluchestan Torrential Rain in 2020. Water 12, 2047. https://doi.org/10.3390/w12072047.

Richards, J.A., Woodgate, P.W., Skidmore, A.K., 1987. An explanation of enhanced radar backscattering from flooded forests. Int. J. Remote Sens. 8, 1093–1100. https://doi.org/10.1080/01431168708954756.

Ringeval, B., de Noblet-Ducoudré, N., Ciais, P., Bousquet, P., Prigent, C., Papa, F., Rossow, W.B., 2010. An attempt to quantify the impact of changes in wetland extent on methane emissions on the seasonal and interannual time scales. Global Biogeochem. Cycles 24. https://doi.org/10.1029/2008GB003354.

Rodriguez-Alvarez, N., Podest, E., Jensen, K., McDonald, K.C., 2019. Classifying Inundation in a Tropical Wetlands Complex with GNSS-R. Remote Sens. 11, 1053. https://doi.org/10.3390/rs11091053.

Rousseeuw, P.J., 1987. Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. J. Comput. Appl. Math. 20, 53–65. https://doi.org/10.1016/0377-0427(87)90125-7.

Ruf, C.S., Atlas, R., Chang, P.S., Clarizia, M.P., Garrison, J.L., Gleason, S., Katzberg, S.J., Jelenak, Z., Johnson, J.T., Majumdar, S.J., O'Brien, A., Posselt, D.J., Ridley, A.J., Rose, R.J., Zavorotny, V.U., 2016. New ocean winds satellite mission to probe hurricanes and tropical convection. Bull. Am. Meteorol. Soc. 97, 385–395. https://doi.org/10.1175/BAMS-D-14-00218.1.

Ruffini, G., Soulat, F., Caparrini, M., Germain, O., Martín-Neira, M., 2004. The Eddy Experiment: Accurate GNSS-R ocean altimetry from low altitude aircraft. Geophys. Res. Lett. 31 https://doi.org/10.1029/2004GL019994.

Russo, I.M., Bisceglie, M.d., Galdi, C., Lavalle, M., Zuffada, C., 2022. Entropy-based coherence metric for land applications of gnss-r. IEEE Trans. Geosci. Remote Sens. 60, 1–13. https://doi.org/10.1109/TGRS.2021.3125858.

Sakamoto, T., Van Nguyen, N., Kotera, A., Ohno, H., Ishitsuka, N., Yokozawa, M., 2007. Detecting temporal changes in the extent of annual flooding within the Cambodia

and the Vietnamese Mekong Delta from MODIS time-series imagery. Remote Sens. Environ. 109, 295–313. https://doi.org/10.1016/j.rse.2007.01.011.
Santoro, M., 2018. GlobBiomass - global datasets of forest biomass. doi:10.1594/PANGAEA.894711.
Santoro, M., Cartus, O., Carvalhais, N., Rozendaal, D.M.A., Avitabile, V., Araza, A., de Bruin, S., Herold, M., Quegan, S., Rodríguez-Veiga, P., Balzter, H., Carreiras, J., Schepaschenko, D., Korets, M., Shimada, M., Itoh, T., Moreno Martínez, A., Cavlovic, J., Cazzolla Gatti, R., da Conceição Bispo, P., Dewnath, N., Labrière, N., Liang, J., Lindsell, J., Mitchard, E.T.A., Morel, A., Pacheco Pascagaza, A.M., Ryan, C. M., Slik, F., Vaglio Laurin, G., Verbeeck, H., Wijaya, A., Willcock, S., 2021. The global forest above-ground biomass pool for 2010 estimated from high-resolution satellite observations. Earth Syst. Sci. Data 13, 3927–3950. https://doi.org/10.5194/essd-13-3927-2021.
Saunois, M., Stavert, A.R., Poulter, B., Bousquet, P., Canadell, J.G., Jackson, R.B., Raymond, P.A., Dlugokencky, E.J., Houweling, S., Patra, P.K., et al., 2020. The global methane budget 2000–2017. Earth Syst. Sci. Data 12, 1561–1623. https://doi.org/10.5194/essd-12-1561-2020.
Senyurek, V., Lei, F., Boyd, D., Kurum, M., Gurbuz, A.C., Moorhead, R., 2020. Machine learning-based CYGNSS soil moisture estimates over ISMN sites in CONUS. Remote Sens. 12, 1–24. https://doi.org/10.3390/rs12071168.
Sippel, S.J., Hamilton, S.K., Melack, J.M., Choudhury, B.J., 1994. Determination of inundation area in the Amazon River floodplain using the SMMR 37 GHz polarization difference. Remote Sens. Environ. 48, 70–76. https://doi.org/10.1016/0034-4257(94)90115-5.
Suk, J.E., Vaughan, E.C., Cook, R.G., Semenza, J.C., 2020. Natural disasters and infectious disease in Europe: a literature review to identify cascading risk pathways. Eur. J. Pub. Health 30, 928–935. https://doi.org/10.1093/eurpub/ckz111.
Tavenard, R., Faouzi, J., Vandewiele, G., Divo, F., Androz, G., Holtz, C., Payne, M., Yurchak, R., Rußwurm, M., Kolar, K., Woods, E., 2020. Tslearn, A Machine Learning Toolkit for Time Series Data. J. Mach. Learn. Res. 21, 1–6.
Tootchi, A., Jost, A., Ducharne, A., 2019. Multi-source global wetland maps combining surface water imagery and groundwater constraints. Earth Syst. Sci. Data 11, 189–220.
Ulaby, F., Moore, R., Fung, A., 1982. Microwave remote sensing: Active and passive. volume 2-radar remote sensing and surface scattering and emission theory.
Unwin, M.J., Pierdicca, N., Cardellach, E., Rautiainen, K., Foti, G., Blunt, P., Guerriero, L., Santi, E., Tossaint, M., 2021. An introduction to the hydrognss gnss reflectometry remote sensing mission. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 14, 6987–6999. https://doi.org/10.1109/JSTARS.2021.3089550.
Wan, W., Liu, B., Zeng, Z., Chen, X., Wu, G., Xu, L., Chen, X., Hong, Y., 2019. Using CYGNSS Data to Monitor China's Flood Inundation during Typhoon and Extreme Precipitation Events in 2017. Remote Sens. 11, 854. https://doi.org/10.3390/rs11070854.
Webb, E.B., Smith, L.M., Vrtiska, M.P., Lagrange, T.G., 2010. Effects of Local and Landscape Variables on Wetland Bird Habitat Use During Migration Through the Rainwater Basin. J. Wildlife Manage. 74, 109–119. https://doi.org/10.2193/2008-577.
Whalen, S., 2005. Biogeochemistry of Methane Exchange between Natural Wetlands and the Atmosphere. Environ. Eng. Sci. 22, 73–94. https://doi.org/10.1089/ees.2005.22.73.
Wigneron, J.-P., Li, X., Frappart, F., Fan, L., Al-Yaari, A., De Lannoy, G., Liu, X., Wang, M., Le Masson, E., Moisy, C., 2021. SMOS-IC data record of soil moisture and L-VOD: historical development, applications and perspectives. Remote Sens. Environ. 20.
Xiao, X., Boles, S., Liu, J., Zhuang, D., Frolking, S., Li, C., Salas, W., Moore, B., 2005. Mapping paddy rice agriculture in southern China using multi-temporal MODIS images. Remote Sens. Environ. 95, 480–492. https://doi.org/10.1016/j.rse.2004.12.009.
Yan, Q., Huang, W., Jin, S., Jia, Y., 2020. Pan-tropical soil moisture mapping based on a three-layer model from CYGNSS GNSS-R data. Remote Sens. Environ. 247, 111944 https://doi.org/10.1016/j.rse.2020.111944.
Zavorotny, V.U., Gleason, S., Cardellach, E., Camps, A., 2014. Tutorial on remote sensing using GNSS bistatic radar of opportunity. IEEE Geosci. Remote Sens. Mag. 2, 8–45. https://doi.org/10.1109/MGRS.2014.2374220. Publisher: IEEE.

4.3 Contribution to this work and perspectives

This study highlighted the great potential of GNSS-R and CYGNSS for studying inundation dynamics in the entire pan-tropical area. All the methods and results presented are the fruit of my work on developing the algorithms, process and analyze CYGNSS data. I still benefited from very constructive interactions with my supervisors and also with other people, in particular Catherine Prigent who has shown a great interest in the potential of CYGNSS for mapping flood dynamics and also in further comparisons with GIEMS-2 dataset. Most of the conclusions and perspectives to this work are already discussed in the article. Here, I will just add some new insights on the evolution of my investigations after this work, to introduce the next (and last) chapter.

The output of our methodology using the K-means / DTW clustering was basically a static map showing, for each 0.1° (∼11 km at the equator) CYGNSS pixel, whether it was subjected to floods. The output clusters were related to the reflectivity maps and time series to show the evolution of the signal in space and time. It highlighted that all around the world, a high reflectivity and clusters C3 and C4 are linked with the presence of surface water (see Figure 6 on the paper above). These results were a step on exploiting the potential of spaceborne GNSS-R for hydrology, but not our destination at all. The hydrological community requires pan-tropical estimations of inundation extent and dynamics, with high spatiotemporal mapping capability. CYGNSS has now the potential for providing these estimations.

The objective of our last study, consecutive to this work, was to estimate the fraction of water (or Water Fraction) contained in each CYGNSS pixel, at every time step of the 0.1° and 7-day spatiotemporal grid. In the article presented in Section 4.2, we computed several statistical parameters derived from CYGNSS reflectivity with a coherent reflection assumption: Γ_{mean}, Γ_{std}, Γ_{median}, $\Gamma_{90\%}$ and their difference $\Gamma_{90\%-50\%}$, among others. $\Gamma_{90\%}$ or $\Gamma_{90\%-50\%}$ were found to be the best for detecting the presence of water, even a small fraction in the pixel (Section 4.2). For the next part, we used Γ_{mean} which is the weighted average of observations from a 1-month moving window. Interestingly, it shows a (rough) linear increase while the Water Fraction increases in the pixel (see Section 5.3.2). The results obtained through this idea are quite promising, and are presented in the next chapter.

Finally, the spatial resolution constraints and its possible improvement need to be assessed. The 0.1° resolution of our estimations seems coarse, but it was selected to ensure a high number of observations per pixel with the bistatic configuration of the acquisition. The resolution of a CYGNSS coherent observation over surface water is higher (Camps (2019) and Section 2.5.2.1), so pixels with a non negligible Water Fraction also contain some observations with high power. Hence, the main river streams and floodplains are visible in the reflectivity maps from Section 4.2, and the Water Fraction maps in Chapter 5. Only small water bodies cannot be monitored using CYGNSS Level 1 DDM, but raw IF tracks with complex information have this capability down to a hundred of meters. Furthermore, the spatial resolution of our estimations could be improved in different manners:

- CYGNSS incoherent average time was set to 1 s before July, 2019 and 0.5 s after. Observations after the 1-year study period we chose are thus twice more sampled.
- The use of a 1-month Gaussian moving average in our grid gives an average number of observations per time step around 20. This is optimal for deriving robust statistical parameters as it was performed (Section 4.2), but could be reduced while improving the spatial resolution.
- Finally, other publications chose to perform moderate-resolution estimations at 1 km or 3 km. This is made either at the expense of the temporal variations (Al-Khaldi et al., 2021a), or using spatial interpolations (Gerlein-Safdi et al., 2021; Chew, 2021). The spatial interpolation techniques fill numerous gaps where no information exist, and were only applied over selected target areas. There is no evidence of their reliability at global scale in a bistatic acquisition configuration with a large lack of information.

Due to these reasons, we estimate that our estimations could be further performed at ∼5 km. For now, we focused on developing the methodology presented in Chapter 5. These ideas are pointed out for future versions, as they require large time and energy-consuming computations to be produced and evaluated.

Chapter 5

Surface Water Extent retrieval using spaceborne GNSS-R

Contents

5.1 Introduction

In this section, I present the methodology developed to estimate the fraction of water contained in every 0.1° CYGNSS pixel. It is a natural extension of the work presented in Chapter 4. The results shown here largely highlight the capability of CYGNSS to estimate these Water Fractions, and also the robustness of the estimations when compared to ancillary datasets.

Some minor adjustments should be implemented soon to properly finish this work (see Section 5.6). These modifications will concern mostly Section 5.5, with visual and statistical comparisons to be performed against some new and independent MODIS-based SWE maps. The ideas for fine-tuning this work to be published soon are presented in the conclusion (Section 5.6). Also, this chapter do not include for now a discussion section. It will be written once all the results are included, before submitting this work. The conclusions and perspectives (both short and long-term prospects) are discussed in Section 5.6 and in the general conclusion and perspectives.

5.2 Data

5.2.1 CYGNSS Level 1 DDM

CYGNSS is a constellation dedicated to study the formation of tropical cyclones, through the retrieval of ocean wind speed (Ruf et al., 2016). It is composed of 8 Low Earth Orbit (LEO) micro-satellites, with a high revisit in the pan-tropical region (coverage at ±38° of latitude). The Delay Doppler Mapping Instrument (DDMI) onboard CYGNSS satellites comprises 4 channels to record simultaneously 4 reflected signals. Each signal was incoherently averaged over a second before July 2019 (1 Hz sampling), and over 0.5 s afterwards (2 Hz sampling). So, the entire constellation records up to 32 (64) observations per second. Over a pixel with 25 km resolution, as for the CYGNSS-derived wind speed product (Clarizia and Ruf, 2016), the mean and median revisit times are estimated to 7 hours and 3 hours, respectively (Ruf et al., 2016).

The spatial resolution of an observation is quite low over the ocean, due to the diffuse scattering induced by surface roughness (waves). However, it is much higher when the coherent component of the reflected signal dominates, mostly over calm water surfaces, near saturated soils (very high SM content), and flat deserts. The spatial resolution of a coherent reflection is approximately the size of the First Fresnel Zone (FFZ), with a large returned power coming from an even smaller surface (Camps, 2019). The FFZ is at least 0.6 km × 0.6 km for an incidence angle $\theta = 0°$, and increases with θ. Furthermore, the spatial resolution is degraded along-track up to 6.6 km for a 1 Hz sampling, due to the motion of CYGNSS satellites during the incoherent average of the received power. A small fraction of water inside the footprint still contributes to most of the reflected power. So, river streams and lakes down to few hundred meters are detectable using CYGNSS Level 1 DDM, according to both models

(Loria et al., 2020) and observations (Al-Khaldi et al., 2021a).

In this study, we used the DDM and metadata from CYGNSS Level 1 Science Data Records, v3.0 (CYGNSS, 2020). The daily files are distributed in the Physical Oceanography Distributed Active Archive Center (PODAAC). As for a previous study (Zeiger et al., 2022), a 1-year period of time was chosen to evaluate the methodology and results, from August 2018 to July 2019. Except for the last month, CYGNSS observations during the study period have a 1 Hz sampling. Twice more observations are available each month for the next years.

5.2.2 GIEMS Surface Water Extent

For comparison with CYGNSS Water Fraction estimations, external surface water extent (SWE) datasets are required. The Global Inundation Extent from Multi-Satellite version 2 (GIEMS-2, Prigent et al. (2020)) and the Surface Water Microwave Product Series (SWAMPS, Schroeder et al. (2015)) are the two global, long-term products of SWE, derived from passive and active microwave observations. They are based on multi-frequency (mostly Ku-band) brightness temperatures from Special Sensor Microwave/Imager (SSM/I) and Special Sensor Microwave Imager / Sounder (SSMIS) at 0.25° spatial resolution. We used GIEMS as the comparison of both datasets highlighted numerous issues with SWAMPS (Pham-Duc et al., 2017), some of which were further corrected in next releases (Jensen and Mcdonald, 2019). Still, seasonal cycles in SWAMPS have low amplitude and were sometimes found to be in opposition of phase with GIEMS, other SWE products and river discharges in the Niger basin. We finally used the new GIEMS global SWE maps for years 2018 and 2019, as part of the GIEMS-2 release (Prigent et al., 2020). They were not employed in the retrieval methodology. The Water Fractions derived from CYGNSS and GIEMS-2 are thus totally independent, and feed a comprehensive comparison of the inundation dynamics in both datasets (Section 5.5).

5.2.3 Regional dynamic flood maps

The main issue with GIEMS is its low spatial resolution (0.25°). CYGNSS was found to be much more sensitive to small river streams and water bodies than GIEMS (Zeiger et al., 2022). We therefore employed high resolution surface water maps in our CYGNSS Water Fraction retrieval methodology (see Section 5.3). No global information exists at high resolution unless optical datasets like Pekel et al. (2016), which are unable to detect water below the canopy. Our approach thus used regional SWE maps representing distinct climate regions.

First, we used MODIS-based flood maps at 500 m spatial resolution. They are derived from the surface reflectances contained in the MOD09A1 product (8-day binned level 3, version 6), acquired at seven spectral bands from visible/ near infrared (NIR) to shortwave infrared (SWIR). The flood maps are derived from a multi-threshold approach as in Frappart et al. (2018); Normandin et al. (2018), based on the values of Enhanced Vegetation Index (EVI, Huete et al. (1997)) and land surface water index (LSWI, Xiao et al. (2005)). It is a simplification of the method proposed by Sakamoto et al. (2007). The MODIS-based flood

maps were generated for 2018-2019 over three complementary test regions: the Inner Niger Delta (IND) with low vegetation/ bare soils, the Lower Mekong Basin (LMB) with seasonal floods around the Tonle Sap and paddy fields, and the Parana/ La Plata basin with flooded savannas and higher vegetation in the north.

As MODIS cannot be used to detect floods below the canopy, we need other reference SWE maps to evaluate flood dynamics in dense tropical forests. Very few high-resolution SWE maps are available in equatorial areas, and they are mainly based on L-band SAR images. In the Amazon basin, a wetland extent and vegetation classification was achieved based on a L-band JERS-1 SAR double-mosaic at 100 m spatial resolution (Hess et al., 2003, 2015). Although the reference year is 1995-1996, it provides wetland masks at both low (October-November 1995) and high (May-June 1996) water stages. We used the results of this wetland classification that are distributed (Hess et al., 2015), as no other precise and recent flood map exists over the Amazon. The delineated flooded areas at high and low water stages are used similarly than MODIS maps, both in the Water Fraction retrieval methodology (Section 5.3) and to compare with CYGNSS SWE (Section 5.5).

5.2.4 GlobBiomass AGB

The vegetation attenuates or even extinct the coherent component of GNSS-R signals (Loria et al., 2020), causing a decrease in CYGNSS reflectivity especially in dense vegetation cover areas (Carreno-Luengo et al., 2020). Previous investigations have shown that water signatures can still be identified using CYGNSS reflectivity in areas where the Above Ground Biomass (AGB) is in the range of 200-300 Mg/ha, in the Congo and Amazon basins (Zeiger et al., 2022). To consider the vegetation attenuation in our Water Fraction retrieval methodology, we thus employed the GlobBiomass AGB map (Santoro, 2018; Santoro et al., 2021). It has a spatial resolution of 3.2" (∼100 m at the equator) and is given for reference year 2010. The AGB was essentially derived from ALOS-PALSAR L-band and ENVISAT Advanced Synthetic Aperture Radar (ASAR) C-band SAR backscatters. The full methodology is described in Santoro et al. (2021). The AGB map presents a high level of detail, reproduces well the spatial patterns of AGB, and shows a bias over regions with very dense vegetation, where the dynamic range of the observations is not well calibrated and ASAR data are not available.

5.2.5 Resampling of the datasets

To match the spatial resolution of our CYGNSS estimates (0.1°, see Section 5.3.1), the ancillary datasets were resampled. The mean and standard deviation of GlobBiomass AGB were computed in every pixel. Also, the regional reference flood maps at 500 m (MODIS, Normandin et al. (2018)) and 100 m (JERS-1, Hess et al. (2015)) were upscaled into the 0.1° grid. The reference Water Fractions were derived as the number of flooded sub-pixels in the reference datasets, divided by the total number of sub-pixels.

5.3 Methods

The preprocessing of CYGNSS dataset was fully described in (Zeiger et al., 2022). The most important steps are the extraction of the peak power in CYGNSS DDM, and the application of a set of flags to mask out observations of low quality or located over the ocean. We also filtered out the observations when the peak power is not correctly centered in the DDM (*i.e.* located in the first or last three delay rows), as it is an indicator of topography effects which may induce low-quality observations.

5.3.1 CYGNSS reflectivity computation

The GNSS reflected signals are composed of a coherent component, associated to specular scattering inside the FFZ, and an incoherent component, associated to diffuse scattering from surrounding areas. The bistatic scattering models have studied both the coherent and the incoherent returned power (De Roo and Ulaby, 1994; Zavorotny and Voronovich, 2000; Voronovich and Zavorotny, 2018). In this study, the surface reflectivity was computed while supposing a coherent reflection model, mostly found over inland water with low roughness (Chew and Small, 2020a; Loria et al., 2020). Let P_{DDM} be the peak of a CYGNSS Level 1 DDM. Considering that P_{DDM} corresponds to the coherent received power, the surface reflectivity Γ was computed as (De Roo and Ulaby, 1994; Gleason et al., 2020):

$$\Gamma = \Gamma(\theta_i = 0°) = \frac{\Gamma(\theta_i)}{cos(\theta_i)} = \left(\frac{4\pi}{\lambda}\right)^2 \frac{P_{DDM}(R_r + R_t)^2}{G_r G_t P_t} \times \frac{1}{cos(\theta_i)} \tag{5.1}$$

where R_r and R_t are the receiver and the transmitter distances to the specular point (*ranges*), G_r is the receiver antenna gain, $P_t G_t$ is the GPS Equivalent Isotropically Radiated Power (EIRP), λ is the GPS L1 signal wavelength ($\lambda_{L1} = 0.1903$ m), and θ_i is the incidence angle of the wave at the specular point. To account for the influence of the incidence angle, the reflectivity was normalized at a 0° incidence using the factor $\frac{1}{cos(\theta_i)}$ (Zeiger et al., 2022).

CYGNSS reflectivity were then gridded at 7-day, 0.1° spatiotemporal resolution. A 30-day moving window was employed in each pixel, to fill gaps induced by the pseudo-random distribution of CYGNSS observations. This windowing was weighted using a Gaussian distribution with $\sigma = 7$ days, such as the central observations (if exist) mostly determined the values of the further parameters. The mean and standard deviation of the observations (Γ_{mean} and Γ_{std}, respectively) along with the nineteenth percentile ($\Gamma_{90\%}$) were computed for each pixel and at every time step. Γ_{mean} was particularly important for the retrieval methodology (see Section 5.3.2).

5.3.2 Linearity between Γ_{mean} and the Water Fraction

The retrieval of the fractions of water (hereafter named *Water Fractions* or WF) takes advantage of the linearity observed between Γ_{mean} and the reference WF. This is highlighted in

Figure 5.1a, where the average Γ_{mean} per 1% WF bin is plotted for different AGB intervals. Two distinct patterns can be observed:

- A quasi-linear relationship over the entire WF range for pixels with moderate to high vegetation cover (AGB $\geq$ 50 Mg/ha). For low AGB pixels (*i.e.* AGB $\leq$ 50 Mg/ha), it is only observed up to a 50% WF in the sample. The greater is the AGB, the higher are the slope and intercept of this linear relation.
- In low AGB intervals, Γ_{mean} remains constant for a Water Fraction greater than 50%, and even decreases in the pixels with the lowest AGB values (AGB $\leq$ 10 Mg/ha). However, very few samples are concerned as highlighted by the density plot in Figure 5.1b. These pixels with a high WF and a low AGB may be exposed to a higher surface roughness caused by wind speed, to explain the lower Γ_{mean} values observed.

Figure 5.1b also highlights the large deviation of Γ_{mean} around its average value for each 1% WF bin. This is the consequence of a global approach. The deviation is lower when plotting the same relationships at regional scale, due to a lower influence of changes in geophysical properties on the GNSS-R signal. We used reference datasets from the Amazon, Niger, Mekong and Parana basins to retrieve pan-tropical Water Fractions with the highest quality at global scale, while regional deviations should be further explored.

5.3.3 Inversion of the Water Fractions

The retrieval methodology is based on the linear relationships between Γ_{mean} and the WF illustrated in Figure 5.1a, although Figure 5.1b highlighted the large deviation around the theoretical fit. To establish a linear model, the choice of the training and validation datasets is crucial. A flowchart summarizing the different steps for estimating CYGNSS WFs is pre-

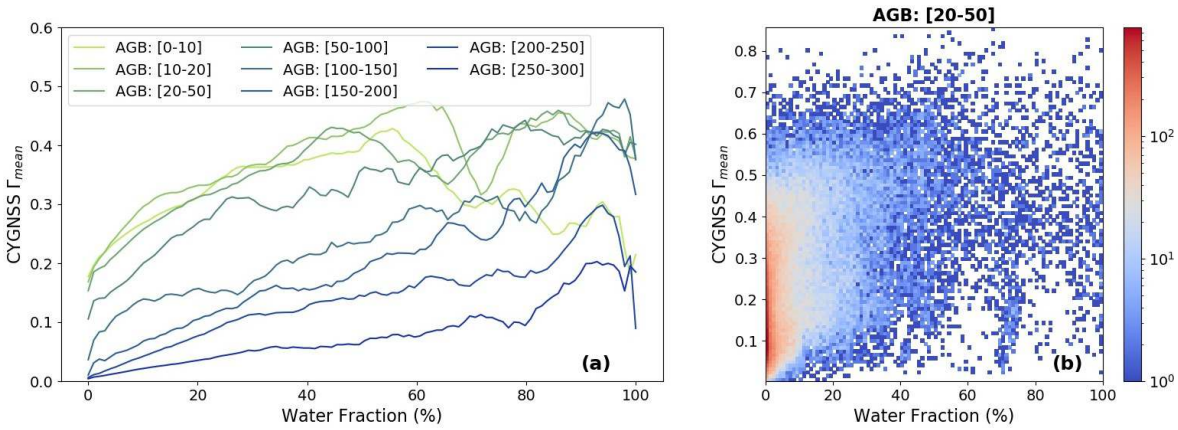

Figure 5.1: Relationships between Γ_{mean} and the Water Fraction. **(a)** Average Γ_{mean} per 1% WF value, for different AGB intervals. The curves are smoothed with a running mean of size 5% to better visualize the linear relations between the two variables. **(b)** Density plot of WF and Γ_{mean}, for the AGB interval [20-50]. The reference WF values are from MODIS and JERS-1 flood maps defined in Section 5.2.

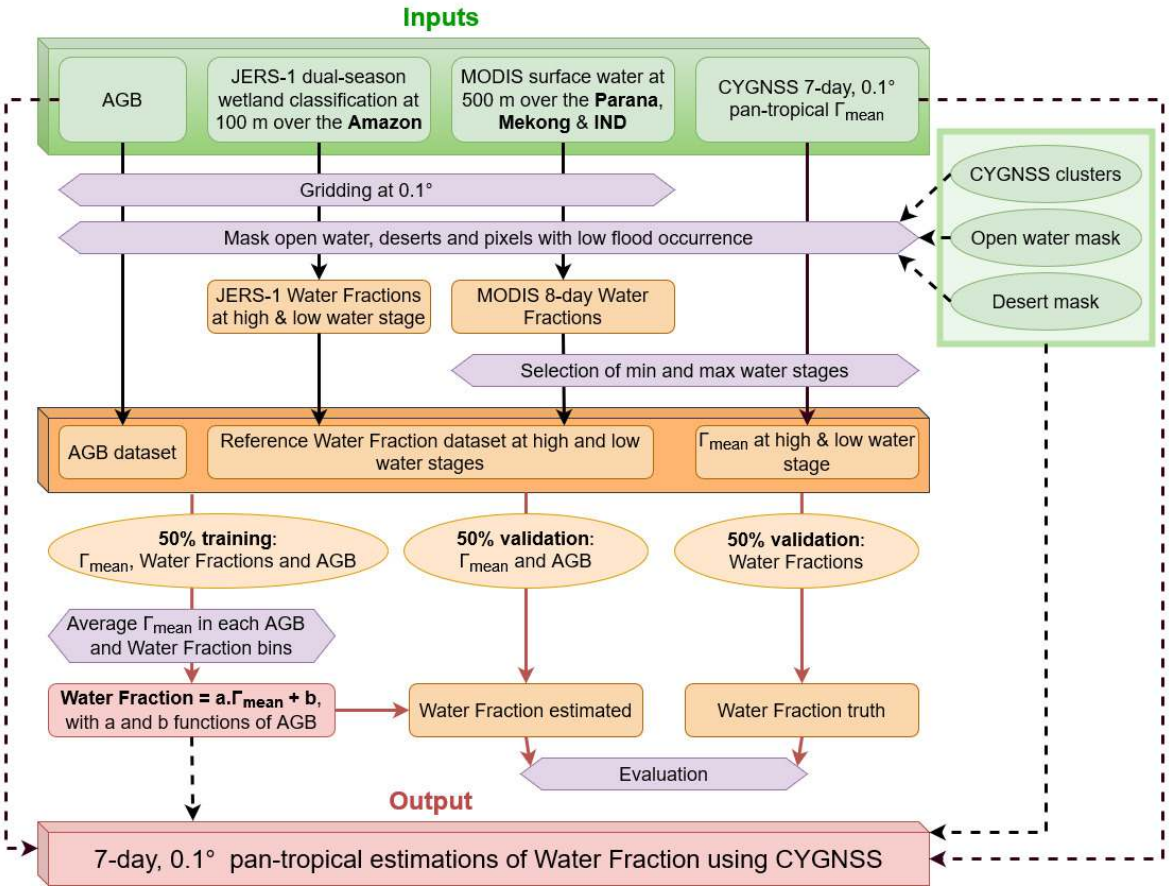

Figure 5.2: Flowchart of the retrieval methodology for estimating CYGNSS Water Fractions at 0.1° spatial and 7 days temporal resolutions.

sented in Figure 5.2. Also, specific points are illustrated in Figure 5.3 to add some graphical explanations.

First, we selected the pixels likely to contain surface water at least a part of the year. For this, we used the clustering results from Zeiger et al. (2022), as pixels in clusters C3 and C4 were shown to correspond to flood or permanent water classes (see Figure 6 in this publication). We also filtered out the deserts prior to the inversion of Water Fractions, as high reflectivity values were found over flat arid regions. The desert mask was determined considering pixels containing more than 90% of bare soils, and where the flood occurrence is null (Figure 5.3a). We used for this the land cover map from European Space Agency's (ESA) Climate Change Initiative (CCI) (ESA, 2017), and the Regularly Flooded Wetland (RFW) dataset from (Tootchi et al., 2019). Furthermore, a low Γ_{mean} was regularly found over large lakes such as Lake Victoria and Tonle Sap (Zeiger et al., 2022), because the roughness (waves) caused by wind attenuates the coherent reflected signal (Chew and Small, 2020b; Loria et al., 2020; Al-Khaldi et al., 2021a). So, pixels covered by at least 80% of open water (from the RFW dataset) were filtered out. As they are already monitored using various remote sensing instruments, there is few interest in including this confounding class into our WF retrieval

algorithm.

The reference Water Fractions come from regional SWE datasets over the Amazon, Parana, Niger and Mekong basins. The dataset from Hess et al. (2015) over the Amazon only provides water extent at high and low water stages during the 1995-1996 hydrological cycle. So, we extracted similar maps from the MODIS 8-day WF estimates over the Parana, LMB and

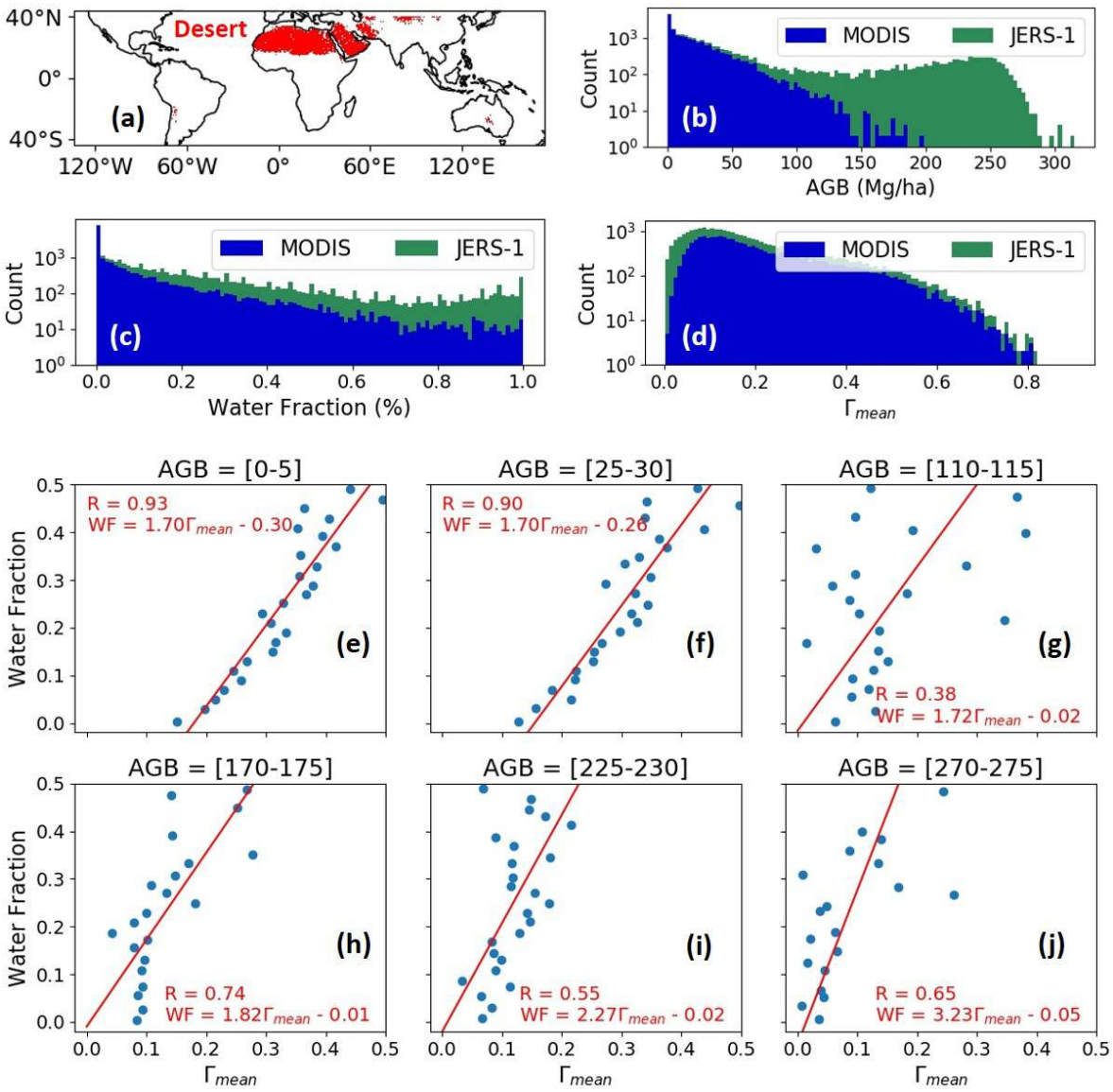

Figure 5.3: Illustration of the Water Fraction retrieval methodology. **(a)** Mask of non-flooded bare soils, **(b)** Distribution of the AGB for samples associated to MODIS or JERS-1 reference WF, **(c-d)** similar histograms for the WF and Γ_{mean}, respectively. **(e-j)** Scatterplots of the average Γ_{mean} computed in 2% WF bins. To remove the influence of the vegetation, a distinct linear relation was calculated in AGB intervals of length 5 Mg/ha. Six examples are shown here for an AGB ranging from 0 to 275 Mg/ha.

IND. The values of Γ_{mean} are similarly extracted for both high and low water stages in every basin (Figure 5.2). As the relation between Γ_{mean} and the WF depends on the vegetation cover, each sample used for training and validation also comprised the AGB derived from GlobBiomass (Santoro, 2018). The distribution of AGB is shown in Figure 5.3b, for samples with a reference WF derived from either MODIS or JERS-1. Although the latter owns less samples (11,878 vs 19,228), it complements the lack of observations from MODIS over pixels with moderate to high biomass. Merging both data sources created a consistent reference WFs dataset for AGB up to 270-280 Mg/ha. Similar histograms for the reference WF and Γ_{mean} samples are shown in Figure 5.3c-d.

For the regression, the input Γ_{mean}, reference WF and AGB (see Figure 5.2, box in the middle) were first separated into 50% training and 50% validation datasets (∼15,500 samples each). The training dataset was binned according to both the AGB and the WF of the samples, using 5 Mg/ha AGB and 2% WF intervals. The average Γ_{mean} in every bin was then computed. The influence of the vegetation on GNSS-R signals was supposed to be constant inside a single AGB interval. Furthermore, the binning along the Water Fraction dimension permits to extract pairs of Γ_{mean} and WF values equally sampled. This was preferred to the use of all samples, as low values dominate the distributions of both variables (Figure 5.3c-d) and would have induced biases in high WF estimations. For every AGB interval, a linear regression was performed between Γ_{mean} and the WF using these pairs of variables, as illustrated in Figure 5.3e-j. These representative plots highlight:

- The linearity between Γ_{mean} and the WF, for both low and high AGB intervals. In a few bins, lower correlations were found as in Figure 5.3g.
- The slopes and intercepts of the linear relations vary depending on the AGB interval. In particular, the slopes are the lower for a low AGB and then increase exponentially from medium to high AGB. Conversely, the intercepts are large (always negative) for a low biomass, and increase toward zero for an increasing AGB.

According to these observations, the CYGNSS Water Fractions (WF_{CYGNSS}) can be computed as:

$$WF_{CYGNSS} = a_{AGB} \times \Gamma_{mean} + b_{AGB} \tag{5.2}$$

where the slope a_{AGB} and the intercept b_{AGB} depend on the average AGB in the pixel. The relationships between a_{AGB}, b_{AGB} and the AGB were evaluated using first, second and third order polynomial fits. We found the latter to better represent the large increase in the slope observed for AGB values above 150 - 200 Mg/ha. So, a_{AGB} and b_{AGB} are computed as:

$$a_{AGB} = a_3 \times AGB^3 + a_2 \times AGB^2 + a_1 \times AGB + a_0 \tag{5.3}$$

$$b_{AGB} = b_3 \times AGB^3 + b_2 \times AGB^2 + b_1 \times AGB + b_0 \tag{5.4}$$

where a_i and $b_i, i \in [0, 3]$ are the coefficients of the third order polynomial regressions between the AGB and a_{AGB} or b_{AGB}, respectively. These coefficients were averaged across the results of 100 random selections of the training and validation datasets, to guarantee the robustness

of the regression (see Section 5.4 and Table 5.1). WF_{CYGNSS} is finally estimated as:

$$\begin{aligned} WF_{CYGNSS} \;=\; & AGB^3\,(a_3\Gamma_{mean} + b_3) + AGB^2\,(a_2\Gamma_{mean} + b_2) \\ & + AGB\,(a_1\Gamma_{mean} + b_1) + a_0\Gamma_{mean} + b_0 \end{aligned} \tag{5.5}$$

5.4 Validation of the methodology

We computed the coefficients a_i and $b_i, i \in [0, 3]$ from a random 50% training dataset. Using Equation (5.5), CYGNSS Water Fractions can then be estimated and compared to the reference WF. Figure 5.4 presents the density plots of WF_{CYGNSS} and $WF_{reference}$, for both the training and the validation datasets. Both show similar performances, and highlight the large deviation between predicted and reference Water Fractions, as expected looking at Figure 5.1b. The Root Mean Square Error (RMSE) and the Pearson's correlation coefficient were computed for the validation dataset, and present consistent values (RMSE = 0.21 and R = 0.63). Figure 5.4 also highlights the large range of WF_{CYGNSS} values obtained for pixels with a very high $WF_{reference}$. The variations of GNSS-R returned power in ∼100% flooded pixels are certainly linked with changes in water roughness (Chew and Small, 2020b).

The robustness of the regression against AGB was assessed by running 100 iterations of the methodology previously described, starting with the random selection of the training and validation datasets. The results are presented in Table 5.1. The correlation coefficient ranged from 0.60 to 0.63, and the RMSE ranged from 21% (0.21) to 23% (0.23), respectively.

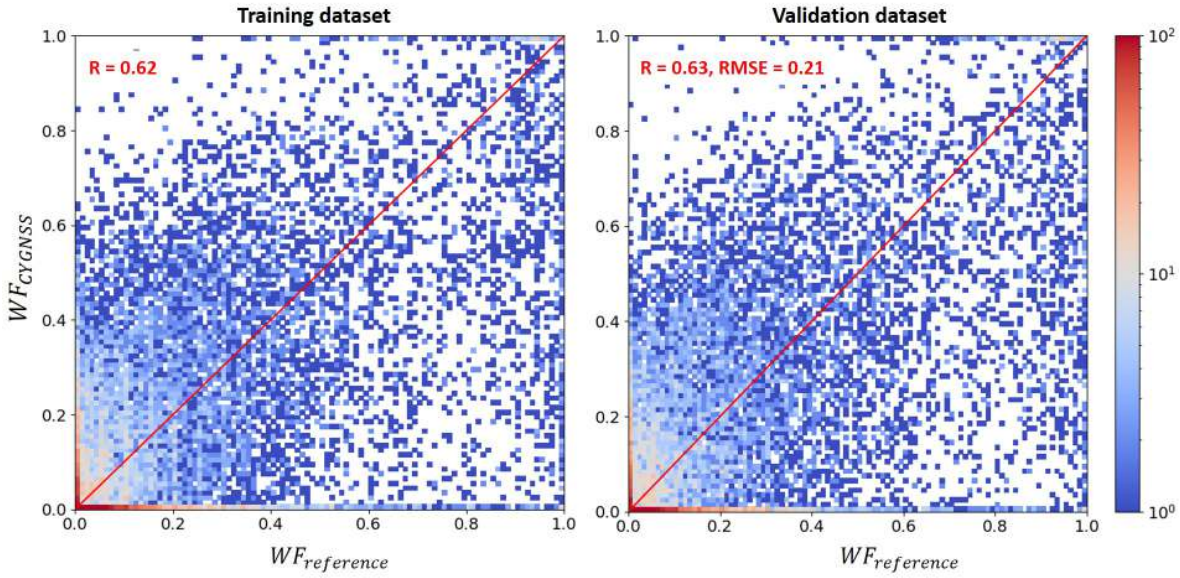

Figure 5.4: Density plots of the estimated CYGNSS Water Fractions (WF_{CYGNSS}) against the reference WF_{target}, for both the training dataset (left) and the validation dataset (right). The correlations are shown in red, with the RMSE which is computed for the validation dataset only.

Table 5.1: Output of 100 iterations on estimating and validating CYGNSS Water Fractions following a random selection of the training and validation datasets. The correlation (R) and the RMSE are indicated, along with the parameters a_i and $b_i, i \in [0,3]$ from Equation (5.5). The minimum, maximum and average values of each parameter are indicated in the first three rows of the table. Finally, the relative standard deviation (rel. std) in the fourth row is the std of the variable divided by its average value, expressed in percentage. It is a good and readable indicator of whether a parameter is stable or not across the 100 iterations.

	R	**RMSE**	$\mathbf{a_0}$	$\mathbf{a_1}$	$\mathbf{a_2}$	$\mathbf{a_3}$	$\mathbf{b_0}$	$\mathbf{b_1}$	$\mathbf{b_2}$	$\mathbf{b_3}$
min	0.60	0.21	1.48	$-1.8e^{-3}$	$-18.8e^{-5}$	$0.7e^{-7}$	-0.43	$4.7e^{-3}$	$-4.2e^{-5}$	$2.6e^{-8}$
max	0.63	0.23	1.90	$14.8e^{-3}$	$0.2e^{-5}$	$6.2e^{-7}$	-0.31	$6.8e^{-3}$	$-2.2e^{-5}$	$8.9e^{-8}$
mean	0.62	0.22	1.70	$5.8\ e^{-3}$	$-8.3e^{-5}$	$3.2e^{-7}$	-0.37	$5.8e^{-3}$	$-3.0e^{-5}$	$4.9e^{-8}$
rel. std	0.8 %	1.1 %	5.4 %	63.0 %	52.1 %	40.4 %	6.9 %	7.9 %	14.9 %	26.1 %

The relative standard deviations (*i.e.* the std divided by the average) are shown for each parameter. It is very low (∼1 %) for both the correlation and the RMSE. Conversely, the relative std of the coefficients $a_i, i \in [0,3]$ (slope of the linear regression, see Equation (5.5)) range from 5.4% to 63.0%. The deviation is lower for the coefficients of the intercept ($b_i, i \in [0,3]$) with relative std ranging from 6.9% to 26.1%.

The average value of a_i and $b_i, i \in [0,3]$ across 100 iterations were found to be robust, and were finally fixed as the best fit parameters for the regression. Note that although Figure 5.4 highlights the deviation of WF_{CYGNSS} compared to $WF_{reference}$, both the RMSE and R are stable across 100 iterations whatever the choice of the training and validation samples. Figure 5.5 shows the evolution of a_{AGB} (slope) and b_{AGB} (intercept) as a function of the AGB. For low vegetation covers, the slope is ∼constant and the intercept increases rapidly while increasing the AGB. But for high vegetation covers, the intercept is near 0 and the slope increases exponentially while increasing the AGB. These relations were used along with global AGB estimates from Santoro (2018) to compute the Water Fractions in the full coverage of CYGNSS.

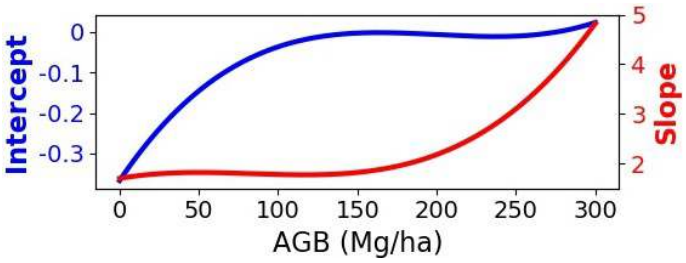

Figure 5.5: Slope and intercept of the linear relation between Γ_{mean} and the Water Fraction, as a function of the AGB (a_{AGB} and b_{AGB} in Equation (5.2), respectively).

5.5 Evaluation of the pan-tropical CYGNSS Water Fraction

Now, weekly Water Fractions were estimated in the full coverage of CYGNSS using all the average coefficients listed in Table 5.1. The results are compared to the regional and global Water Fraction datasets presented in Section 5.2, including GIEMS-2. While ∼31,000 samples from the regional reference SWE datasets were used in the retrieval methodology, GIEMS and CYGNSS Water Fractions are fully independent and their comparison is particularly interesting.

5.5.1 Global comparison with GIEMS

We first compared our results against GIEMS-2 Water Fractions (hereafter WF_{GIEMS}) in the full coverage of CYGNSS (±38° latitude). The WF_{CYGNSS} estimations are gridded at 0.1° and 7 days spatiotemporal resolution, while GIEMS provides its estimations on a 0.25°, monthly basis. The version of GIEMS we used was upsampled at ∼10 days to provide thrice-monthly information, still at 0.25° spatial resolution.

Figure 5.6 presents monthly averaged WF from both CYGNSS and GIEMS. The two top panels show WF_{CYGNSS} and WF_{GIEMS} for August 2018, with large floods in the Orinoco and northern Amazon, and irrigated crops in Southeast Asia. The two middle panels show WF_{CYGNSS} and WF_{GIEMS} for December 2018, with floods especially in the Congo basin. Finally, the two bottom panels show WF_{CYGNSS} and WF_{GIEMS} for April 2019, during high water stage in the southern Amazon and the Parana. The large areas covered by seasonal water are detected using both datasets. The spatial correspondence is good, with WF_{CYGNSS} more disseminated and WF_{GIEMS} more concentrated in the large flooded areas. This is due to the spatial resolution of the sensors used to derive both products: 25 km for SSM/I and SSMIS, against ∼1 × 6 km for a coherent CYGNSS observation. Also, the Water Fractions in coastal regions are larger in CYGNSS than GIEMS. This may be due to a contamination of coastal pixels by GNSS-R observations over the ocean, but also to the presence of mangroves. A further comparison with global high-resolution mangrove maps like the one from Hamilton and Casey (2016) will help in disentangling these two effects.

5.5.2 Regional comparisons

We compared CYGNSS, GIEMS and the regional reference WF at the scale of the river basin to better evaluate their respective performances. First, Figure 5.7 shows a comparison of WF_{CYGNSS} and WF_{GIEMS} with the estimated flood extent from JERS-1 over the Amazon (WF_{JERS}), at both low and high water stages in October-November 1995 and May-June 1996, respectively (Hess et al., 2015). These plots highlight three interesting points:

- Spatially, WF_{CYGNSS} detects water over most of the floodplains and along the streams of the Amazon and its main tributaries. It corresponds well to the flooded areas detected by JERS-1 two decades before.

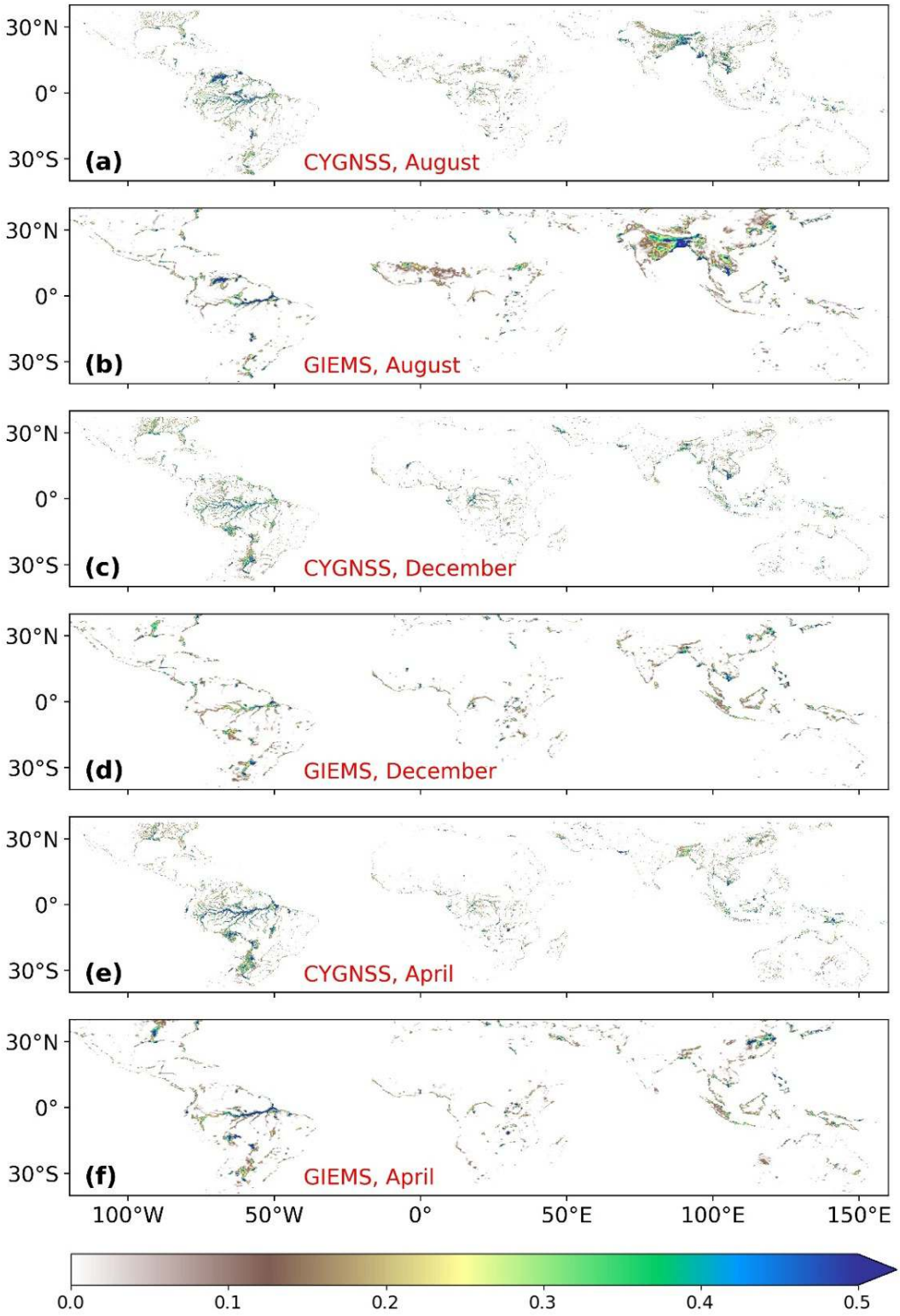

Figure 5.6: Monthly averaged Water Fractions estimates from CYGNSS and GIEMS: **(a)** CYGNSS in August 2018, **(b)** GIEMS in August 2018, **(c)** CYGNSS in December 2018, **(d)** GIEMS in December 2018, **(e)** CYGNSS in April 2019, and **(f)** GIEMS in April 2019.

- The amplitude of the signal at high water stage is lower in WF_{CYGNSS} than in WF_{JERS}. In particular, JERS-1 detects areas with 100% floods in the Llanos de Mojos in Bolivia (South of the basin, around 65°W and 15°S), while WF_{CYGNSS} estimations are mostly lower, around 50%. Some pixels along the main stream of the Amazon also show a lower Water Fraction with CYGNSS.
- CYGNSS and GIEMS agree for the amplitude and the seasonality of the signal (see also Figure 5.9b). However, WF_{CYGNSS} shows a much higher level of detail than WF_{GIEMS}, which suffers from the low spatial resolution of SSM/I and SSMIS.

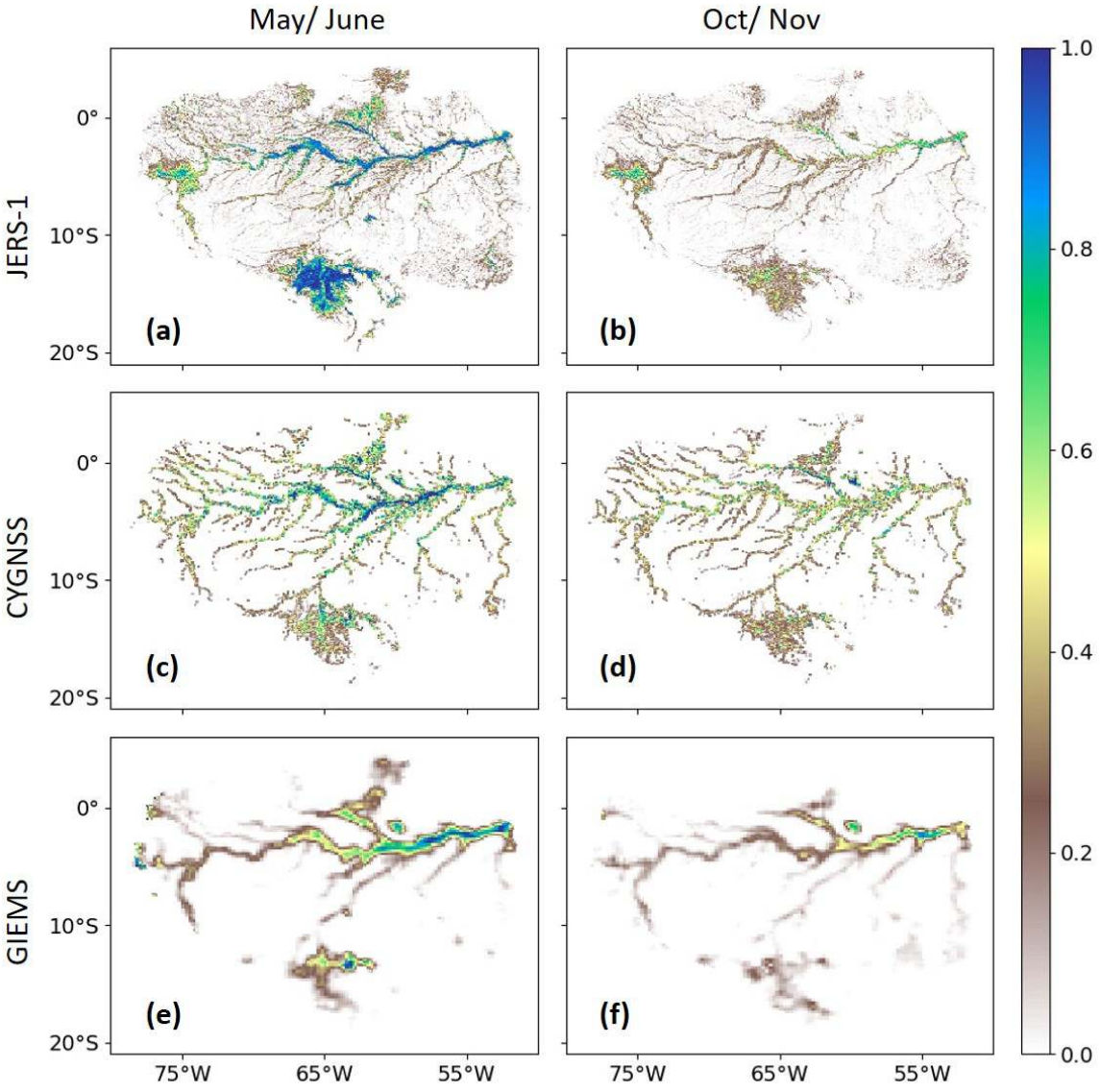

Figure 5.7: Maximum Water Fraction at high water stage (May/ June) and low water stage (October/ November) in the Amazon basin, from JERS-1 (in 1995-1996), CYGNSS and GIEMS (both in 2018-2019). **(a)** JERS-1 high water, **(b)** JERS-1 low water, **(c)** CYGNSS high water, **(d)** CYGNSS low water, **(e)** GIEMS high water, and **(f)** GIEMS low water.

These visual inspections were also conducted over the Parana, Niger and Mekong basins against MODIS WF estimations (WF_{MODIS}), at the time of maximum flood extent (Figure 5.8). WF_{CYGNSS} and WF_{MODIS} look similar spatially, especially over the IND and the LMB (including the Tonle Sap and the Delta of Mekong). GIEMS is coarser but also detects the major floodplains, although the large spread in WF_{GIEMS} may be related to an overestimation of flooded areas during the Sahelian and Southeast Asian monsoons. Finally, over the Parana / La Plata basin, WF_{GIEMS} and WF_{CYGNSS} are very close spatially and in amplitude (see also Figure 5.9c). On the contrary, WF_{MODIS} in the Parana shows fewer flooded pixels with a lower average Water Fraction.

To evaluate the capability of each dataset to retrieve the dynamics of surface water, we computed and compared the average SWE per time step at the scale of the river basin.

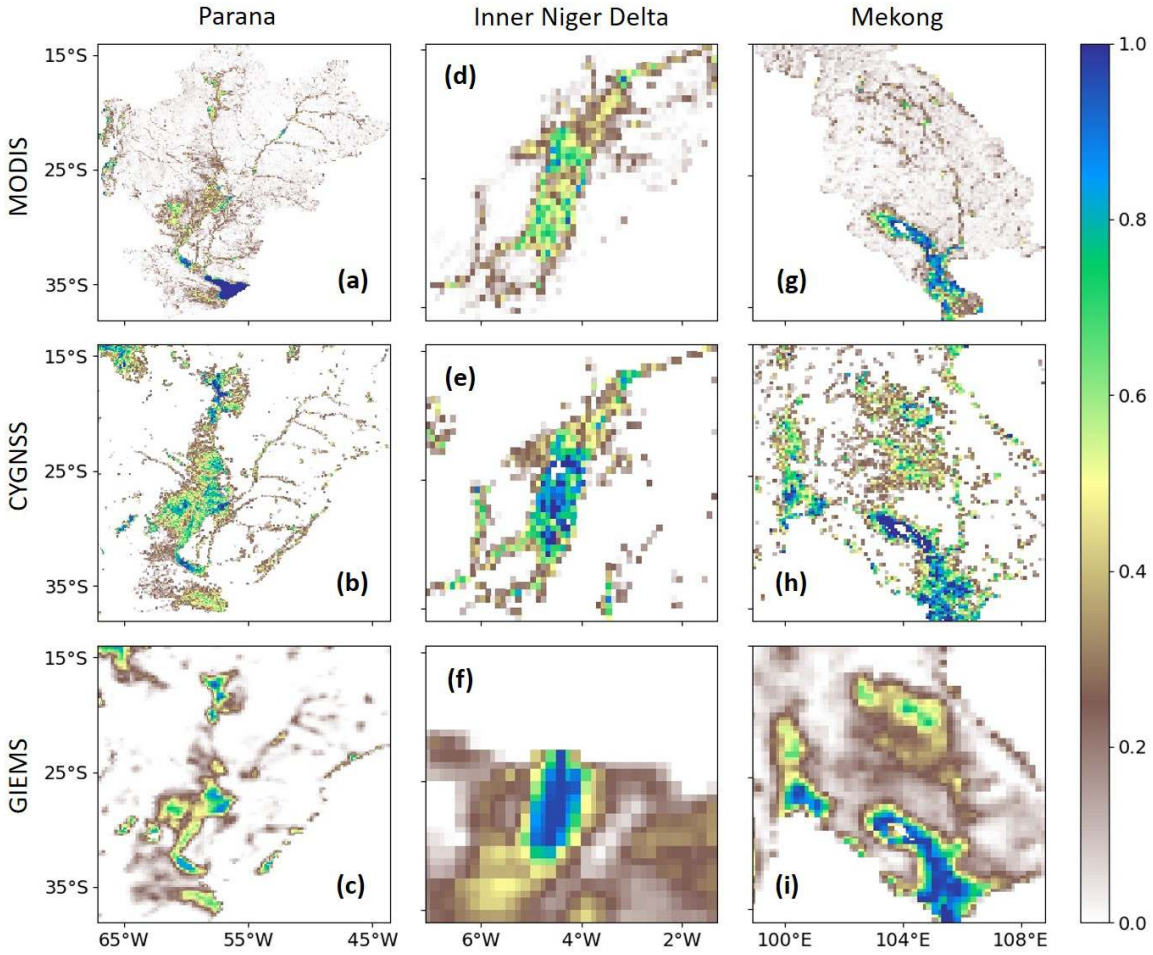

Figure 5.8: Water Fraction at high water stage in the Parana, Inner Niger Delta (IND) and Lower Mekong Basin (LMB), derived from MODIS, CYGNSS and GIEMS in 2018-2019. **(a)** MODIS in the Parana, **(b)** CYGNSS in the Parana, **(c)** GIEMS in the Parana, **(d)** MODIS in the IND, **(e)** CYGNSS in the IND, **(f)** GIEMS in the IND, **(g)** MODIS in the LMB, **(h)** CYGNSS in the LMB, and **(i)** GIEMS in the LMB.

The WF were simply multiplied by the size of each pixel (around 100 km^2, depending on the latitude), and the results in the boundaries of each watershed were added up. Figure 5.9 shows the results in the Orinoco, Amazon and Parana/ La Plata basins in South America, the IND, Chad and Congo basins in Africa, and the Ganges-Brahmaputra, LMB and Yangtze basins in Asia. When comparing CYGNSS (SWE_{CYGNSS}) to GIEMS (SWE_{GIEMS}), their estimates are consistent and show an identical seasonality. Moreover, the amplitude of the signal is near the same in the Orinoco, Amazon, Parana, Chad and Congo basins (5 out of 9). On the contrary, the peak SWE is higher in GIEMS in the IND, Ganges-Brahmaputra, Mekong and Yangtze basins. In these regions, intense seasonal rainfall occur due to the monsoon regimes. GIEMS likely overestimated the presence of surface water, due to a confusion with soil moisture.

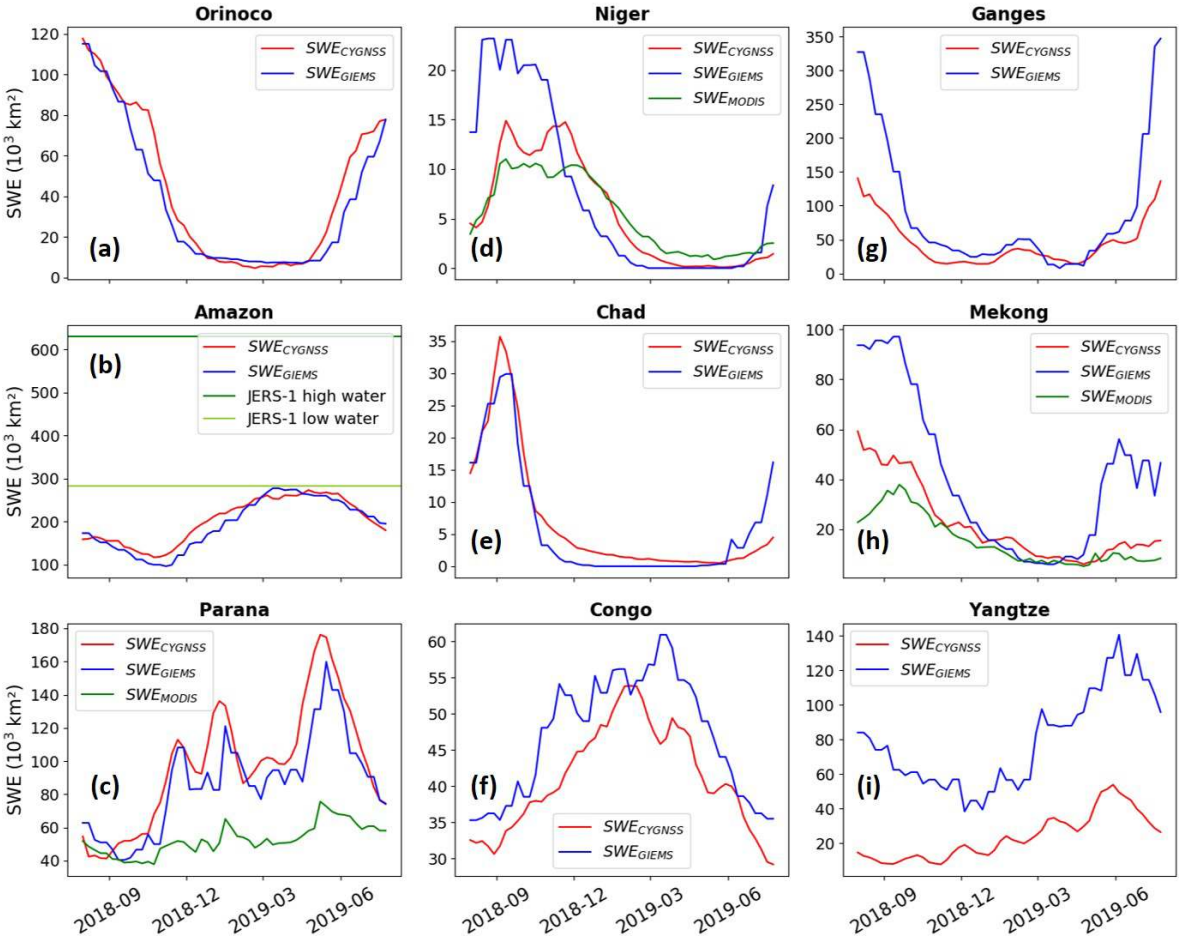

Figure 5.9: Time series of Surface Water Extent (SWE) estimated from CYGNSS and GIEMS averaged for 9 River basins: **(a)** Orinoco, **(b)** Amazon, **(c)** Parana/ La Plata, **(d)** IND, **(e)** Chad **(f)**, Congo, **(g)** Ganges-Brahmaputra, **(h)** LMB, and **(i)** Yangtze. Also shown, the MODIS SWE time series for the Parana, IND and LMB, and the SWE at high and low water stages in the Amazon from Hess et al. (2015).

The comparison with the regional datasets highlights contrasting tendencies. Over the IND and the LMB, both the seasonality and the amplitude of SWE_{CYGNSS} match well the MODIS-based estimations (SWE_{MODIS}). For these two regions, the similar spatial (see Figure 5.8) and temporal evolutions of SWE estimated with MODIS and CYGNSS puts high confidence in the estimations from the latter, while GIEMS likely overestimates seasonal floods. On the contrary, a very high agreement between CYGNSS and GIEMS was found over the Parana, where MODIS estimations are quite low. This can be due to: i) an overestimation of surface water in both CYGNSS and GIEMS due to the influence of SM, or ii) an underestimation of SWE_{MODIS} due to a higher vegetation cover in the Parana basin than over the LMB and the IND. Small seasonal variations and a low amplitude of the peak SWE are found in the time series of SWE_{MODIS} in the Parana. Moreover, the Water Fractions at maximum flood extent are relatively low in MODIS over both the flooded pampas (around 58°W, 38°S), and the Pantanal wetlands (around 58°W, 18°S), as illustrated in Figure 5.8.

Finally, SWE_{CYGNSS} and SWE_{GIEMS} in the Amazon are twice lower than the SWE estimated from JERS-1 (SWE_{JERS}), at both high and low water stages. This could indicate a limitation of both products in the densely forested regions. The vegetation can be the cause of either a loss of the GNSS-R coherent component or its attenuation (Loria et al., 2020), causing lower received power. Nevertheless, these conclusions should be treated cautiously, as the wetland maps over the Amazon date back to 1995-1996. All the long-term trends, interannual and decadal variability could have caused changes in the water storage in the Amazon basin.

5.6 Conclusion and perspectives on CYGNSS Water Fractions

The methods and results presented in this chapter demonstrated the correct retrieval of Water Fractions in 0.1° pixels using CYGNSS. Although a deviation between the estimated and reference WF was noted, the method is robust and the comparisons with ancillary datasets are consistent for both the spatial and temporal dynamics. These results should be submitted soon, with some minor adjustments:

- Presentation of more detailed statistical performances of the regression.
- Statistical comparison of WF_{CYGNSS} with the ancillary datasets (from GIEMS, MODIS and JERS-1), at several spatial scales.
- Comparisons with MODIS-based SWE maps in other river basins, in particular over the Lake Chad (environment similar to the IND) and Yangtze (environment similar to the LMB) basins. These estimations will be fully independent from WF_{CYGNSS}.
- Use of the NDVI (dynamic estimations, unlike AGB maps) to study whether the vegetation explains the deviation between WF_{CYGNSS} and WF_{MODIS}.
- Eventual comparison of the coastal WF_{CYGNSS} with global mangrove inventories.

Beyond this imminent publication, some trails now emerge for further studies. With the conclusive proof of concept of CYGNSS-based Water Fraction estimations, the main developments should be: i) mitigating the observed biases and limitations, and ii) extending this

dataset for the use of the hydrological community. Here are the main pathways to be discussed:

- Does the deviation between the estimated and reference Water Fraction come from the changing geophysical properties of the reflecting surface? Beyond the use of AGB to correct the vegetation attenuation, the role of surface roughness and SM and so the performances of the algorithm when considering their influence should be explored.
- Can CYGNSS WF estimations be performed at higher spatial resolution, while not affecting the robustness of the methodology?
- Does other space GNSS-R missions provide complementary data to improve the spatiotemporal resolution or the coverage of the estimated WF?
- Extending WF_{CYGNSS} estimations to the full ∼6-year record of CYGNSS.

These trails are further discussed in the general conclusion and perspectives.

Conclusion and perspectives

Conclusion

My thesis work focused on GNSS reflectometry techniques (GNSS-R) applied to hydrology. Two different applications were investigated, corresponding to two different spatial scales. First, I worked on water level retrieval in rivers using data acquired by a geodetic receiver. It is a very localized application, as the receiver records GNSS reflected signals from an area of a few tens to hundreds of meters around the antenna. Then, I worked on the detection of floods using spaceborne GNSS-R data from the NASA Cyclone CNSS (CYGNSS) mission (Ruf et al., 2016). This is oriented towards large spatial scales due to the relatively low resolution of coherent observations over smooth water ($\sim$1 $\times$ 6 km, Camps (2019)). Of particular interest is the potential of CYGNSS to map inundations in the densely forested tropical Amazon and Congo basins, because the GNSS L-band signals penetrate better dense forests than other microwave frequencies.

The first application on river water levels was based on a GNSS-R acquisition on the Garonne River in Podensac, using a geodetic receiver recording the signal-to-noise ratio (SNR). The frequency of the SNR oscillations is used to calculate the water heights. The initial objective was to study a very high-frequency phenomena, *i.e.* the tidal bore. Early results faced inherent limitations of the technique but the wave frequency was extracted from the SNR time series (Chapter 3). Finally, we studied the water level variations of the Garonne River during an asymmetric tidal cycle. The amplitude of these very powerful tides in Podensac (at $\sim$120 km from the estuary) sometimes exceeds six meters and they present a large asymmetry, as the rising tide lasts $\sim$3 hours and the falling tide $\sim$9 hours. The water surface varies rapidly (vertical velocity or $\dot{h}$ $\sim 10^{-3}$ m/s). This is the main limitation of SNR techniques developed for coastal applications, and applied to low-amplitude tides (Larson et al., 2013a; Geremia-Nievinski et al., 2020). The dynamic SNR method of Roussel et al. (2015) allows to measure large water level variations. However, it has shown limitations in our case study for two reasons:

- Fewer GNSS satellites in visibility over the river as compared to the ocean. We had a 180° azimuth mask and elevation angles below 10-15° were masked by either the banks or trees. Low elevation tracks are generally used because the SNR oscillations are larger, and thus less subject to noise (Larson et al., 2013a).
- We found multiple frequencies in the SNR oscillations, likely due to the signal being affected by the riverbanks and the vegetation.

I finally modified the dynamic SNR method to adapt it to these noisy measurements. The noise filtering is performed in two steps. The first step consists in removing the multipeaks in the frequency spectrum from the Lomb-Scargle Periodogram (LSP). A prediction interval is used in case of multipeaks to decide which frequency should be extracted. Then, the addition

of an iterative method in the least squares estimation eliminates low-quality observations to compute the final values of h and $\dot{h}$.

The results demonstrate the determination of water levels in the Garonne River, for vertical velocities up to 0.1 - 1 mm/s. In particular, we highlighted that the number of GNSS satellites visible from the receiver mainly explains the variability of the output water heights (h), which presents good results with at least 4 satellites in visibility (R = 0.99, ubRMSD $\leq$ 32 cm), and an increased accuracy for 7 or more satellites (ubRMSD $\leq$ 12 cm). The use of multi-frequency (L1 and L2) and multi-constellation (GPS and GLONASS) data was also successfully tested. This opens the way to land applications such as monitoring flash floods, with geodetic receivers. The acquisition must be carefully configured to minimize azimuth and elevation masks. The results of this study were published in *Remote Sensing* and are presented in Chapter 3.

During the second part of my PhD, I analyzed CYGNSS data to study the spatiotemporal dynamics of inundations. CYGNSS covers the intertropical band (±38° latitude) and provides a large number of GNSS-R observations. Despite its potential for global or large-scale applications, CYGNSS was mainly used at the regional scale. The objective of my work is two-fold: first, to map floods over the entire CYGNSS spatial coverage, and second, to study their temporal dynamics. To do so, I aggregated CYGNSS observations with a spatial resolution of 0.1° and a temporal resolution of 7 days, in order to extract statistical parameters describing the distribution of the coherent reflectivity in each pixel. This served all subsequent work. The study was decomposed in two parts, with an article already published in *Remote Sensing of Environment*.

This publication highlights the capability of CYGNSS to detect the presence of floods over its entire coverage. I used an unsupervised classification method (K-means) applied dynamically to the time series of CYGNSS statistical parameters, thanks to the implementation of the Dynamic Time Warping (DTW) distance metric. The same output cluster represented pixels showing very different flood dynamics, even in opposition of phase between wetlands in the Northern and Southern hemispheres. The result is a static map at 0.1° spatial resolution, whose 4 classes are associated with the presence or absence of water. An important focus was also made on studying how CYGNSS parameters and the output clusters depend on the geophysical properties of the reflecting surface. Several sources of error were identified in accordance with the state of the art: the presence of wind and waves on large lakes (Chew and Small, 2020b), specular reflection on deserts (Al-Khaldi et al., 2021a), or attenuation of the GNSS signal by vegetation (Loria et al., 2020; Carreno-Luengo et al., 2020). Still, surface water was detected in pixels with an average Above ground Biomass (AGB) of up to 300 Mg/ha, corresponding to dense tropical forests in the Amazon and Congo basins.

The second part of my CYGNSS study follows from a direct perspective of the first. We wanted to estimate for each time step (every 7 days) the flood extent or the fraction of water in CYGNSS pixels. To do so, I found a linearity between the fraction of water and the average reflectivity value in each pixel (Γ_{mean}). Of course, a large deviation of the estimates was observed due to the influence of other geophysical parameters. In particular, a dense vegetation strongly attenuates the GNSS signals. To correct for this effect, the slope and

intercept of the linear relationship between Γ_{mean} and the Water Fraction were determined as a function of AGB, using a third order polynomial fit. The reference water fractions used for both learning and validation were taken from several regional maps of surface water extent with higher spatial resolution.

The regional comparisons between CYGNSS estimated Water Fractions, reference SWE maps and GIEMS are very interesting. CYGNSS provides in most cases a clear representation of the maximum flood extent and of its temporal evolution. The comparison between CYGNSS and GIEMS Water Fractions highlights a much higher level of detail in CYGNSS and some biases of GIEMS. Still, GIEMS is highly valuable as it provides 30 years of global SWE estimations. CYGNSS can be useful in evaluating the biases of these long time series, and in understanding the spatial structure of large floods at higher spatial resolution. Finally, CYGNSS Water Fractions can be extended over the ∼6 years of observations, which feeds some of the perspectives discussed below.

Perspectives for in-situ water level measurement

Water level in rivers

The use of SNR measured by a geodetic receiver for coastal sea level measurements has been widely documented (Larson et al., 2013a,b; Löfgren and Haas, 2014; Roussel et al., 2015; Geremia-Nievinski et al., 2020; Tabibi et al., 2020). The fact that as little studies were conducted on river case studies is quite surprising. The dynamic SNR method (Roussel et al., 2015) adapted to noisy observations in Podensac allowed to filter out most of the multipeaks (mainly high frequencies). This considerably reduced the noise on the output water levels. However, an accurate measurement requires at least 4 GNSS satellites in visibility (see Figure 6 in Zeiger et al. (2021) or Chapter 3).

The most interesting prospect for this work is the application of SNR-based water level determinations to extreme hydrological events, especially flash floods. A receiver was already installed in Salsigne (Aude, France) to acquire GNSS-R data on the Orbiel river and its tributaries. The water level rose from ∼8 m on this river during the extreme floods of 2018, generating significant pollution, human and material damage. It is a complex study area with large topography, and therefore elevation masks (see Figure 3.3). It will indicate whether the improved dynamic SNR method can be applied to a very small river with topography around, *i.e.* an extremely unfavourable configuration. The wider the river, the simpler the configuration of the acquisition with many reflections on the water surface. Therefore, an application to large rivers is also of particular interest. The main advantages of the SNR technique are the low cost of the instrumentation, and real time measurements possible if the receiver is connected to internet.

Coastal water levels and sea state

The filtering of multipeaks and the use of the wavelets to analyze multiple frequencies in the SNR time series are also of major interest in the coastal domain. First, the developed method could reduce the noise observed on SSH estimations with a rough sea (Roussel et al., 2015). Then, the study of the sea state itself is interesting to provide information on significant wave height (SWH) and wind speed. For this, further analysis will be necessary. A receiver was installed in the port of Bilbao, at the extremity of the pier protecting the inner port from the rough sea. It will gather SNR data to complement the existing stations in Biarritz and Saint-Jean-De-Luz, for conducting coastal GNSS-R altimetry and possibly sea state monitoring in the Bay of Biscay. The idea is to separate the low frequency component of the signal related to tides from the higher frequencies related to sea state. The comparison with reference data will determine whether a relation with SWH can be established.

Perspectives on monitoring flood dynamics with spaceborne GNSS-R

Uncertainties of CYGNSS Water Fractions

At this time, we did not evaluate the uncertainties of CYGNSS Water Fraction estimates presented in Chapter 5. The linear model validation reports a root mean square error (RMSE) of ∼20% between the estimated and reference Water Fractions. However, these references are also subject to biases, as are the biomass dataset used in the retrieval methodology (Santoro et al., 2021). It is thus not obvious how to quantify the error related to both the propagation of these uncertainties, and the GNSS-R signal itself. The current Water Fraction estimations (a «version 0») show a significant deviation around the expected reference values (see Figure 5.1 and Figure 5.4). Many pixels with a theoretical Water Fraction of 0 actually have non-zero estimates. These pixels could be used to quantify the noise floor - or uncertainty -, in the same way as was done for GIEMS for example (Prigent et al., 2020). This noise floor could then be used as an indicator of the confidence to give to water fraction estimates.

Influence of other geophysical parameters on the Water Fractions

A significant part of the deviation between estimated and reference Water Fractions probably comes from the reflection of the GNSS signal by the Earth's surface. As a reminder, it was shown that the coherent signal can be strongly attenuated by high soil and water roughness (Chew and Small, 2020b; Loria et al., 2020) and by vegetation (Carreno-Luengo et al., 2020; Loria et al., 2020). On the contrary, high soil moisture (SM $\geq$ 40 - 50%) may reflect the GNSS signal in the same way as surface water (Chew and Small, 2020b; Collett et al., 2022). The influence of all these variables is considered in the coherent bistatic reflection models (De Roo and Ulaby, 1994; Voronovich and Zavorotny, 2018), which were used to compute CYGNSS reflectivity. However, the methodology presented in Chapter 5 only takes into account the vegetation attenuation. Some issues were related to changes in the roughness of the reflecting

surface, and as a consequence large lakes (with more wind-related roughness) as well as deserts (not flooded but producing specular reflection) were filtered out before the computation of the Water Fractions. These findings were observed in Chapter 4 and confirm previous results (Al-Khaldi et al., 2021a). As for soil moisture, its role was so far neglected in the proposed method. Future versions of the Water Fraction estimates could implement SM products from SMAP (Chan et al., 2016b, 2018) or SMOS (Wigneron et al., 2021) in the processing chain, as well as proxies for soil roughness such as the SMAP *h-parameter*. It will then determine whether the addition of this information improves the linear regression or not.

Time series of CYGNSS Water Fractions since 2017

As of now, only one year of data was used for both the K-means / DTW classification of CYGNSS observations (Chapter 4), and for the estimation of Water Fractions (Chapter 5). This is related to the large volume of data analyzed and stored as well as the time required for their processing. As long as the processing chain is not finalized, we prefer to study only a relatively short period of data (August 2018 to July 2019) presenting a complete hydrological cycle. For later works, CYGNSS data are available from 2017 up to now and give the possibility to analyze a 5-year dataset, or even more thanks to the continuity of the mission. With a 5-year Water Fraction dataset, we could remove the seasonality of the signal in order to obtain its anomalies, for studying the interannual variability of the inundation extent. This would be a valuable source of information for hydrology with a greater level of detail than GIEMS.

Spatial resolution

Some factors make the 0.1° spatial resolution of our estimates amenable to improvement (see the conclusion in Chapter 4):

- Twice as many observations after July 2019 as during our study period.
- High number of observations per pixel (∼20 each time step on average) to compute robust statistical parameters.

It should be possible to estimate CYGNSS Water Fractions with a spatial resolution of 0.05° (∼5.5 km at the equator) for the period after July 2019. This would perhaps cost a larger number of data gaps. Nevertheless, it is interesting to evaluate this CYGNSS potential at higher resolution. Maps at 0.1° already show a much higher level of detail than GIEMS at 0.25°, thanks in particular to the aggregation of coherent GNSS-R observations with a higher resolution (Camps, 2019). With a reduction of the pixel's area by a factor 4, a better representation of the spatial heterogeneity of floods is likely. Finally, it is worth noting the subsequent possibility to downscale the estimations, *i.e.* to improve their spatial resolution using ancillary high resolution datasets. This is typically based on Digital Elevation Models (DEM): within a neighborhood, lower elevation areas have a higher probability of being flooded. These flood probability maps combined with GIEMS allowed for example to compute

the GIEMS-D15 (Fluet-Chouinard et al., 2015) and GIEMS-D3 (Aires et al., 2017) products downscaled at ∼500 m and ∼90 m spatial resolution, respectively.

Surface water volume

As stated in Chapter 1, surface water volumes are commonly derived using radar altimetry time series and GIEMS SWE (Frappart et al., 2008, 2011, 2012; Becker et al., 2018; Tourian et al., 2018). The main limitation is the low spatial resolution of GIEMS. The same methods were applied to SWE maps at higher resolution from optical imagery (Frappart et al., 2018; Pham-Duc et al., 2020), which obviously do not detect floods in forested areas. CYGNSS water fractions will therefore be useful to infer the surface water volumes stored in tropical wetlands. This will initially be focused on seasonal dynamics, and later be extended to interannual variability once longer CYGNSS Water Fraction time series will be computed.

Additionally to this perspective on CYGNSS, we also worked on improving the water level estimates based on radar altimetry. The *virtual stations* (VS) represent the intersections of altimetry tracks with water bodies. The radar altimeters overpass the same tracks with a revisit time ranging from 10 to 35 days, to monitor time series of water levels. Few VS existed over floodplains, and most were manually identified. We developed an approach to automatically detect the water bodies along altimetry tracks based on the radar backscatter (other parameters were considered). The results were published in *Remote Sensing* and are presented in Section 5.6. It largely increase the number of altimetry VS, especially over floodplains. It contributes in improving the spatial sampling of VS that are interpolated to derive a water level map, and further compute the water volumes stored in wetlands.

Hydrological models and methane emissions

The impact of tropical wetlands on hydrological and biogeochemical cycles is very important, as illustrated by the state of the art in Chapter 1. Their hydrological function regulates river discharges (Junk et al., 2013), as their storage capacity attenuates flood peaks and contributes to maintain the flow in low water periods (Fossey et al., 2016). Hydrodynamic models also highlight the importance of representing stream connectivity with floodplains to evaluate flood dynamics (Fleischmann et al., 2018; Neal et al., 2012). Currently, the unavailability of reliable flood estimates in the Congo and Amazon basins, in particular, hinders the understanding of hydrological dynamics in tropical wetlands. Although CYGNSS water fractions do not have a high spatial resolution, they are much more accurate than the estimates currently provided by GIEMS or SWAMPS at 0.25°.

Wetlands also represent the largest natural source of methane emissions. However, the intercomparison of models show large disagreements in the spatial and temporal distribution of methane emissions from wetlands (Melton et al., 2013). A large part of the uncertainty comes from the SWE estimations provided by GIEMS or SWAMPS (Poulter et al., 2017). In particular, the computation of methane emissions from wetlands is based on models mostly

forced by SWAMPS (Poulter et al., 2017; Saunois et al., 2017). For reminder, a comparison showed that SWAMPS is in opposition of phase with other flood products and also with river discharges in the Niger basin (Pham-Duc et al., 2017). Thus, wetland methane emission models are very poorly constrained. This is reflected in the large difference in the methane budget between atmospheric inversions (*top-down*), and inventory-based estimates (*bottom-up*; see Saunois et al. (2020) and Chapter 1). CYGNSS Water Fractions at 0.1° or 0.05° could help in better constraining the inventories in tropical wetlands.

HydroGNSS, SWOT and NISAR

The conclusion of Chapter 1 mentions the future space missions dedicated to hydrology: SWOT (Biancamaria et al., 2016), NISAR and HydroGNSS (Unwin et al., 2021). The hydrological products supplied by these missions will require validation. In particular, the imminent launch of SWOT (end of 2022) will allow a global monitoring of flooded areas and their water levels thanks to the Ka-band swath altimeter, KaRin (Fjørtoft et al., 2013). A comparison of CYGNSS Water Fractions with SWOT estimated inundation extent will be of great interest, especially in the presence of vegetation to evaluate SWOT performance for tropical wetland hydrology. Both the flood extent and surface water volumes will be better monitored, starting in 2023.

Finally, the future ESA spaceborne GNSS-R mission dedicated to hydrology, HydroGNSS, should be mentioned. It will provide averaged DDMs allowing the detection of large water objects, in a similar way than CYGNSS. It will also include multi-constellation and multi-frequency observations (Unwin et al., 2021), as well as a large number of complex measurements obtained before the incoherent averaging of the signal (see Chapter 2). This will not only allow the continuity with current CYGNSS observations, but also the detection of much smaller water bodies (∼100 m) due to the higher spatial resolution of complex observations (Li et al., 2021, 2022). HydroGNSS will also enable much more accurate phase-delay altimetry (Cardellach et al., 2004), including over inland water (Li et al., 2018). In this context, the development of HydroGNSS products can benefit from the experience gained in the processing of CYGNSS data. Their respective performances could even be compared in case of a common acquisition period for both GNSS-R missions.

Conclusion et perspectives

Conclusion

Mon travail de thèse s'est principalement focalisé sur l'utilisation des techniques de réflectométrie GNSS (GNSS-R) appliquées à l'hydrologie. Deux applications différentes, correspondant à deux échelles spatiales distinctes, ont été au cœur de mes recherches. J'ai tout d'abord travaillé sur la problématique des niveaux d'eau dans les rivières à l'aide de données acquises par un récepteur géodésique conventionnel. Cette application est très localisée, car le récepteur enregistre les signaux GNSS réfléchis provenant d'une zone de quelques dizaines à quelques centaines de mètres autour de l'antenne. Dans un second temps, j'ai travaillé sur la détection des inondations par réflectométrie GNSS spatiale, à l'aide des données fournies par la mission Cyclone CNSS (CYGNSS) de la NASA (Ruf et al., 2016). Cette problématique est tournée vers les grandes échelles spatiales du fait de la résolution relativement faible des observations cohérentes sur l'eau ($\sim$1 $\times$ 6 km, Camps (2019)). Il est particulièrement intéressant d'évaluer le potentiel de CYGNSS pour cartographier les inondations dans les grands bassins tropicaux de l'Amazone et du Congo, présentant des forêts très denses, que les signaux GNSS en bande L pénètrent mieux que les autres fréquences radar.

La première application sur le niveau d'eau des rivières s'est basée sur une acquisition GNSS-R sur la Garonne, à Podensac. Le récepteur géodésique fournit des mesure du rapport signal-sur-bruit (*Signal-to-Noise Ratio*, SNR), dont les oscillations permettent de calculer la hauteur d'eau. Comme nous l'avons vu dans le Chapitre 3, l'objectif initial était de travailler sur un évènement ultra-rapide : le mascaret. Les premiers résultats ont montré que cela n'était pas réaliste, mais que la fréquence des vagues du mascaret pouvait être extraite des séries temporelles de SNR. Nous avons donc cherché à étudier les variations de niveaux d'eau de la Garonne lors d'un cycle de marée asymétrique. Ces marées, très puissantes à Podensac ($\sim$120 km de l'embouchure), ont une amplitude dépassant parfois six mètres et une asymétrie marquée : la marée montante dure environ 3 heures, contre 9 heures pour la marée descendante. La surface d'eau peut varier avec une vitesse importante ($\dot{h}$ ou vitesse verticale), de l'ordre de 10^{-3} m/s. Cela constitue la principale limitation des méthodes existantes, qui ont été développées pour le milieu côtier, et appliquées à des marées de faible amplitude (Larson et al., 2013a; Geremia-Nievinski et al., 2020). La méthode SNR dynamique de Roussel et al. (2015) permet de mesurer d'importantes variations du niveau d'eau. Elle a cependant montré ses limites dans notre cas d'étude, pour deux raisons:

- Moins de satellites GNSS visibles sur la rivière par rapport au cas océanique, avec un masque de 180° en azimut, et les angles d'élévation de moins de 10° masqués par les berges. Or, les basses élévations sont généralement utilisées car les oscillations du SNR sont plus importantes, et donc moins sujettes au bruit (Larson et al., 2013a).
- De multiples fréquences d'oscillations dans les séries temporelles de SNR, sans doute du fait des interactions avec les berges et la végétation, ou des perturbations de la surface

du fleuve (*e.g.* passages de bateaux).

J'ai finalement modifié la méthode dynamique initiale pour l'adapter à ces mesures fortement bruitées. Le filtrage du bruit est réalisé en deux étapes. La première consiste à éliminer les multifréquences dans le spectre fréquentiel issu du périodogramme de Lomb-Scargle (LSP). Un intervalle de prédiction est utilisé dans ce cas pour décider quelle fréquence doit être extraite, et quelle fréquence est issue de perturbations du signal et doit être éliminée. Les fréquences sélectionnées dépendent des variations de la surface d'eau et de l'élévation du satellite GNSS. Ensuite, l'ajout d'une méthode itérative lors de l'estimation par les moindres carrés fournit les valeurs finales de h et $\dot{h}$, en éliminant les observations de moins bonne qualité.

Le résultat final montre la possibilité de déterminer les niveaux d'eau de la Garonne, pour des vitesses verticales de l'ordre de 0,1 - 1 mm/s. En particulier, j'ai mis en évidence que le principal facteur influançant la variabilité des résultats est le nombre de satellites GNSS visibles depuis le récepteur. La hauteur d'eau calculée est de bonne qualité à partir de 4 satellites en visibilité (R = 0.99, ubRMSD $\leq$ 32 cm), et la précision de la méthode augmente pour 7 satellites ou plus (ubRMSD $\leq$ 12 cm). L'utilisation de données multifréquences (L1 et L2) et multi-constellations (GPS et GLONASS) a également été testée avec succès. Cela ouvre la voie à des applications continentales comme le suivi de crues rapides et de grande amplitude, grâce à la technique d'acquisition SNR à bas coût. L'acquisition doit être configurée avec soin, afin de limiter au maximum les masques en azimut et en élévation. Cette méthode altimétrique permet d'étudier de nombreux phénomènes géophysiques présentant une vitesse verticale de l'ordre de 1 mm/s ou moins, en milieu continental comme océanique. Les résultats finaux ont été publiés dans la revue *Remote Sensing* et sont présentés dans le Capitre 3.

Dans la seconde partie de ma thèse, j'ai analysé les données CYGNSS pour étudier les dynamiques spatiotemporelles des eaux de surface, et plus particulièrement celle des inondations. Cette mission couvre la bande intertropicale (±38° latitude) et fournit un grand nombre d'observations GNSS-R. Malgré son potentiel pour les applications globales ou à large échelle, CYGNSS a surtout été utilisé à l'échelle régionale. L'objectif de mon étude est double : d'abord cartographier les inondations sur toute la couverture spatiale de CYGNSS, puis en étudier la dynamique temporelle, qui est mal connue dans la plupart des grands bassins versants tropicaux. Pour cela, j'ai agrégé les observations CYGNSS avec une résolution spatiale de 0,1° et une résolution temporelle de 7 jours, afin d'en dériver dans chaque pixel des paramètres statistiques sur la distribution de la réflectivité cohérente. Cela a servi de base à tout travail postérieur. Cette étude a été décomposé en deux parties qui ont été ou sont en cours de valorisation, avec un article déjà publié dans *Remote Sensing of Environment*.

Dans cette première partie de l'étude, j'ai montré que les variables dérivées de CYGNSS permettent de détecter la présence d'eau sur toute sa couverture spatiale. Pour cela, j'ai utilisé une méthode de classification non-supervisée (K-means) appliquée de manière dynamique aux séries temporelles de variables CYGNSS, grâce à l'implémentation du *Dynamic Time Warping* (DTW) comme mesure de distance entre deux classes. Une même classe représente alors des pixels montrant des dynamiques d'inondations très différentes, voire en opposition de phase entre les zones humides de l'hémisphère Nord et celles de l'hémisphère Sud. Le résultat produit est une carte statique à 0,1° de résolution spatiale, présentant 4 classes associées à

la présence ou non d'inondations. Un axe important a aussi consisté à étudier l'influence qu'ont le type de sols et le couvert végétal sur les variables CYGNSS, et sur le résultat de la classification. Plusieurs sources d'erreur ont été identifiées, conformément à l'état de l'art : la présence de vent et de vagues sur les grands lacs (Chew and Small, 2020b), la réflexion spéculaire sur les déserts (Al-Khaldi et al., 2021a), ou encore l'atténuation du signal GNSS par la végétation (Loria et al., 2020; Carreno-Luengo et al., 2020). Les eaux de surface sont toutefois détectées dans des pixels présentant un *Above ground Biomass* (AGB) moyen allant jusqu'à 300 Mg/ha, correspondant aux forêts tropicales denses des bassins de l'Amazone et du Congo.

La seconde partie de mon étude sur CYGNSS découle d'une perspective directe de la première. Afin de ne pas se restreindre à une cartographie statique des zones inondées, nous voulions estimer pour chaque pas de temps (tous les 7 jours) la présence d'inondations. Il s'est également avéré que le pourcentage (ou fraction) d'eau dans chaque pixel pouvait être estimé approximativement grâce à CYGNSS. Pour cela, j'ai tiré parti de la linéarité observée entre la fraction d'eau du pixel, et la valeur moyenne de réflectivité dans chaque pixel (Γ_{mean}), qui est l'une des variables extraites des observations CYGNSS. Bien entendu, cette linéarité n'est pas parfaite, et on observe une large dispersion des estimations du fait de l'influence d'autres variables géophysiques sur le signal. En particulier, la végétation dense atténue fortement le signal GNSS. Pour corriger cet effet, la pente et l'ordonnée à l'origine de la relation linéaire entre Γ_{mean} et la fraction d'eau ont été fixées en fonction de l'AGB. Les fractions d'eau de référence, utilisées à la fois pour l'apprentissage et la validation, proviennent de différentes cartographies régionales dont la résolution spatiale est plus élevée.

Les comparaisons menées à l'échelle du bassin versant entre la fraction d'eau estimée par CYGNSS, les références régionales et GIEMS sont très positives. CYGNSS permet dans la plupart des cas d'obtenir clairement une représentation spatiale de l'étendue des inondations, mais aussi leur évolution temporelle. La comparaison entre CYGNSS et GIEMS montre aussi un niveau de détail largement supérieur avec CYGNSS, et met en lumière certains biais de GIEMS. Ce dernier garde cependant une valeur inestimable, car il fournit 30 ans d'estimations globales de l'étendue des inondations, avec un signal dont la saisonnalité est correcte mais dont l'amplitude est discutable. CYGNSS peut justement être utile pour évaluer les biais de ces longues séries temporelles, et pour améliorer la compréhension de la structure spatiale des grandes inondations grâce à une meilleure résolution spatiale. Enfin, il sera également possible d'étendre les estimations CYGNSS sur plusieurs années d'observations, ce qui alimente les perspectives discutées ci-dessous.

Perspectives pour la mesure de niveaux d'eau in-situ

Niveaux d'eau des rivières

L'utilisation du SNR mesuré par un récepteur géodésique pour des applications altimétriques en milieu côtier a été largement documenté (Larson et al., 2013a,b; Löfgren and Haas, 2014;

Roussel et al., 2015; Geremia-Nievinski et al., 2020; Tabibi et al., 2020). Il est presque étrange qu'aussi peu de travaux existent sur des cas d'étude en rivière. L'adaptation de la méthode SNR dynamique de Roussel et al. (2015) à des données bruitées, pour le site de Podensac, a permis de considérablement filtrer les multipics en fréquence, et de réduire le bruit final sur la détermination de la hauteur d'eau. Une mesure précise requiert cependant suffisamment de satellites GNSS visibles (voir la Figure 6 dans Zeiger et al. (2021) ou le Chapitre 3).

La perspective la plus intéressante pour ce travail est l'application des méthodes de mesure du SNR et de calcul de la hauteur d'eau à des évènements hydrologiques extrêmes, en particulier les crues éclairs. Un récepteur a déjà été installé à Salsigne, dans l'Aude, pour acquérir des données GNSS-R sur l'Orbiel et ses affluents lorsque le niveau d'eau monte. Lors des crues extrêmes de 2018, la montée des eaux avait atteint 8 m sur cette rivière, générant d'importantes pollutions, des dégâts humains et matériels. Ce site est complexe avec une topographie importante, et donc de forts masques en élévation (voir Figure 3.3). L'étude des séries temporelles de SNR à Salsigne permettra d'avoir d'avantage d'informations sur l'applicabilité de la méthode dynamique améliorée à des sites d'étude variés. Il faut également noter que plus la rivière est large, plus la configuration de l'acquisition est simple avec de nombreuses réflexions sur la surface d'eau. Aussi, une application aux grands fleuves revêt un intérêt tout particulier. Les gros avantages de la technique SNR sont le faible coût de l'instrumentation, et la possibilité de réaliser des relevés en temps réels si le récepteur est connecté au réseau.

Niveaux d'eau côtier et état de la mer

Le filtrage des fréquences doubles et l'utilisation des analyses en ondelettes du signal ont également un intérêt majeur en domaine côtier. Tout d'abord, la méthode développée pourrait permettre de réduire le bruit conséquent sur la mesure de hauteur d'eau lorsque l'état de la mer est mauvais (Roussel et al., 2015), avec une forte hauteur significative des vagues (*Significant Wave Height*, SWH). Ensuite, l'étude de l'état de la mer lui-même est une application intéressante, à même de fournir des relevés sur la hauteur des vagues ou encore la vitesse des vents. Pour cela, des analyses ultérieures seront nécessaires. Un récepteur a été installé dans le port de Bilbao, où la mer agitée est séparée du port interne plus calme, car protégé par une grande jetée. Ces données s'ajouteront à des stations disponibles côté français, à Biarritz et Saint-Jean-De-Luz, pour fournir des relevés d'altimétrie GNSS-R côtière et possiblement d'état de la mer dans le Golfe de Gascogne. L'idée est de séparer la composante basse fréquence du signal liée aux marées des plus hautes fréquences liées à l'état de la mer, pour analyser celles-ci. La comparaison avec des données de référence permettra de savoir si une relation avec le SWH peut être établie.

Perspectives sur les dynamiques d'inondations par GNSS-R spatial

Incertitudes sur les fractions d'eau CYGNSS

Pour l'instant, aucune incertitude n'est fournie sur les estimations de fractions d'eau CYGNSS présentées dans le Chapitre 5. La validation du modèle linéaire fait état d'une erreur moyenne quadratique (*Root Mean Squared Error*, RMSE) de ∼20% entre les valeurs estimées et les valeurs de référence. Cependant, ces références sont également sujettes à des biais, de même que les données de biomasse utilisées dans le calcul des fractions d'eau (Santoro et al., 2021). Il n'est ainsi pas évident de quantifier l'erreur liée à la fois à la propagation de ces incertitudes, et au signal GNSS-R lui-même. L'estimation actuelle de fractions d'eau, que l'on pourrait qualifier de «version 0», montre une dispersion importante autour des valeurs de référence attendues (voir les Figure 1 et Figure 5 du Chapitre 5). De ce fait, de nombreux pixels dont la fraction d'eau théorique est de 0 ont en réalité des estimations non nulles. Ces pixels pourraient servir à quantifier le bruit - ou l'incertitude - observé sur les fractions d'eau, de la même manière que cela fut réalisée pour GIEMS par exemple (Prigent et al., 2020). Ce niveau de bruit plancher pourrait alors être utilisé comme indicateur de la confiance à donner à des estimations de fractions d'eau.

Influence d'autres paramètres géophysiques sur les fractions d'eau

Une part non négligeable des incertitudes provient sans doute de la réflexion du signal GNSS par la surface de la Terre, et de son lien avec différents paramètres géophysiques. Pour rappel, il a été montré auparavant que le signal cohérent peut être fortement atténué par une rugosité importante des sols ainsi que de l'eau (Chew and Small, 2020b; Loria et al., 2020), et par la végétation (Carreno-Luengo et al., 2020; Loria et al., 2020). Au contraire, une humidité des sols élevée (SM $\geq$ 40 - 50%) peut réfléchir le signal GNSS de la même manière que les eaux de surface (Chew and Small, 2020b; Collett et al., 2022). L'influence de toutes ces variables se retrouve de plus dans les modèles de réflexion bistatique cohérente (De Roo and Ulaby, 1994; Voronovich and Zavorotny, 2018), que j'ai utilisé pour calculer la réflectivité CYGNSS. Cependant, la méthodologie présentée dans le Chapitre 5 prend uniquement en compte l'atténuation du signal par la végétation, qui est modélisée pour ensuite calculer la fraction d'eau dans chaque pixel. Afin de régler certains problèmes liés à des changements de rugosité de la surface réfléchissante, les grands lacs (présentant plus de rugosité liée au vent) ainsi que les déserts (non inondés mais produisant une réflexion spéculaire) ont été filtrés en amont du calcul des fractions d'eau. Ces constats proviennent d'observations effectuées dans l'article présenté au Chapitre 4, et confirmant d'autres résultats antérieurs (Al-Khaldi et al., 2021a). Quant à l'humidité des sols, son rôle a jusqu'à présent été négligé dans la méthode proposée. De futures versions des estimations de fractions d'eau pourront implémenter dans la chaîne de traitement des produits d'humidité des sols de SMAP (Chan et al., 2016b, 2018) ou SMOS (Wigneron et al., 2021), ainsi que des proxy de la rugosité des sols comme le h-

parameter de SMAP. Il sera alors possible de savoir si l'ajout de ces informations améliore la régression linéaire, comme la prise en compte de l'atténuation par la végétation a pu le faire.

Séries temporelles de fractions d'eau CYGNSS depuis 2017

Pour l'instant, seule une année de données a été utilisée à la fois pour la classification K-means / DTW des observations CYGNSS (Chapitre 4), et pour l'estimation des fractions d'eau (Chapitre 5). Cela est lié au grand volume de données analysées, à l'espace de stockage ainsi qu'au temps de calcul nécessaires pour leur traitement. Tant que la chaîne de traitement n'est pas finalisée, notamment via l'ajout des idées listées précédemment, nous avons préféré nous restreindre à une période de données certes relativement courte (Août 2018 à Juillet 2019), mais suffisante pour étudier la dynamique saisonnière des inondations sur un cycle hydrologique. Cependant, les données CYGNSS disponibles depuis début 2017 donnent la possibilité d'étendre prochainement notre analyse sur 5 ans minimum, voire plus grâce à la continuité de la mission. Cela permettrait de supprimer la saisonnalité du signal afin d'obtenir ses anomalies, pour comparer celles-ci avec des jeux de données existants (notamment GIEMS). L'étude des dynamiques spatiotemporelles et de la variabilité interannuelle des fractions d'eau CYGNSS peut être une source d'information précieuse pour l'hydrologie, avec un niveau de détail supérieur aux étendues d'inondations fournies par GIEMS.

Résolution spatiale des estimations

Comme je l'ai déjà mentionné (voir la conclusion du Chapitre 4), certains facteurs font que la résolution spatiale de 0,1° de nos estimations pourrait être améliorée, tout en gardant un échantillonnage de 7 jours:

- Deux fois plus d'observations après Juillet 2019 que durant notre période d'étude;
- Un nombre élevé d'observations par pixel (∼20 à chaque pas de temps) permettant d'extraire des paramètres statistiques robustes.

Il devrait être possible d'estimer les fractions d'eau CYGNSS avec une résolution spatiale de 0.05° (∼5,5 km à l'équateur) pour la période postérieure à Juillet 2019. Cela se ferait peut être au prix d'un nombre de trous plus importants dans les données. Néanmoins, il apparaît intéressant d'évaluer ce potentiel de CYGNSS à plus haute résolution. Les cartes à 0,1° montrent déjà un niveau de détail bien plus important que GIEMS à 0,25°, grâce notamment à l'agrégation d'observations GNSS-R cohérentes dont la résolution est supérieure (Camps, 2019). La diminution d'un facteur 4 de la surface de chaque pixel permettra sans doute de rendre encore mieux compte de l'hétérogénéité spatiale des dynamiques d'inondation. Enfin, il faut noter la possibilité ultérieure de *downscaler* les estimations, c'est à dire d'améliorer leur résolution spatiale par la combinaison avec d'autres jeux de données haute résolution. Cela est typiquement réalisé en utilisant des modèles numériques de terrain (*Digital Elevation Model*, DEM) : dans un même voisinage, les zones de plus faible altitude ont une probabilité plus élevée d'être inondées. La combinaison de ces cartes de probabilités d'inondations

avec GIEMS a par exemple permis de calculer les produits GIEMS-D15 (Fluet-Chouinard et al., 2015) et GIEMS-D3 (Aires et al., 2017) à ∼500 m and ∼90 m de résolution spatiale, respectivement.

Calcul des volumes d'eau stockés

Comme je l'ai précisé dans le Chapitre 1, le calcul des stocks d'eau est couramment réalisé en utilisant la combinaison de longues séries temporelles d'altimétrie radar avec GIEMS (Frappart et al., 2008, 2011, 2012; Becker et al., 2018; Tourian et al., 2018). La limitation principale vient de la faible résolution spatiale de GIEMS, malgré l'utilisation alternative des courbes hypsographiques basées sur un DEM (Papa et al., 2015; Kitambo et al., 2022b). Les mêmes méthodes ont été appliquées à des étendues d'inondations estimées à plus haute résolution à partir d'imagerie optique (Frappart et al., 2018; Pham-Duc et al., 2020), qui ne peuvent évidemment pas fournir d'information sur la végétation inondée. Pour les grands bassins tropicaux, en présence de végétation ou de cultures inondées, les fractions d'eau CYGNSS seront donc utile pour étudier la dynamique des stocks d'eau dans les plaines d'inondation. Cette étude pourra porter dans un premier temps sur la dynamique saisonnière, avant d'être ultérieurement étendue à la variabilité interannuelle lorsque les séries temporelles de fractions d'eau CYGNSS auront été calculées depuis 2017.

Modèles hydrologiques et émissions de méthane

L'impact des zones humides tropicales sur les cycles hydrologiques et biogéochimiques est très important, comme l'illustre l'état de l'art dans le chapitre 1. Leur rôle hydrologique permet de réguler les débits des rivières (Junk et al., 2013), car leur capacité de stockage atténue les pics de crue, et contribue à maintenir le débit en période de basses eaux (Fossey et al., 2016). Les modèles hydrodynamiques montrent également l'importance de représenter la connectivité des cours d'eau avec les plaines d'inondations pour représenter de manière réaliste les dynamiques de crues (Fleischmann et al., 2018; Neal et al., 2012). Actuellement, l'indisponibilité d'estimations fiables des inondations dans les bassins du Congo et de l'Amazone, notamment, est un frein à la compréhension de la dynamique hydrologique des zones tropicales. Bien que les fractions d'eau CYGNSS n'aient pas une résolution spatiale élevée, elles sont bien plus précises que les estimations fournies actuellement par GIEMS ou SWAMPS à 0,25°.

Les zones humides représentent également la première source naturelle d'émissions de méthane (Saunois et al., 2020). Toutefois, les comparaisons entre modèles montrent de larges désaccords dans la distribution spatiale et temporelle des émissions de méthane des zones humides (Melton et al., 2013). Une large part des incertitudes provient des estimations d'inondations fournies par GIEMS ou SWAMPS (Poulter et al., 2017). En particulier, le calcul des émissions de méthane des zones humides est basée sur des modèles qui sont, pour la plupart, forcés par SWAMPS (Poulter et al., 2017; Saunois et al., 2017). Or, une comparaison sur le bassin du Niger a montré que SWAMPS pouvait être en opposition de phase avec

d'autres produits d'inondation, mais également avec les débits des rivières (Pham-Duc et al., 2017). Les modélisations d'émissions de méthane dans les zones humides sont donc très mal contraintes. Cela se retrouve dans la grande différence existant dans le budget du méthane entre les inversions atmosphériques (*top-down*), et les estimations basées sur des inventaires (*bottom-up*; voir (Saunois et al., 2020) et le Chapitre 1). Les fractions d'eau CYGNSS à 0,1° ou 0,05° pourraient constituer une information intéressante pour mieux contraindre ces inventaires dans les zones humides tropicales.

HydroGNSS, SWOT et NISAR

Dans la conclusion du Chapitre 1, j'ai évoqué les futures missions spatiales à même d'améliorer la connaissance des étendues et des volumes d'eau: SWOT (Biancamaria et al., 2016), NISAR et HydroGNSS (Unwin et al., 2021). Une évaluation des produits hydrologiques de ces missions qui font, pour la première fois, partie des objectifs scientifiques principaux, sera logiquement nécessaire. En particulier, le lancement imminent de SWOT (fin 2022) permettra un suivi global des zones inondées et de leurs niveaux d'eau grâce à l'altimètre à fauchée en bande Ka, KaRin (Fjørtoft et al., 2013). La comparaison des fractions d'eau CYGNSS avec l'extension des inondations SWOT sur les zones humides sera très intéressante, en particulier en présence de végétation pour déterminer les performances de SWOT sur l'hydrologie des zones humides tropicales. Cela permettra de combiner une connaissance plus précise de l'étendue des inondations et des stocks d'eau, à partir de 2023.

Pour finir, il faut mentionner la future mission GNSS-R spatiale de l'ESA dédiée à l'hydrologie, HydroGNSS. Elle fournira de la même manière que CYGNSS des DDM moyennées permettant la détection de grands objets d'eau. Surtout, elle incluera comme innovations des observations multi-constellations et multifréquences (Unwin et al., 2021), ainsi qu'un grand nombre de *raw IF*, soit les mesures complexes obtenues avant la moyenne incohérente du signal (voir le Chapitre 2). Cela permettra non seulement la continuité avec les observations CYGNSS actuelles, mais également de détecter des objets d'eau bien plus petits ($\sim$100 m), grâce à la résolution spatiale plus élevée des observations complexes (Li et al., 2021, 2022). HydroGNSS permettra aussi de réaliser des applications altimétriques basées sur la phase bien plus précises (Cardellach et al., 2004), y compris sur l'hydrologie continentale (Li et al., 2018). Dans ce cadre, le développement de produits HydroGNSS peut bénéficier de l'expérience acquise dans le traitement des données CYGNSS. Une comparaison directe de leurs performances respectives pourrait même être menée, en cas de période d'acquisition commune aux deux missions GNSS-R.

Appendix A: Publication in Remote Sensing on radar altimetry

remote sensing

MDPI

Article

Automatic Detection of Inland Water Bodies along Altimetry Tracks for Estimating Surface Water Storage Variations in the Congo Basin

Frédéric Frappart [1,2,*], Pierre Zeiger [1], Julie Betbeder [3], Valéry Gond [3], Régis Bellot [3,4], Nicolas Baghdadi [5], Fabien Blarel [1], José Darrozes [6], Luc Bourrel [6] and Frédérique Seyler [7]

1 LEGOS, Université de Toulouse, CNES, CNRS, IRD, UPS—14 Avenue Edouard Belin, 31400 Toulouse, France; pierre.zeiger@legos.obs-mip.fr (P.Z.); fabien.blarel@legos.obs-mip.fr (F.B.)
2 INRAE, UMR1391 ISPA, 33140 Villenave d'Ornon, France
3 CIRAD, Forêts et Sociétés, 34398 Montpellier, France; julie.betbeder@cirad.fr (J.B.); valery.gond@cirad.fr (V.G.); regis.bellot@orange.fr (R.B.)
4 IGN, 94160 Saint-Mandé, France
5 INRAE, UMR TETIS, University of Montpellier, 500 rue François Breton, CEDEX 5, 34093 Montpellier, France; nicolas.baghdadi@teledetection.fr
6 GET, Université de Toulouse, CNRS, IRD, UPS—14 Avenue Edouard Belin, 31400 Toulouse, France; jose.darrozest@get.omp.eu (J.D.); luc.bourrel@ird.fr (L.B.)
7 ESPACE-DEV, Université Montpellier, IRD, Université Guyane, Université Réunion, Université Antilles, Université Avignon, 500 rue Jean-François Breton, 34393 Montpellier, France; frederique.seylert@ird.fr
* Correspondence: frederic.frappart@legos.obs-mip.fr

Citation: Frappart, F.; Zeiger, P.; Betbeder, J.; Gond, V.; Bellot, R.; Baghdadi, N.; Blarel, F.; Darrozes, J.; Bourrel, L.; Seyler, F. Automatic Detection of Inland Water Bodies along Altimetry Tracks for Estimating Surface Water Storage Variations in the Congo Basin. *Remote Sens.* **2021**, *13*, 3804. https://doi.org/10.3390/rs13193804

Academic Editor: Luca Brocca

Received: 6 August 2021
Accepted: 14 September 2021
Published: 23 September 2021

Abstract: Surface water storage in floodplains and wetlands is poorly known from regional to global scales, in spite of its importance in the hydrological and the carbon balances, as the wet areas are an important water compartment which delays water transfer, modifies the sediment transport through sedimentation and erosion processes, and are a source for greenhouse gases. Remote sensing is a powerful tool for monitoring temporal variations in both the extent, level, and volume, of water using the synergy between satellite images and radar altimetry. Estimating water levels over flooded area using radar altimetry observation is difficult. In this study, an unsupervised classification approach is applied on the radar altimetry backscattering coefficients to discriminate between flooded and non-flooded areas in the Cuvette Centrale of Congo. Good detection of water (open water, permanent and seasonal inundation) is above 0.9 using radar altimetry backscattering from ENVISAT and Jason-2. Based on these results, the time series of water levels were automatically produced. They exhibit temporal variations in good agreement with the hydrological regime of the Cuvette Centrale. Comparisons against a manually generated time series of water levels from the same missions at the same locations show a very good agreement between the two processes (i.e., RMSE $\leq$ 0.25 m in more than 80%/90% of the cases and R $\geq$ 0.95 in more than 95%/75% of the cases for ENVISAT and Jason-2, respectively). The use of the time series of water levels over rivers and wetlands improves the spatial pattern of the annual amplitude of water storage in the Cuvette Centrale. It also leads to a decrease by a factor of four for the surface water estimates in this area, compared with a case where only time series over rivers are considered.

Keywords: radar altimetry; wetlands; surface water storage; Congo

1. Introduction

Floodplains and wetlands cover at least 12.1 $\times$ 10^6 km^2 (~8%) of the land surfaces of the Earth [1,2]. They play a major role in the water cycle through river flow variability, flood mitigation, groundwater recharge and water quality improvement [2–4]. They were identified as one of the most productive ecosystems as well as a major contributor of biodiversity within a landscape [5–10]. They also have an important role in the global

carbon cycle as 16 to 33% of the soil carbon pool is stored in the wetlands [11,12] and as 20 to 25% of the methane emissions originated from the wetlands [13–16].

In spite of their importance, the water stored in floodplains and wetlands and its temporal variations are poorly known from regional to global scales. Before the launch of NASA/CNES Surface Water and Ocean Topography (SWOT) in 2022, which will provide surface water elevation over inland water bodies at a spatial resolution of 100 m [17], anomalies of surface water storage are currently derived by: (i) estimating water level changes using an interferometric synthetic aperture radar (SAR) (InSAR) [18,19], (ii) filling a digital elevation model (DEM) with surface water extent estimates from remotely sensed observations through a hypsometric curve [20,21], (iii) combining surface extent products derived from satellite images with radar altimetry based water levels [22,23], or (iv) solving the water balance equation combining various remotely sensed observations (i.e., anomalies of terrestrial water storage (TWS) from the Gravity Recovery And Climate Experiment (GRACE), water levels from radar altimetry, rainfall from Global Precipitation Climatology Project (GPCP), shuttle radar terrain model (RTM), digital elevation model (DEM), synthetic aperture radar (SAR) from Japanese Earth Resources -1 (JERS-1) and multi-spectral from MODerate resolution imaging spectroradiometer (MODIS) images) [24,25]. The latter approach needs the creation of a network of altimetry-based stations of water levels or virtual stations (VSs). Currently, these VSs are mostly obtained through a manual process using dedicated softwares such as the Virtual ALtimetric Stations (VALS) software [26], the Multi-mission Altimetry Processing Software (MAPS) [27,28], or the Altimetry Time Series (AlTiS) software [29], except for large rivers and lakes. This processing is time consuming and does not allow a complete coverage of extensive areas such as the wetlands present in the large river basins, especially when considering a large number of satellite missions. Automatic processing to produce RA-based water levels relies on the identification of the cross-section between a permanent waterbody and the RA ground-tracks (e.g., [30,31]). Therefore, these processes are applied to cross-sections located on rivers, lakes and reservoirs.

Yet, several studies already shown that radar altimetry (RA) can be used for the monitoring of land surface properties. Spatio-temporal changes in radar altimetry backscattering coefficients (σ_0) were related to land cover types (ice, arid and semi-arid areas, wetlands, forests, ...) and hydrological changes (floods, soil moisture, ...) from regional to global scales [32–37]. Over West Africa, signature of floods under the dense vegetation canopy of the Congo River Basin was observed at S, C, Ku and Ka bands through an increase in the backscattering coefficient of various altimetry missions (i.e., ENVISAT, Jason-2 and SARAL) [36]. These results are very encouraging for using the backscattering coefficient to automatically identify radar altimetry measurements over water under dense vegetation cover.

The objectives of this study are to (i) demonstrate the possibility to identify water under the altimeter ground-tracks using the backscattering coefficient, important but poorly used information derived from the RA echo [38], (ii) automatically create time series of water levels over inland waterbodies, including floodplains and wetlands which are not monitored, (iii) to measure the impact of the densification of the networks of altimetry-based water levels on the surface water estimates currently performed combining nadir RA and satellite images. This assessment was achieved on the Cuvette Centrale of Congo, applying an unsupervised classification technique (k-means) to radar altimetry backscattering coefficients acquired at Ku-band (ENVISAT and Jason-2) to automatically discriminate acquisitions made over water bodies (e.g., open water such as rivers, ponds, oxbow lakes, and lakes and water under vegetation as forested floodplains and wetlands) in the extensive floodplain area of the Cuvette Centrale in the Congo River basin. The optimal selection of the cluster numbers was performed using the Calinski–Harabasz criterion [39]. The results of the clustering technique were validated against an unsupervised classification of the Cuvette Centrale validated using several data sources [40]. Based on the results of the clustering technique, the time series of water levels were generated at the cross-sections

with the river network and over the floodplains to construct a network of VSs. Over rivers, the resulting time series of water level were compared, due to the lack of in situ data in this area, this study used the altimetry-based water stages from the Hydroweb database [41]. Then, water level maps were obtained combining surface water extent from PALSAR-1 images (see [40] for details about the processing of the SAR images) and altimetry-based water levels over rivers and floodplains.

2. Study Area

The Congo River Basin (Figure 1a) is the largest drainage basin in Africa. Its shape is almost circular, covering an area of ~4 million km^2 [42]. It is the second river in the world, after the Amazon River, in terms of discharge (~40,000 m^3·s^{-1}) [43]. Its main waterway is the Congo River, with a length of ~6650 km from its source in the southeast of the Democratic Republic of Congo to the Atlantic Ocean [44]. The Congo River Basin is mainly located in the equatorial and tropical savanna climate area according to the Köppen–Geiger climate classification [45], but also in the humid subtropical climate in its southern part, and, marginally, in the semi-arid and desert climates in the east (Figure 1a).

The Cuvette Centrale is a vast floodplain located in the center of the Congo River Basin which extends from 3° S to 3° N in latitude and from 16° E to 22° E in longitude (Figure 1b). It covers an area of 1,176,000 km^2 with a wetland extent of 360,000 km^2 (32% of its area) [46]. It is the remnant of a lake which occupied the area during the Tertiary geological period that is now surrounded by mountains and plateaus [47]. As this area is difficult to access, the vegetation has been poorly studied [48]. It is mostly covered with tropical evergreen forests, and vegetation adapted to soils saturated with water (e.g., flooded forests and inundated grasslands) [40,46,49,50]. In the Cuvette Centrale are located the confluences between the Congo River and two of its major northern tributaries, the Ubangi River, and downstream, the Sangha River. The annual rainfall in the Cuvette Centrale ranges from 1400 to 1800 mm·year^{-1} and the potential evapotranspiration reaches 1280 mm·year^{-1} [51]. The topography is flat with an average slope lower than 7 cm·km^{-1} between Kisangani and Kinshasa [42], on the 50 km upstream is the outlet of the Kasai River [52] and on the 450 km upstream, is the downstream part of the Ubangi River [53]. The Cuvette Centrale is covered with sandy lacustrine Quaternary sediments where are located the forested wetlands [54,55]. These forested wetlands are subject to the flood pulse of the Congo River [56]. Water levels of the Congo River are characterized by a bimodal flooding pattern, with a main peak of high water in November–December, and a secondary one in April–May, a more pronounced low water stage in August and one of lower intensity occurring in February–March in the Cuvette Centrale. The hydrological regime of the Ubangi (main tributary in the north of the Cuvette Centrale) is characterized by a high water period peaking in November and low water levels in March [57]. Permanently inundated areas are located alongside the rivers. Other areas are inundated either during both flood events (discontinuously during a long time period of 6–7 months) or only during the maximum of the largest inundation event occurring in November–December (during a shorter time period of 2–3 months). More details about the Cuvette Centrale, its vegetation and hydrological regime can be found in [40].

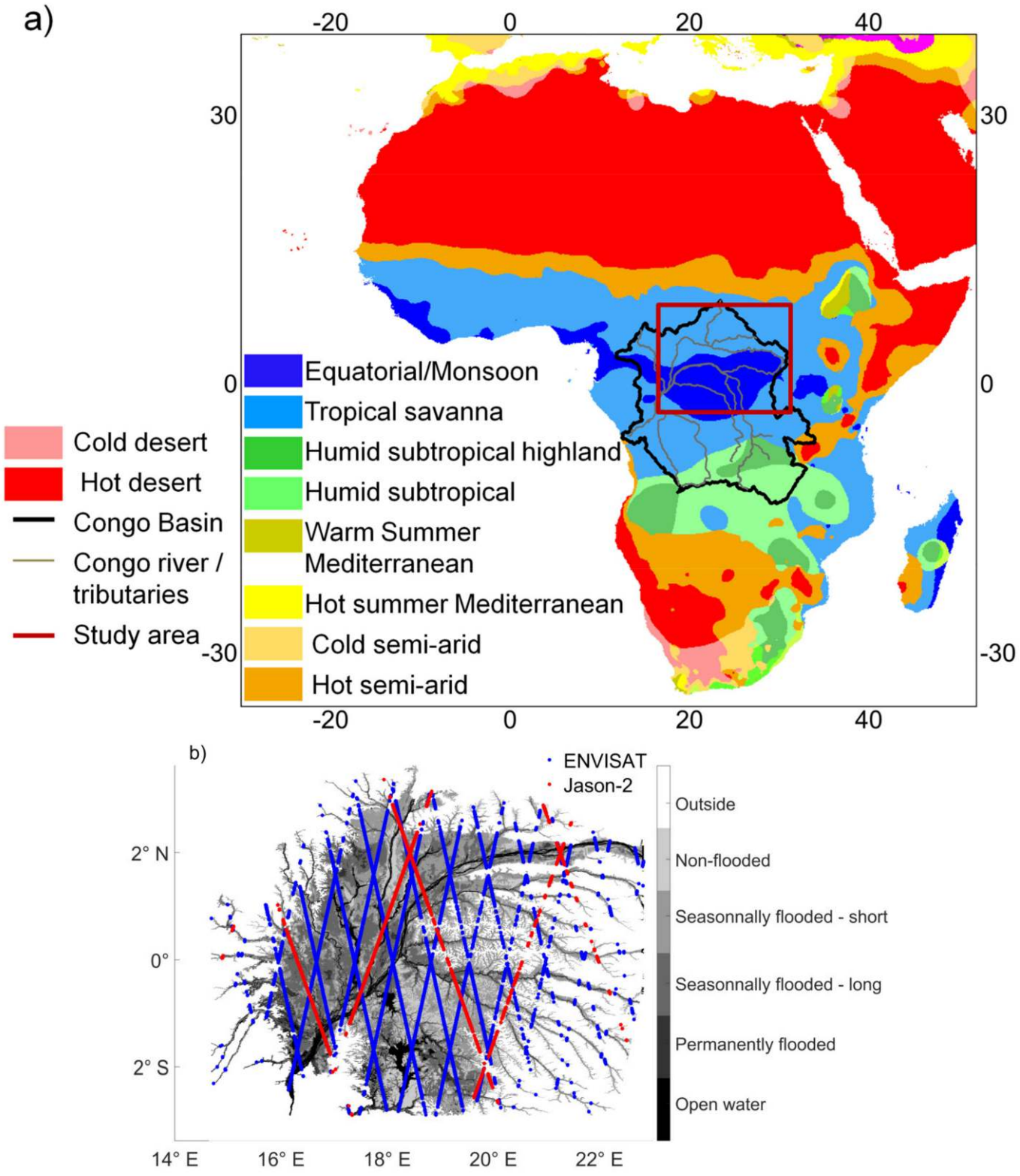

Figure 1. (**a**) The Congo River Basin and its location in the climate zones of Africa from the updated Köppen–Geiger classification [43], the red rectangle encompasses the Cuvette Centrale; (**b**) locations of the radar altimetry ground-tracks (Envisat in red, Jason-2 in blue) inside the Cuvette Centrale. Land types come from [38].

3. Datasets

3.1. Radar Altimetry Data

Data acquired by two altimetry missions (ENVISAT and Jason-2) were considered in this study. These two missions were operating on their nominal orbits from March 2002 to October 2010 and from June 2008 to September 2016 respectively. These two missions were chosen because they were operating when PALSAR-1 onboard ALOS-1 was in operation (see Section 3.2). A complete description of these two altimetry missions can be found in [58] and in [59]. In this study, the following parameters, available at high frequency (i.e., at 18 Hz for ENVISAT and 20 Hz for Jason-2, that is to say ~350 m along the altimetry ground-tracks), were used: the time and geographical locations of acquisition, the distance between the satellite and the surface or range, the orbit of the satellite, the corrections applied to the range to account for the delays caused by the path through the atmosphere (ionosphere, dry and wet atmosphere corrections) and some geophysical effects (solid earth and pole tides), the geoid model and the backscattering coefficient (σ_0) at Ku-band. Radar altimeter ranges and backscattering coefficients at Ku-band are derived from the offset center of gravity (OCOG) retracking algorithm [60], found to be well-suited for land surface studies in several previous studies (e.g., [26,61]). All these parameters, contained in the geophysical data records (GDR) of each mission, are made available by Centre de Topographie des Océans et de l'Hydrosphère (CTOH) [62], as well as indexes of the normalized tracks [63,64] that were used to compute statistics along the tracks (see Section 4.2). The location of the ground-tracks in the Cuvette Centrale is presented in Figure 1b.

3.2. Land Cover Map of the Cuvette Centrale

It results from the merging of two unsupervised classifications obtained using the k-means clustering technique:

(i) From the enhanced vegetation index (EVI) [65] from the L3 Global 500 m 16-Day moderate resolution imaging spectroradiometer (MODIS) onboard the NASA Terra satellite product (MOD13A1) from 2001 to 2009;
(ii) From a series of six PALSAR-1 images at 100 m of spatial resolution acquired at L-band, in HH polarization and at various viewing angles (from 18° to 43°), in ScanSAR mode by the phased array type synthetic aperture radar (PALSAR) sensors onboard the Advanced Land Observing Satellite (ALOS) on 7 September 2007, 25 October 2008, 25 January 2009, 27 April 2009, 13 December 2009, and 15 March 2010.

Four classes of forest types and four classes of hydrological status were obtained from the Thorndike index [66]. Through cross-comparisons with several external datasets, the topographic map of northern Congo, ICESat lidar data of elevation and canopy height and aerial photographs, five classes were identified attributed to:

(i) Open water;
(ii) Permanently flooded forests;
(iii) Seasonally flooded forests during the two Congo River flood pulses and located alongside the river;
(iv) Seasonally flooded forests during a short time corresponding to the maximum of the largest flood pulse, farther from the river;
(v) Non-flooded forests.

More details about the classification results can be found in [40].

3.3. Altimetry-Based Time Series of Water Levels

As no in situ data of water stage are available in the study area during the altimetry period, time series of water levels derived from ENVISAT and Jason-2 data over rivers in the Cuvette Centrale were used to validate the automatic estimates of river heights derived from the unsupervised classification applied to the radar altimetry backscattering coefficients. They were obtained from the Hydroweb database [41]. In this study, 24 and

32 time series from ENVISAT (35-day repeat period) and Jason-2 (10-day repeat period), respectively, were used. The length of the cross-sections between the river and the altimetry tracks range from 100 to 150 m in the upstream parts of the Cuvette Centrale to 10 km in its downstream part.

4. Methods

4.1. Radar Altimetry Data Pre-Processing

Radar altimetry data were preprocessed in order to:

(i) Derive the altimeter height (h) from the parameters contained in the GDR using the following equation (e.g., [67]):

$$h = H - R - \Sigma(\Delta R_{propagation} + \Delta R_{geophysical}) + N, \quad (1)$$

where H is the height of the center of mass of the satellite above the ellipsoid estimated using the Precise Orbit Determination (POD) technique, R is the altimeter range (i.e., the distance from the center of mass of the satellite to the surface taking into account instrumental corrections equals to $c\Delta t/2$ where c is the velocity of light in the vacuum and Δt is the two-way travel time of the electromagnetic wave emitted by the radar), $\Delta R_{propagation}$ and $\Delta R_{geophysical}$ are the sum of the geophysical and environmental corrections applied to the range, respectively, and N is the geoid.

The corrections to apply to the range for the propagation are the following:

$$\Delta R_{propagation} = \Delta R_{ion} + \Delta R_{dry} + \Delta R_{wet}, \quad (2)$$

where ΔR_{ion} is the atmospheric refraction range delay due to the free electron content associated with the dielectric properties of the ionosphere, ΔR_{dry} is the atmospheric refraction range delay due to the dry gas component of the troposphere, ΔR_{wet} is the atmospheric refraction range delay due to the water vapor and the cloud liquid water content of the troposphere.

And for the geophysical effects:

$$\Delta R_{geophysical} = \Delta R_{solid \cdot Earth} + \Delta R_{pole}, \quad (3)$$

where $\Delta R_{solid \cdot Earth}$ and ΔR_{pole} are the corrections accounting for crustal vertical motions due to the solid Earth and pole tides, respectively.

(ii) Obtain along-track time series of backscattering coefficient presenting no missing data in entry of the clustering method. To do so, monthly climatologies (i.e., the average of all data from each month averaged over the whole study period) of backscattering at Ku-band for every normalized index along the altimetry ground-tracks were computed.

4.2. Radar Altimetry Data Clustering

K-means clustering is a statistical technique designed to assign objects to a fixed number of groups or clusters based on the analysis of a set of specified variables so that the within-cluster sum of squares is minimum [68]. The initial cluster centers K are selected and then iteratively refined:

(i) Each object is assigned to its closest cluster in terms of distance to the center of the cluster;

(ii) At every iteration, each cluster center or centroid is updated to be the average of the member of the cluster.

This process is repeated until convergence.

K-means++ algorithm, which improves the quality of the final solution in terms of intra-cluster distances [69] was used. It benefits from repeated new initial cluster centroid positions. The number of times to repeat clustering using the new initial cluster centroid

positions was set to 5 and the maximum number of iterations before convergence to 100. The distance used for the k-means approach is the Euclidean distance (Ed).

In this study, the k-means approach is used to perform an unsupervised classification of the radar altimetry data using the seasonality of the backscattering at Ku-band. Data used as an input of the k-means clustering are climatological monthly variations in the backscattering at Ku-band for every normalized index along the altimetry ground-tracks [37]. The average ($\overline{\ }$) and standard deviation (*std*) of the radar altimetry backscattering coefficients (σ_0) are computed on every along-track grid-point as follows [36,37]:

$$\overline{\sigma_0}\ (dB) = 10\log_{10}\left(\overline{10^{\ \sigma_0/10}}\right), \tag{4}$$

$$std(\sigma_0)\ (dB) = 10\log_{10}\left(1 + \frac{std\left(10^{\ \sigma_0/10}\right)}{\overline{10^{\ \sigma_0/10}}}\right), \tag{5}$$

The optimal number of clusters was determined using the Calinski–Harabasz criterion (I_{CH}) and the silhouette coefficient (6 and 7, respectively). I_{CH} or variance ratio criterion was defined as the between-cluster variance and the overall within-cluster variance [39]:

$$I_{CH} = \frac{N_S - C}{C - 1}\frac{\sum_{i=1}^{C} d(u_i, U)}{\sum_{i=1}^{C}\sum_{x_j \in CL_i} Ed\left(x_j, u_i\right)}, \tag{6}$$

where N_S is the number of samples, C the number of clusters, Ed the Euclidian distance between two elements, u_i the i^{th} centroid, U the center of gravity of the whole dataset, CL_i the i^{th} cluster, x_j is the j^{th} element of the dataset. I_{CH} is the maximum value obtained for a given maximum number of clusters C. Time series of water levels were then obtained computing the median of the consecutive acquisitions from river and wetland classes. In the case of long continuous classes, time series were computed on a maximum distance of 5 km.

The silhouette coefficient provides an assessment of partitioning validity and can be used for determining the optimal number of classes [70].

The validity of the clustering is measured by computing the similarity or cohesion of each cluster member and its separation to the other clusters. The silhouette coefficient for the i^{th} cluster member is defined as follows:

$$s(i) = \frac{b(i) - a(i)}{max\{a(i), b(i)\}}, \tag{7}$$

where $a(i)$ is the average Ed between the i^{th} member in cluster A and the other members of this cluster; $b(i)$ is the average Ed between member i and the members in its second closest cluster B.

The silhouette ranges between -1 and 1. If $s(i)$ close to 1, the member i is well-assigned to cluster A. If $s(i)$ is close to 0, the object can be in-between of A and B. If $s(i)$ is close to -1, the object is badly assigned to its cluster and on average closer to members of B [70].

4.3. Automatic Generation of Time Series of Water Levels

Based on the results of the unsupervised classification, time series of water levels are generated from RA data over the classes corresponding to open water and inundated areas. A time series is generated if there are at least five non-flagged altimetry heights within a distance of 3 km non-separated between each other by more than 1 km. The minimum distance between two different VSs is 3 km.

4.4. Validation

Two types of evaluation of the classification of altimetry data are performed. Global validation is performed against the land cover map of the Cuvette Centrale from [40]

considered as the reference. The confusion matrix is estimated when comparing the classes derived from the k-means clustering technique to the classes of land types.

Local validation parameters (bias, RMSE, R) were also performed comparing the time series of water level automatically produced using the results from the unsupervised to time series of water levels from Hydroweb.

4.5. Surface Water Volume Estimates

Maps of surface water levels were obtained combining inundation extent derived from the land cover map of the Cuvette Centrale [40] with radar altimetry-based water levels from ENVISAT and Jason-2. The methodology defined in [22,23] and applied to different types of satellite images, such as SAR [22], multispectral [71] and passive microwaves [72], was used in this study. It consists in interpolating altimetry-based water stages over the inundated areas using an inverse distance weighting (IDW) technique. The wetlands inundation map was used to simulate the inundation extent over 2006–2010 as the PALSAR images used to achieve this inundation map do not allow to reach a monthly temporal resolution. During a low water period (February to April and July–August), water extent is limited to open water and permanent inundation classes (numbered 1 and 2 in Section 3.2). For the first high water period and the rising and decline of the second and larger high water period (May–June and September–October), seasonally flooded areas during the two flood pulses (class 3) are also considered, and for the peak of high water period (November–January), all classes related to water (i.e., the previously mentioned and the seasonally flooded areas during a short time corresponding to the maximum of the largest flood pulse, class 4) are merged. A map of minimum water levels recorded during the whole common observation period between SAR and altimetry was computed using a hypsometric approach to take into account the difference of altitude between the river and the floodplain (see [23] for more details on the method).

5. Results

5.1. Automatic Generation of Time Series of Water Levels

The optimal number of classes of backscattering coefficients was chosen by analyzing the Calinski–Harabasz criterion and the silhouette index from the number of classes ranging from 2 to 10 for both ENVISAT and Jason-2 data. The evolution of these indices as a function of the number of class is presented in Figure 2 for ENVISAT and Jason-2 backscattering. As silhouette coefficients are far from 1 for both ENVISAT (maximum value of 0.58) and Jason-2 (maximum of 0.51) backscattering coefficients, they cannot be used for determining the optimal number of classes from altimetry backscattering coefficients. On the contrary, I_{CH} presents a clear peak for both ENVISAT and Jason-2 for 4 and 3 classes respectively (Figure 2). Nevertheless, very similar values of I_{CH} (varying from less than 5% from the maximum value) can be observed for a number of classes ranging from 3 to 5 for both altimeters. In the following, all the results are be presented using the number of classes varying from 3 to 5.

5.2. Unsupervised Classification Results

K-means unsupervised classification was applied to ENVISAT and Jason-2 backscattering at Ku-band considering 3 to 5 classes. Figures 3 and 4 present the seasonal distribution of backscattering at Ku-band from ENVISAT and Jason-2 respectively for class numbers ranging from 3 to 5. The different classes obtained from the k-means approach exhibit a quite similar temporal behavior with a maximum occurring in December, a minimum in August, a secondary maximum in May and a secondary minimum in March. This seasonal change is in phase with the temporal variations in the water levels and discharge of the Congo River (e.g., [38,65]) but in opposite phase with the backscattering at L-band from PALSAR-1 images [38]. Larger seasonal amplitudes are observed for the classes characterized by a higher backscattering. Better separability is observed for 3 and 4 than for 5 classes (i.e., less overlap of the *std*). In this latter case, the distances between the center

of the consecutive classes can be lower than the sum of the standard deviations of these two classes. This lower separability affects more the classes with the lower backscattering coefficients and more the Jason-2 data than the ENVISAT data.

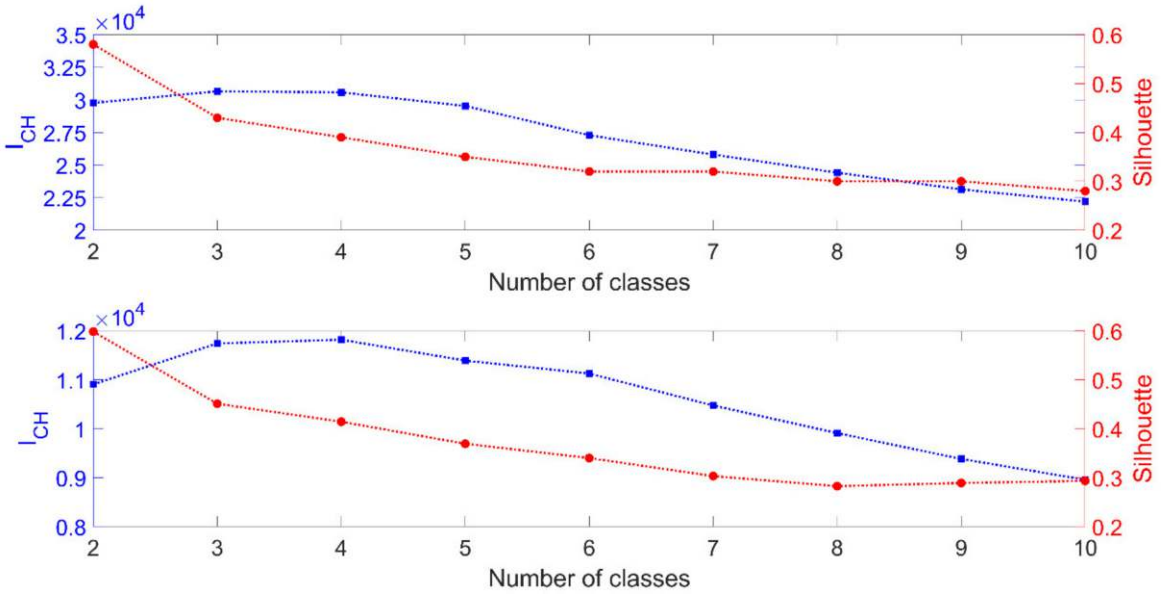

Figure 2. Evolution of the Calinski–Harabasz criterion (I_{CH} in blue) and the silhouette (in red) criterion as a function of the number of classes for ENVISAT (upper panel); Jason-2 (lower panel) backscattering coefficients.

The following classes are present in the classification of the Cuvette Centrale used in this study:

(i) Open water;
(ii) Permanently flooded forests;
(iii) Seasonally flooded forests during the two Congo River flood pulses and located alongside the river;
(iv) Seasonally flooded forests during a short time corresponding to the maximum of the largest flood pulse, farther from the river;
(v) Non-flooded forests.

From the results obtained in Section 5.1, confusion matrices were obtained using 3 to 5 classes from the radar altimetry backscattering datasets. They are presented in Table A1 (3 classes), Table A2 (4 classes) and Table 1 (5 classes) for ENVISAT and Table A3 (3 classes), Table A4 (4 classes) and Table 2 (5 classes) for Jason-2.

Table 1. Confusion matrix between 5 classes of RA backscattering coefficients from ENVISAT and the 5 land type classes identified on the study sites.

ENVISAT Class	Open Water	Flood. Perm.	Flood. Seas. Long	Flood. Seas. Short	Non-Flood.
1	0.42	0.29	0.21	0.07	0.01
2	0.19	0.26	0.36	0.17	0.02
3	0.05	0.09	0.43	0.34	0.09
4	0.00	0.03	0.34	0.49	0.14
5	0.01	0.01	0.15	0.57	0.26

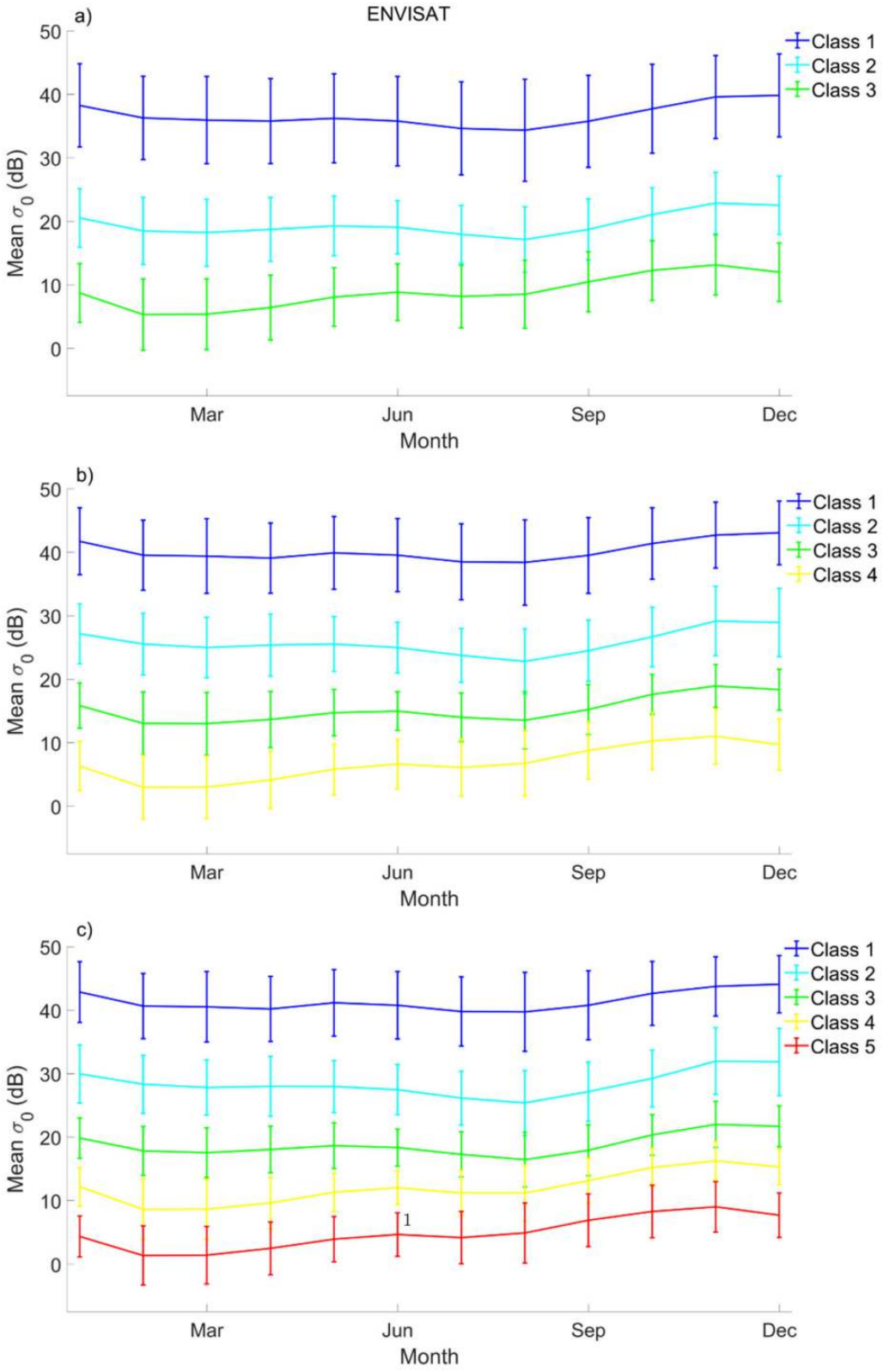

Figure 3. Seasonal variations in the mean backscattering coefficient and its associated deviation at Ku-band from ENVISAT (2003–2010) for each class when considering (**a**) 3, (**b**) 4, (**c**) 5 classes.

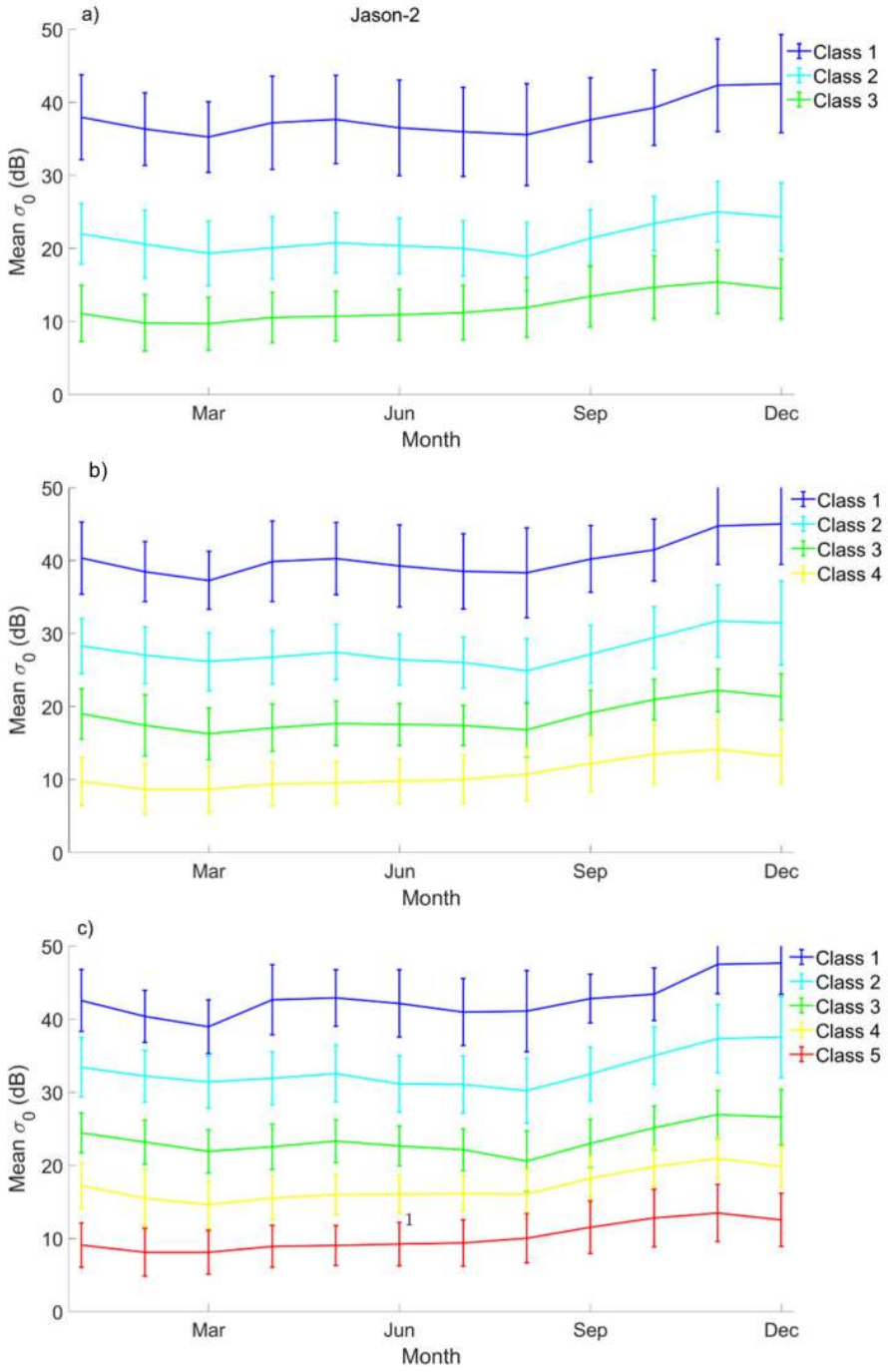

Figure 4. Seasonal variations in the mean backscattering coefficient and its associated deviation at Ku-band from Jason-2 (2008–2016) for each class when considering (**a**) 3, (**b**) 4, (**c**) 5 classes.

Table 2. Confusion matrix between 5 classes of RA backscattering coefficients from Jason-2 and the 5 land type classes identified on the study sites.

Jason-2 Class	Open Water	Flood. Perm.	Flood. Seas. Long	Flood. Seas. Short	Non-Flood.
1	0.60	0.21	0.17	0.01	0.01
2	0.21	0.29	0.31	0.17	0.02
3	0.01	0.10	0.47	0.33	0.09
4	0.00	0.01	0.38	0.45	0.16
5	0.00	0.00	0.10	0.64	0.26

Radar altimetry backscattering classes do not have the same distribution as land type classes. Maximum values range between 0.6 and 0.65, e.g., Jason-2 class 1 data are distributed at 0.6 and 0.62 in the open water class when considering 4 and 5 classes, respectively, at 0.62 and 0.64 for short duration flooded forests (Table 2 and Table A4 respectively). These results can appear low compared with those obtained for images classification. However, as the footprint of the radar altimeter has a radius of several kilometers, it encompasses very heterogenous surfaces in terms of roughness, soil and vegetation types, moisture content and presence or not of surface water.

Previous studies analyzing radar altimetry backscattering coefficients over land surfaces showed that high values of backscattering are observed over rivers and floodplains and low values over vegetation [32,35,36,64,73].

When merging results from predominantly water and vegetation, we obtain:

- The results 0.89, 0.91 and 0.92 for ENVISAT 0.89, 0.93, 0.98 for Jason-2 when considering the sum of the backscattering values in class 1 over open water, permanently and long duration flooded forests;
- The results 0.72, 0.89, 0.93 for ENVISAT and 0.85, 0.79, 0.90 for Jason-2 when considering the sum of the backscattering values in the last class over short duration flooded and non-flooded forests.

The spatial distribution of the backscattering classes is presented for ENVISAT and Jason-2 over the whole Cuvette Centrale in Figure 5 and a smaller area in Figure 6 to illustrate the good agreement between the unsupervised classification and the land types.

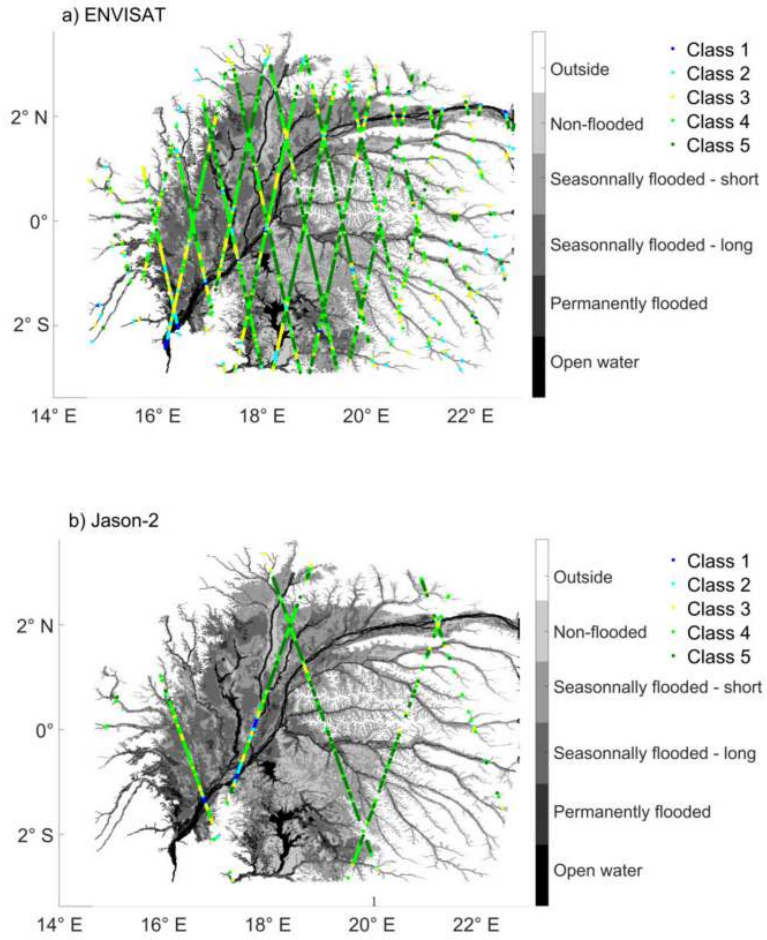

Figure 5. Spatial distribution of the backscattering classes (Ku band) from (**a**) ENVISAT and (**b**) Jason-2 in the Cuvette Centrale of Congo.

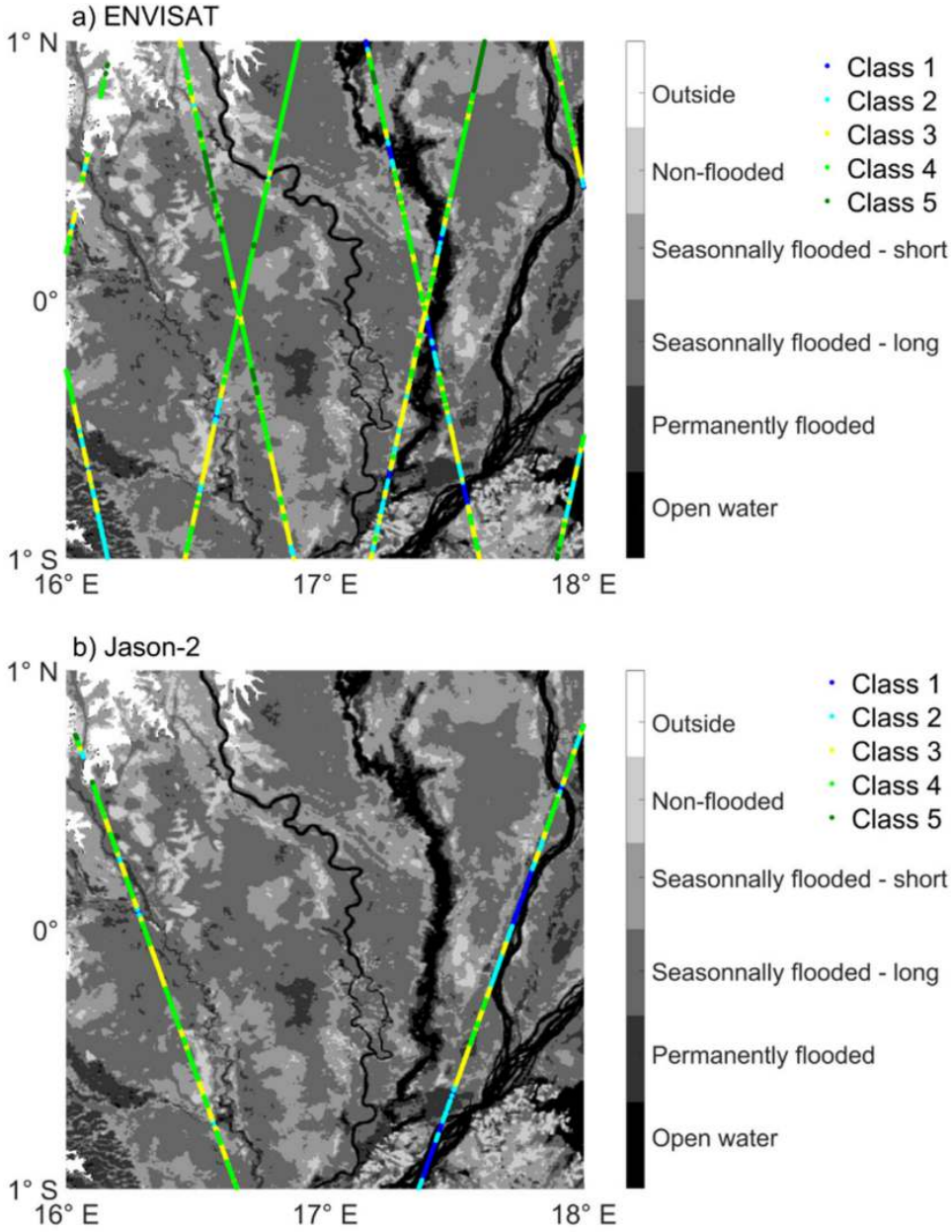

Figure 6. Same as Figure 5 for a smaller area.

5.3. Time Series of Water Levels and Volumes

A total of 250 and 358 time series of water levels were automatically generated over the Cuvette Centrale of Congo using Jason-2 and ENVISAT data respectively for five classes of backscattering (Figure 7). Comparisons were performed between them and the closest location (below 5 km) of a Hydroweb VS. Bias, RMSE and R were estimated for 24 Jason-2 and 32 ENVISAT VSs (Figure 8). Very good agreement was found between the automatically generated time series and the Hydroweb ones with R generally higher than 0.95 (18 out of 24 for Jason-2 and 31 out of 32 for ENVISAT), RMSE lower than 0.25 m (15 out of 24 for Jason-2 VS, 22 out 32 for ENVISAT VSs) and bias lower than 0.25 m (22 out of 24 for Jason-2 VS, 26 out of 32 for ENVISAT VSs).

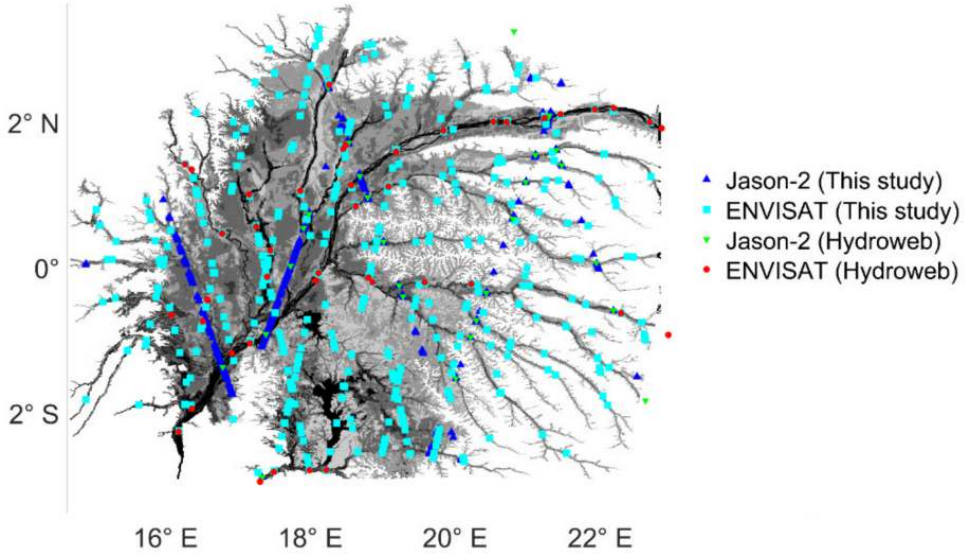

Figure 7. Locations of the VSs automatically created on the Jason-2 (blue triangles) and ENVISAT (light blue squares) ground-tracks and of the VSs from Hydroweb (Jason-2, green triangles; and ENVISAT, red circles) used for comparison.

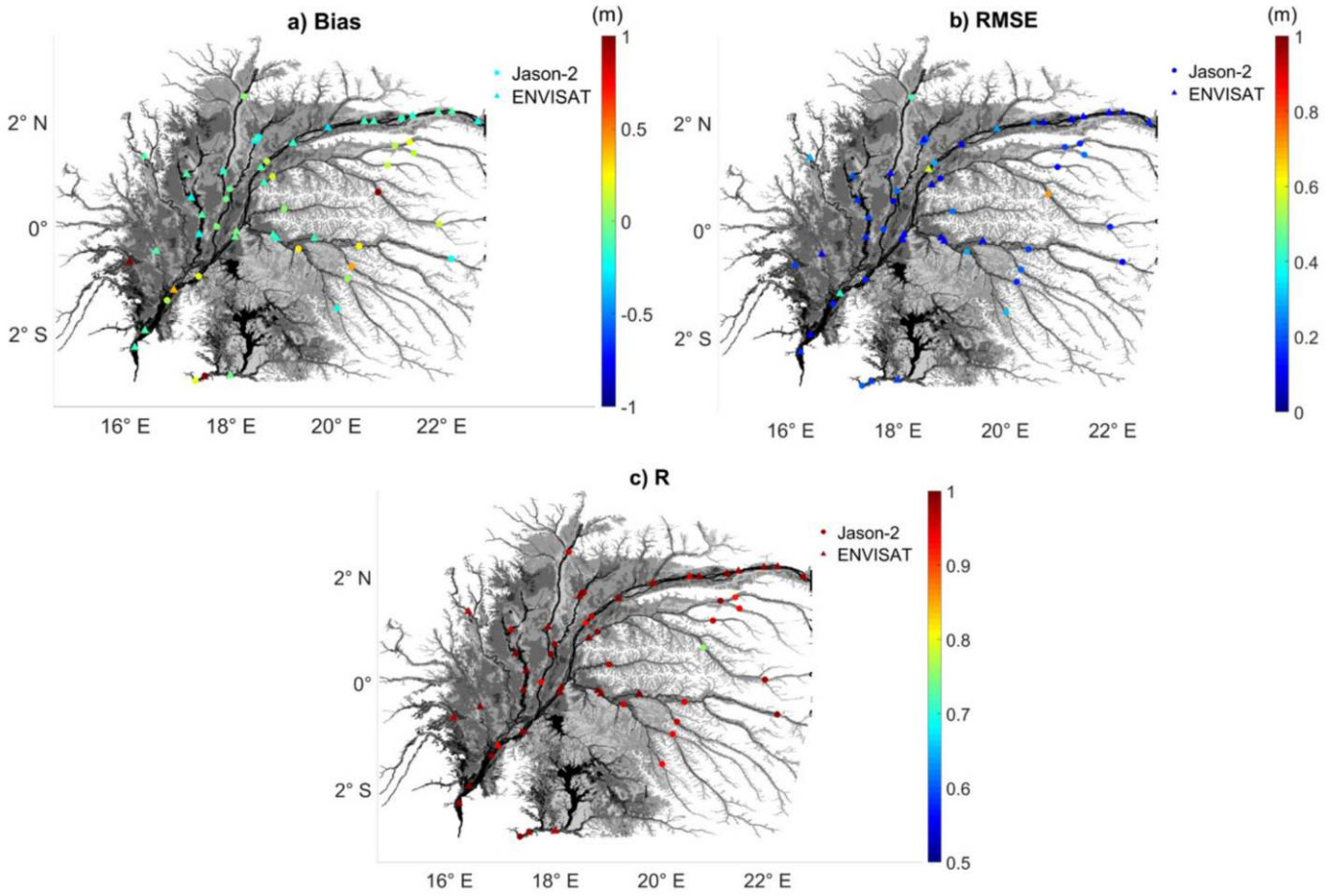

Figure 8. Comparison between altimetry-based of water levels from Hydroweb and automatically generated (this study) in terms of (**a**) bias, (**b**) RMSE and (**c**) r. VS locations are represented using either a dot (Jason-2) or a triangle (ENVISAT).

5.4. Surface Water Storage

Monthly maps of surface water levels and associated water volume variations were derived interpolating altimetry-based water levels on the wetland extent at a monthly temporal resolution. To avoid differences caused by intermission biases, these maps were obtained only using ENVISAT VSs. The VSs used were the ones automatically derived over either the rivers or the rivers and the floodplains. The associated surface water volume variations were computed in both cases. The annual amplitude for the largest flood period (November–December) is around four times higher (above 800 km^3) and the flood peak occurs 1 to 2 months later (December–January) only using VSs on rivers then using VSs on both the rivers and the floodplains (Figure 9). These differences of volume are caused by differences in spatial patterns of the water stage. Maps of the minimum height were obtained from ENVISAT. The minimum height is either the minimum of the water level during the observation period over rivers and permanently flooded pixels, or the bottom topography of floodplains over non-permanently inundated areas. In Figure 10 are presented maps based on the VSs automatically created over the rivers and the floodplains (a), or based on VSs manually created over the rivers and made available by Hydroweb (b). As it can be seen from their difference (c), the minimum height is quite similar on the center of the Cuvette Centrale, and increases up to ~30 m on the edges of the study area. Considering the mean annual amplitude, either automatically created over rivers and floodplains (d) or manually created over rivers only (e), differences up to ~10 m, positive in the north and negative in the south can be observed (f).

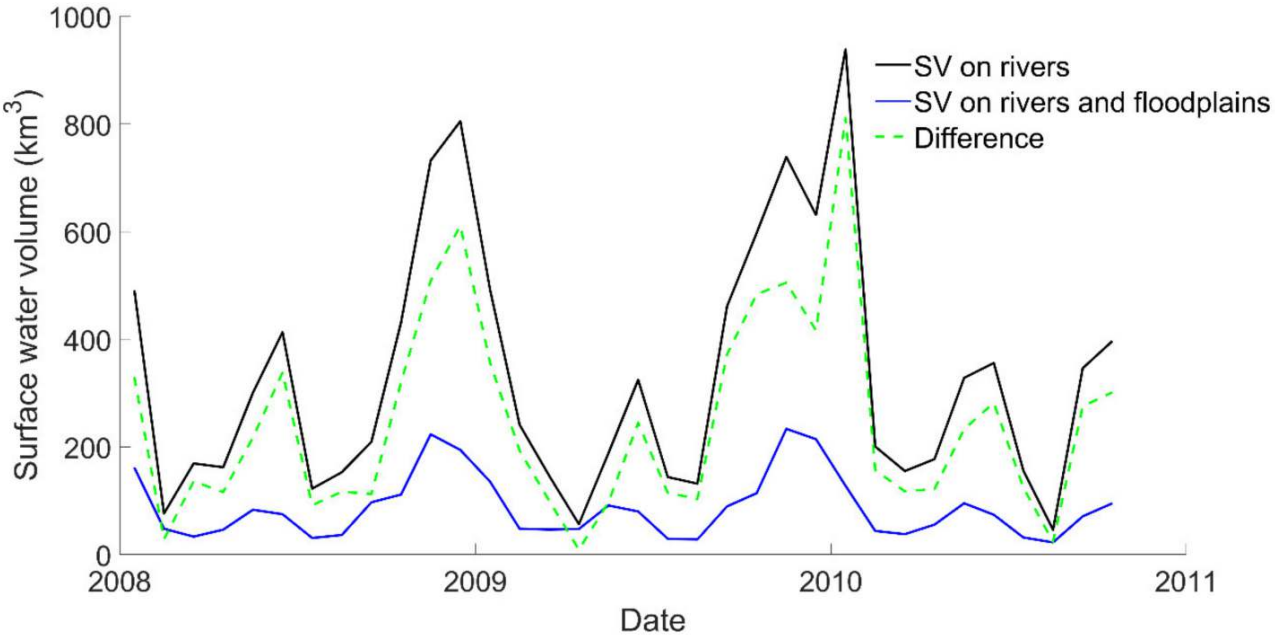

Figure 9. Time series of surface water storage of the Cuvette Centrale of Congo obtained combining the surface water extent and the time series of water levels from the ENVISAT VSs on the rivers (black), on the rivers and floodplains (blue). Their difference is presented in dashed green.

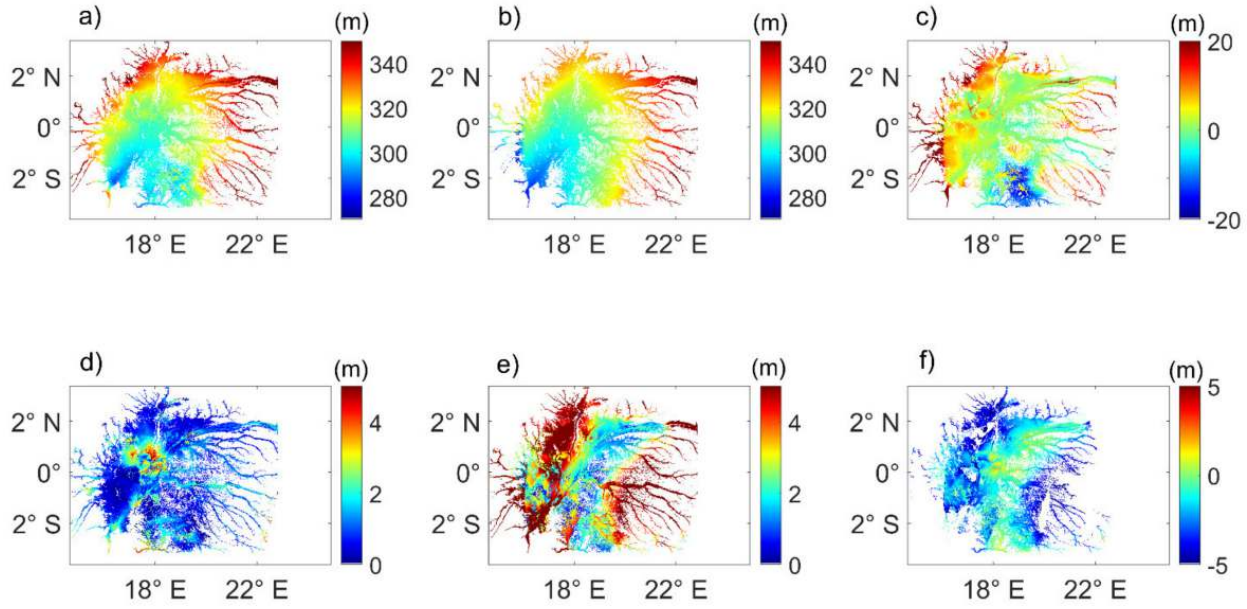

Figure 10. Maps of minimum of height from ENVISAT based on the VSs (**a**) automatically created over the rivers and the floodplains or (**b**) manually created over the rivers and made available by Hydroweb, and (**c**) their difference. Maps of mean annual amplitude using the VSs (**d**) over the rivers and floodplains, (**e**) only over the rivers, and (**f**) their difference.

6. Discussion

6.1. Identification of Open Water on RA Ground-Tracks

The results of the classification of the RA backscattering coefficients show that this parameter can be used to efficiently discriminate between water and vegetation under both ENVISAT and Jason-2 ground-tracks. Contrary to SAR or multi-spectral images, this type of sensor seems to be unable to discriminate between open water and water under vegetations in the floodplains (see the results of the confusion matrices presented in Tables 1, 2 and A1–A4). This is most likely due to:

(i) The size of the illuminated area at Ku-band (several km of diameter in low resolution mode (LRM)), the scene present in the altimeter footprint is very heterogenous and encompasses both rivers and the surrounding floodplains;

(ii) The RA sensor is acquiring data at nadir. As a consequence, the power backscattered by the water dominates the radar echo. RA backscattering are much stronger over rivers and floodplains than over any other types of land cover (e.g., [32,36,37]). Even under forest canopy, water levels can be retrieved (e.g., [22] and this study).

All the classes exhibit a similar temporal behavior even for the classes corresponding to non-inundated areas. The smaller seasonal cycle observed for these classes can be attributed to an increase in soil moisture during the wet periods, which a change in the dielectrical properties of the soil, and hence a rise in σ_0 which can be detected even under a dense vegetation cover [35–37,74].

This good ability allowed to automatically generate time series of water levels on the RA ground-tracks; 358/250 VSs were generated on the ENVISAT/Jason-2 ground-tracks, respectively. The ones on the river were validated against manually created time series of water levels from Hydroweb due to the lack of in-situ data for validation. Very similar results were obtained when using the approach presented in this study versus the manual creation of VSs (RMSE $\leq$ 0.25 m in more than 80%/90% of the 32/24 cases and R $\geq$ 0.95 in more than 95%/75% of the same cases for ENVISAT and Jason-2, respectively). The time series of the anomaly of water levels from ENVISAT and Jason-2, averaged over all the stations located on the river and the permanently inundated areas, and over non-permanently inundated areas are presented on Figure 11. The temporal variations are consistent with the hydrological regimes of the area: a first maximum in April–May, a secondary larger one

in November–December, as well as the annual amplitude around 3–4 m (e.g., [65,67,68]) and lower over the floodplains [47], as it can be seen when considering the average ± 1 *std*. The lower mean annual amplitude is observed on time series from ENVISAT (Figure 11a,c) than Jason-2 (Figure 11b,d). This is most likely due to the longer temporal revisit period of ENVISAT (35 days) than Jason-2 (10 days) and the higher number of VSs on the ENVISAT ground-tracks (358) than Jason-2 (250).

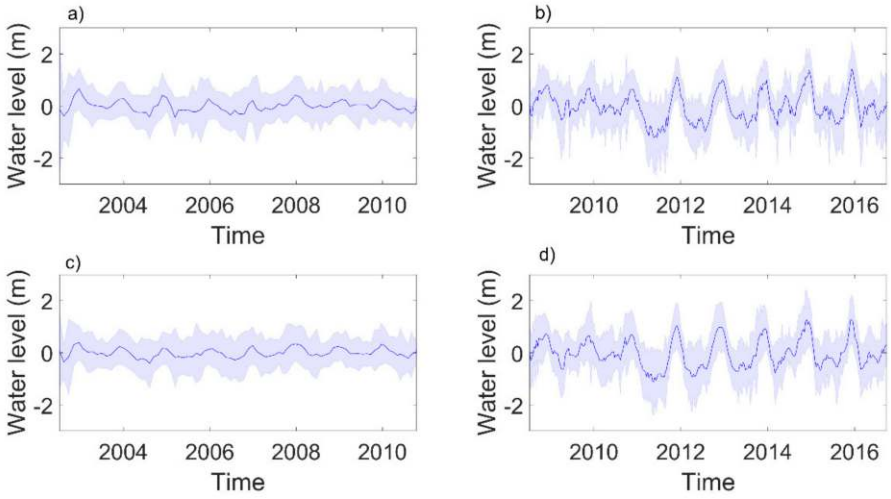

Figure 11. Time series of anomaly of water levels averaged over all the stations located on the river and the permanently inundated areas from (**a**) ENVISAT and (**b**) Jason-2, and over the non-permanently inundated areas from (**c**) ENVISAT and (**d**) Jason-2.

6.2. Impact on Anomaly of Surface Water Storage

The densification of the network of VSs with the definition of new VSs over the floodplains and wetlands offer a unique opportunity to better understand the dynamics of this hydrological reservoir. Two parameters are used to characterize the surface reservoir: the discharge and the storage. The combination of surface extent with water level variations provides access to the surface water storage changes (e.g., [22,71]). In most of the river basins, time series of water level from RA are only recorded at VSs over rivers. This study offers an opportunity to measure the impact of the inclusion of VSs over floodplains on the estimates of surface water storage variations. Even if the real time series of inundation extent is not considered here, realistic monthly variations in this parameter were used. Note that a recent study showed that the use of a climatology of surface water extent was sufficient to retrieve accurate variations in surface water storage [75]. Our study showed that the differences in bottom topography of the floodplains (Figure 10a–c) and in the patterns of the annual amplitude of water level (Figure 10d–f) have large consequences on the amplitude of the variations in surface water storage in the Cuvette Centrale of Congo. The lack of VSs on rivers for the upstream parts in the northwest of the study area leads to an underestimate of the bottom topography from 10 to 30 m, and an overestimate of the annual amplitude above 5 m, causing an overestimate of the storage by a factor of four. Besides, spatial patterns of the annual amplitude are more consistent with the spatial distribution of the wetlands in the Cuvette Centrale. For instance, if we consider the region centered on 18° E and from 0 to 2° N, high annual amplitude is observed in the Ngiri and Eulonga basins (below 1° N), a close to zero annual amplitude above, in the Upper Ubangi basin when using the VSs on the floodplains in the interpolation (Figure 10d), in good agreement with the map of the probability of wetlands [46]. On the contrary, high annual amplitude is found on the whole area when only using the VSs on the rivers (Figure 10e). These differences account for the change in amplitude and temporality (Figure 9) between the study cases (VSs only over the rivers or on the rivers and the floodplains). The influence of the larger changes in water levels observed over the rivers is limited when increasing the number of VSs in the IDW both on the storage and its seasonality (e.g., the signal is

dominated by the largest floodplains and the time-lag of the seasonal cycle of the Congo River tributaries).

7. Conclusions

Radar altimetry is a unique tool for measuring water levels over inland waterbodies, and especially over the non-gauged wetlands and floodplains. Identifying the presence of water in wetlands and floodplains (where and when) makes difficult the construction of VSs over these environments. In this study, applying a k-means unsupervised classification approach to RA backscattering coefficients allows to accurately identify the flooded (good detection over 90%) and non-flooded areas over a complex equatorial area, the Cuvette Centrale of Congo, where floodplains are mostly covered with vegetation. Contrary to other types of satellite products, RA, due to the size of its footprint (several kilometers of diameter at Ku-band), seems unable to discriminate between open water and permanently flooded areas or to provide information on the type of vegetation covering the non-flooded area. Using the classification results, time series of water levels were automatically created over the rivers and the floodplains of the Cuvette Centrale of Congo (358/250 on the ENVISAT and Jason-2 ground-tracks, respectively). The resulting time series exhibit temporal variations in good agreement with the hydrological regime of the study area in terms of temporal variations and annual amplitude. Very similar results to the manually created time series of water levels were obtained: (RMSE $\leq$ 0.25 m in more than 80%/90% of the 32/24 cases and R $\geq$ 0.95 in more than 95%/75% of the same cases for ENVISAT and Jason-2, respectively). The densification of the VSs network by including the floodplains has a strong implication on the water volume estimate.

Owing to the larger number of VSs used when considering the VSs on the floodplains, the interpolation of the water level maps is more constrained, and the spatial patterns of annual amplitudes are more consistent with the information on the wetland extent. With the availability of data from recent RA missions operating in SAR mode such as SENTINEL-3A and B, better discrimination based on the backscattering can be expected, allowing to continue the monitoring with a better accuracy of the time series of water levels over the floodplains. This approach is likely to be applied to any other extensive wetlands or floodplains and will contribute to a better estimate, in combination with satellite images, to the changes in surface water storage over land. As the high precision altimetry era started in the early/mid 1990s with the launch of Topex/Poseidon (1992) and ERS-1/2 (1991 and 1995, respectively), a long-term record of almost 30 years of surface water storage variations can soon be available. All this information will have a large importance for better understanding the hydrological processes in the floodplains and wetlands and for the comparison with data acquired by the future NASA/CNES SWOT mission, to be launched in 2022, which will the first to estimate water levels in a swath.

Author Contributions: Conceptualization, F.F., J.B. and V.G.; methodology, F.F., J.D., L.B., F.S.; processing, F.F., P.Z. and R.B.; data curation, F.B., N.B., F.S.; writing—original draft preparation, F.F.; writing—review and editing, all authors; funding acquisition, F.F., J.B., V.G. and J.D. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by CNES TOSCA grants number CASCHMIR and SWHYM. The APC was offered by MDPI.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Radar altimetry data can be found here: http://ctoh.legos.obs-mip.fr/applications/land_surfaces/altimetric_data/altis/altis. Time series of water levels can be found here: http://hydroweb.theia-land.fr/. Accessed on 1 April 2019.

Acknowledgments: We thank Sylvie Gourlet-Fleury (CIRAD) for helpful discussion during the preparation of the manuscript. We also thank three anonymous reviewers for their meaningful comments which helped us improving the quality of our manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

Table A1. Confusion matrix between 3 classes of RA backscattering coefficients from ENVISAT and the 5 land type classes identified on the study sites.

ENVISAT Class	Open Water	Flood. Perm.	Flood. Seas. Long	Flood. Seas. Short	Non-Flood.
1	0.33	0.30	0.26	0.10	0.01
2	0.06	0.10	0.41	0.34	0.09
3	0.01	0.02	0.25	0.53	0.19

Table A2. Confusion matrix between 4 classes of RA backscattering coefficients from ENVISAT and the 5 land type classes identified on the study sites.

ENVISAT Class	Open Water	Flood. Perm.	Flood. Seas. Long	Flood. Seas. Short	Non-Flood.
1	0.40	0.29	0.22	0.08	0.01
2	0.08	0.17	0.43	0.26	0.06
3	0.00	0.04	0.42	0.41	0.13
4	0.00	0.00	0.11	0.64	0.25

Table A3. Confusion matrix between 3 classes of RA backscattering coefficients from Jason-2 and the 5 land type classes identified on the study sites.

Jason-2 Class	Open Water	Flood. Perm.	Flood. Seas. Long	Flood. Seas. Short	Non-Flood.
1	0.41	0.25	0.24	0.09	0.01
2	0.01	0.07	0.47	0.35	0.10
3	0.00	0.00	0.15	0.62	0.23

Table A4. Confusion matrix between 4 classes of RA backscattering coefficients from Jason-2 and the 5 land type classes identified on the study sites.

Jason-2 Class	Open Water	Flood. Perm.	Flood. Seas. Long	Flood. Seas. Short	Non-Flood.
1	0.48	0.26	0.19	0.06	0.01
2	0.15	0.21	0.39	0.21	0.04
3	0.01	0.05	0.40	0.42	0.12
4	0.01	0.02	0.18	0.57	0.22

References

1. Davidson, N.C.; Fluet-Chouinard, E.; Finlayson, C.M. Global extent and distribution of wetlands: Trends and issues. *Mar. Freshw. Res.* **2018**, *69*, 620. [CrossRef]
2. Mitsch, W.J.; Gosselink, J.G.; Zhang, L.; Anderson, C.J. *Wetland Ecosystems*; John Wiley & Sons: Hoboken, NJ, USA, 2009.
3. Bullock, A.; Acreman, M. The role of wetlands in the hydrological cycle. *Hydrol. Earth Syst. Sci.* **2003**, *7*, 358–389. [CrossRef]
4. Acreman, M.; Holden, J. How wetlands affect floods. *Wetlands* **2013**, *33*, 773–786. [CrossRef]
5. Ghermandi, A.; van der Bergh, J.; Brander, L.M.; Nunes, P. *The Economic Value of Wetland Conservation and Creation: A Meta-Analysis*; Working Paper 79; Fondazione Eni Enrici Mattei: Milan, Italy, 2008.
6. Junk, W.J.; Brown, M.; Campbell, I.C.; Finlayson, M.; Gopal, B.; Ramberg, L.; Warner, B.G. The comparative biodiversity of seven globally important wetlands: A synthesis. *Aquat. Sci.* **2006**, *68*, 400–414. [CrossRef]
7. Mitra, S.; Wassmann, R.; Vlek, P.L.G. An appraisal of global wetland area and its organic carbon stock. *Curr. Sci.* **2005**, *88*, 25–35.
8. Webb, E.B.; Smith, L.M.; Vrtiska, M.P.; Lagrange, T.G. Effects of Local and Landscape Variables on Wetland Bird Habitat Use During Migration Through the Rainwater Basin. *J. Wildl. Manag.* **2009**, *74*, 109–119. [CrossRef]
9. Haukos, D.A.; Smith, L.M. The importance of playa wetlands to biodiversity of the Southern High Plains. *Landsc. Urban. Plan.* **1994**, *28*, 83–98. [CrossRef]

10. Gibbons, J.W. Terrestrial habitat: A vital component for herpetofauna of isolated wetlands. *Wetlands* **2006**, *23*, 630–635. [CrossRef]
11. Maltby, E.; Immirzi, P. Carbon dynamics in peatlands and other wetland soils regional and global perspectives. *Chemosphere* **1993**, *27*, 999–1023. [CrossRef]
12. Mitsch, W.J.; Bernal, B.; Nahlik, A.M.; Mander, Ü.; Zhang, L.; Anderson, C.J.; Jørgensen, S.E.; Brix, H. Wetlands, carbon, and climate change. *Landsc. Ecol.* **2013**, *28*, 583–597. [CrossRef]
13. Whalen, S.C. Biogeochemistry of Methane Exchange between Natural Wetlands and the Atmosphere. *Environ. Eng. Sci.* **2005**, *22*, 73–94. [CrossRef]
14. Bergamaschi, P.; Frankenberg, C.; Meirink, J.F.; Krol, M.; Dentener, F.; Wagner, T.; Platt, U.; Kaplan, J.O.; Körner, S.; Heinmann, M.; et al. Satellite charactography of atmospheric methane from SCIAMACHY on board ENVISAT: 2. Evaluation based on inverse model simulations. *J. Geophys. Res. Atmos.* **2007**, *112*, 73–94. [CrossRef]
15. Bloom, A.A.; Palmer, P.I.; Fraser, A.; David, S.R.; Frankenberg, C. Large-scale controls of methanogenesis inferred from methane and gravity spaceborne data. *Science* **2010**, *327*, 322–325. [CrossRef]
16. Ringeval, B.; De Noblet-Ducoudré, N.; Ciais, P.; Bousquet, P.; Prigent, C.; Papa, F.; Rossow, W.B. An attempt to quantify the impact of changes in wetland extent on methane emissions on the seasonal and interannual time scales. *Glob. Biogeochem. Cycles* **2010**, *24*. [CrossRef]
17. Biancamaria, S.; Lettenmaier, D.P.; Pavelsky, T.M. The SWOT Mission and Its Capabilities for Land Hydrology. *Surv. Geophys.* **2016**, *37*, 307–337. [CrossRef]
18. Alsdorf, D.E.; Smith, L.C.; Melack, J.M. Amazon floodplain water level changes measured with interferometric SIR-C radar. *IEEE Trans. Geosci. Remote Sens.* **2001**, *39*, 423–431. [CrossRef]
19. Lee, H.; Yuan, T.; Jung, H.C.; Beighley, E. Mapping wetland water depths over the central Congo Basin using PALSAR ScanSAR, Envisat altimetry, and MODIS VCF data. *Remote Sens. Environ.* **2015**, *159*, 70–79. [CrossRef]
20. Papa, F.; Frappart, F.; Güntner, A.; Prigent, C.; Aires, F.; Getirana, A.C.V.A.C.V.; Maurer, R. Surface freshwater storage and variability in the Amazon basin from multi-satellite observations, 1993–2007. *J. Geophys. Res. Atmos.* **2013**, *118*, 11951–11965. [CrossRef]
21. Salameh, E.; Frappart, F.; Papa, F.; Güntner, A.; Venugopal, V.; Getirana, A.; Prigent, C.; Aires, F.; Labat, D.; Laignel, B. Fifteen years (1993–2007) of surface freshwater storage variability in the ganges-brahmaputra river basin using multi-satellite observations. *Water* **2017**, *9*, 245. [CrossRef]
22. Frappart, F.; Seyler, F.; Martinez, J.-M.; León, J.G.; Cazenave, A. Floodplain water storage in the Negro River basin estimated from microwave remote sensing of inundation area and water levels. *Remote Sens. Environ.* **2005**, *99*. [CrossRef]
23. Frappart, F.; Papa, F.; Santos Da Silva, J.; Ramillien, G.; Prigent, C.; Seyler, F.; Calmant, S. Surface freshwater storage and dynamics in the Amazon basin during the 2005 exceptional drought. *Environ. Res. Lett.* **2012**, *7*. [CrossRef]
24. Alsdorf, D.; Han, S.C.; Bates, P.; Melack, J. Seasonal water storage on the Amazon floodplain measured from satellites. *Remote Sens. Environ.* **2010**, *114*, 2448–2456. [CrossRef]
25. Lee, H.; Beighley, R.E.; Alsdorf, D.; Jung, H.C.; Shum, C.K.; Duan, J.; Guo, J.; Yamazaki, D.; Andreadis, K. Characterization of terrestrial water dynamics in the Congo Basin using GRACE and satellite radar altimetry. *Remote Sens. Environ.* **2011**, *115*, 3530–3538. [CrossRef]
26. Santos da Silva, J.; Calmant, S.; Seyler, F.; Rotunno Filho, O.C.; Cochonneau, G.; Mansur, W.J. Water levels in the Amazon basin derived from the ERS 2 and ENVISAT radar altimetry missions. *Remote Sens. Environ.* **2010**, *114*, 2160–2181. [CrossRef]
27. Frappart, F.; Papa, F.; Marieu, V.; Malbeteau, Y.; Jordy, F.; Calmant, S.; Durand, F.; Bala, S. Preliminary Assessment of SARAL/AltiKa Observations over the Ganges-Brahmaputra and Irrawaddy Rivers. *Mar. Geod.* **2015**, *38*, 568–580. [CrossRef]
28. Normandin, C.; Frappart, F.; Diepkilé, A.T.; Marieu, V.; Mougin, E.; Blarel, F.; Lubac, B.; Braquet, N.; Ba, A. Evolution of the performances of radar altimetry missions from ERS-2 to Sentinel-3A over the Inner Niger Delta. *Remote Sens.* **2018**, *10*, 833. [CrossRef]
29. Frappart, F.; Blarel, F.; Fayad, I.; Bergé-Nguyen, M.; Crétaux, J.-F.; Shu, S.; Schregenberger, J.; Baghdadi, N. Evaluation of the Performances of Radar and Lidar Altimetry Missions for Water Level Retrievals in Mountainous Environment: The Case of the Swiss Lakes. *Remote Sens.* **2021**, *13*, 2196. [CrossRef]
30. Schwatke, C.; Dettmering, D.; Bosch, W.; Seitz, F. DAHITI—An innovative approach for estimating water level time series over inland waters using multi-mission satellite altimetry. *Hydrol. Earth Syst. Sci.* **2015**, *19*, 4345–4364. [CrossRef]
31. Okeowo, M.A.; Lee, H.; Hossain, F.; Getirana, A. Automated Generation of Lakes and Reservoirs Water Elevation Changes from Satellite Radar Altimetry. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2017**, *10*, 3465–3481. [CrossRef]
32. Papa, F.; Legrésy, B.; Rémy, F. Use of the Topex–Poseidon dual-frequency radar altimeter over land surfaces. *Remote Sens. Environ.* **2003**, *87*, 136–147. [CrossRef]
33. Legrésy, B.; Papa, F.; Remy, F.; Vinay, G.; van den Bosch, M.; Zanife, O.-Z. ENVISAT radar altimeter measurements over continental surfaces and ice caps using the ICE-2 retracking algorithm. *Remote Sens. Environ.* **2005**, *95*, 150–163. [CrossRef]
34. Papa, F.; Prigent, C.; Rossow, W.B.; Legresy, B.; Remy, F. Inundated wetland dynamics over boreal regions from remote sensing: The use of Topex-Poseidon dual-frequency radar altimeter observations. *Int. J. Remote Sens.* **2006**, *27*, 4847–4866. [CrossRef]
35. Fatras, C.; Frappart, F.; Mougin, E.; Frison, P.L.; Faye, G.; Borderies, P.; Jarlan, L. Spaceborne altimetry and scatterometry backscattering signatures at C- and Ku-bands over West Africa. *Remote Sens. Environ.* **2015**, *159*, 117–133. [CrossRef]

36. Frappart, F.; Fatras, C.; Mougin, E.; Marieu, V.; Diepkilé, A.T.; Blarel, F.; Borderies, P. Radar altimetry backscattering signatures at Ka, Ku, C, and S bands over West Africa. *Phys. Chem. Earth* **2015**, *83–84*, 96–110. [CrossRef]
37. Frappart, F.; Blarel, F.; Papa, F.; Prigent, C.; Mougin, E.; Paillou, P.; Baup, F.; Zeiger, P.; Salameh, E.; Darrozes, J.; et al. Backscattering signatures at Ka, Ku, C and S bands from low resolution radar altimetry over land. *Adv. Sp. Res.* **2021**, *68*, 989–1012. [CrossRef]
38. Abdalla, S.; Abdeh Kolahchi, A.; Ablain, M.; Adusumilli, S.; Aich Bhowmick, S.; Alou-Font, E.; Amarouche, L.; Andersen, O.B.; Antich, H.; Aouf, L.; et al. Altimetry for the future: Building on 25 years of progress. *Adv. Sp. Res.* **2021**, *68*, 319–363. [CrossRef]
39. Caliński, T.; Harabasz, J. A Dendrite Method Foe Cluster Analysis. *Commun. Stat.* **1974**, *3*, 1–27. [CrossRef]
40. Betbeder, J.; Gond, V.; Frappart, F.; Baghdadi, N.N.; Briant, G.; Bartholome, E. Mapping of central africa forested wetlands using remote sensing. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2014**, *7*. [CrossRef]
41. Hydroweb. Available online: http://hydroweb.theia-land.fr/ (accessed on 6 April 2020).
42. Devroey, E. *Annuaire Hydrologique du Congo et Ruanda-Urundi, 1959*; Institut Royal Colonial Belge: Brussels, Belgium, 1961.
43. Bricquet, J.P. Les Ecoulements Du Congo a Brazzaville Et La Spatialisation Des Apports. *Gd. Bassins Fluviaux* **1993**, 22–24.
44. Runge, J. The Congo River, Central Africa. In *Large Rivers: Geomorphology and Management*; John Wiley & Sons Ltd: Chichester, UK, 2008; pp. 293–309. ISBN 9780470723722.
45. Peel, M.C.; Finlayson, B.L.; McMahon, T.A. Updated world map of the Köppen-Geiger climate classification. *Hydrol. Earth Syst. Sci.* **2007**, *11*, 1633–1644. [CrossRef]
46. Bwangoy, J.R.B.; Hansen, M.C.; Roy, D.P.; De Grandi, G.; Justice, C.O. Wetland mapping in the Congo Basin using optical and radar remotely sensed data and derived topographical indices. *Remote Sens. Environ.* **2010**, *114*, 73–86. [CrossRef]
47. Robert, M. *Le Congo Physique*; Presses Universitaires de France: Liège, Belgium, 1946.
48. Campbell, D. The Congo river basin. In *The World's Largest Wetlands: Ecology and Conservation*; Cambridge University Press: Cambridge, UK, 2005.
49. Vancutsem, C.; Pekel, J.F.; Evrard, C.; Malaisse, F.; Defourny, P. Mapping and characterizing the vegetation types of the Democratic Republic of Congo using SPOT VEGETATION time series. *Int. J. Appl. Earth Obs. Geoinf.* **2009**, *11*, 62–76. [CrossRef]
50. Verhegghen, A.; Mayaux, P.; De Wasseige, C.; Defourny, P. Mapping Congo Basin vegetation types from 300 m and 1 km multi-sensor time series for carbon stocks and forest areas estimation. *Biogeosciences* **2012**, *9*, 5061–5079. [CrossRef]
51. FAO World-Wide Agroclimatic Data of FAO (FAOCLIM) | Land & Water | Food and Agriculture Organization of the United Nations | Land & Water | Food and Agriculture Organization of the United Nations. Available online: http://www.fao.org/land-water/land/land-governance/land-resources-planning-toolbox/category/details/en/c/1028000/ (accessed on 17 March 2019).
52. Devroey, E. *Le Kasai et Son Bassin Hydrographique*; Goemaere: Brussels, Belgium, 1939.
53. Yayer, J. *Caractéristiques Hydrographiques de l'Oubangui*; Comité Hydrographique du Bassin Congolais: Brussels, Belgium, 1951.
54. Moukolo, N.; Laraque, A.; Olivry, J.C.; Bricquet, J.P. Transport en solution et en suspension par le fleuve congo (zaïre) et ses principaux affluents de la rive droite. *Hydrol. Sci. J.* **1993**, *38*, 133–145. [CrossRef]
55. Censier, C. Caractérisation de processus d'érosion régressive par analyse sédimentologique comparée des sables du chenal et des barres du cours inférieur de l'Oubangui (République Centrafricaine, Congo, Zaïre). In *Proceedings of the Conference L'hydrologie Tropicale: Géoscience el Outil Pour le Développement*; AISH Publication: Wallingford, UK, 1996; Volume 238, pp. 289–303.
56. Junk, W.J.; Bayley, P.B.; Sparks, R.E. The flood pulse concept in river-floodplain-systems. *Can. J. Fish. Aquat. Sci.* **1989**, 110–127. [CrossRef]
57. Wesselink, A.J.; Orange, D.; Feizoure, C.T. Randriamiarisoa Les régimes hydroclimatiques et hydrologiques d'un bassin versant de type tropical humide: L'Oubangui (République Centrafricaine). *IAHS-AISH Publ.* **1996**, *238*, 179–194.
58. Benveniste, J.; Roca, M.; Levrini, G.; Vincent, P.; Baker, S.; Zanife, O.; Zelli, C.; Bombaci, O. The radar altimetry mission: RA-2, MWR, DORIS and LRR. *ESA Bull.* **2001**, *106*, 25101–25108.
59. Lambin, J.; Morrow, R.; Fu, L.L.; Willis, J.K.; Bonekamp, H.; Lillibridge, J.; Perbos, J.; Zaouche, G.; Vaze, P.; Bannoura, W.; et al. The OSTM/Jason-2 Mission. *Mar. Geod.* **2010**, *33*, 4–25. [CrossRef]
60. Wingham, D.J.; Rapley, C.G.; Griffiths, H. New Techniques in Satellite Altimeter Tracking Systems. *Proc. IGARSS Symp. Zurich* **1986**, 1339–1344.
61. Frappart, F.; Calmant, S.; Cauhopé, M.; Seyler, F.; Cazenave, A. Preliminary results of ENVISAT RA-2-derived water levels validation over the Amazon basin. *Remote Sens. Environ.* **2006**, *100*, 252–264. [CrossRef]
62. CTOH. Available online: http://ctoh.legos.obs-mip.fr/ (accessed on 24 October 2017).
63. Blarel, F.; Frappart, F.; Legrésy, B.; Blumstein, D.; Rémy, F.; Fatras, C.; Mougin, E.; Papa, F.; Prigent, C.; Niño, F.; et al. Altimetry backscattering signatures at Ku and S bands over land and ice sheets. In *Remote Sensing for Agriculture, Ecosystems, and Hydrology XVII*; SPIE: Bellingham, WA, USA, 2015; Volume 9637, p. 963727.
64. Frappart, F.; Legrésy, B.; Niño, F.; Blarel, F.; Fuller, N.; Fleury, S.; Birol, F.; Calmant, S. An ERS-2 altimetry reprocessing compatible with ENVISAT for long-term land and ice sheets studies. *Remote Sens. Environ.* **2016**, *184*. [CrossRef]
65. Huete, A.; Justice, C.; Liu, H. Development of vegetation and soil indices for MODIS-EOS. *Remote Sens. Environ.* **1994**, *49*, 224–234. [CrossRef]
66. Thorndike, R.L. Who belongs in the family? *Psychometrika* **1953**, *18*, 267–276. [CrossRef]
67. Crétaux, J.-F.; Nielsen, K.; Frappart, F.; Papa, F.; Calmant, S.; Benveniste, J. Hydrological applications of satellite altimetry: Rivers, lakes, man-made reservoirs, inundated areas. In *Satellite Altimetry over Oceans and Land Surfaces*; Stammer, D., Cazenave, A., Eds.; Earth Observation of Global Changes; CRC Press: Boca Raton, FL, USA, 2017; pp. 459–504.

68. MacQueen, J. Some methods for classification and analysis of multivariate observations. In *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*; University of California Press: Berkeley, CA, USA, 1967; Volume 1, pp. 281–297.
69. Arthur, D.; Vassilvitskii, S. k-means++: The advantages of careful seeding. In *Proceedings of the Eighteenth Annual ACM-SIAM Symposium on Discrete Algorithms*; SIAM: Philadelphia, PA, USA, 2007; pp. 1027–1035.
70. Rousseeuw, P.J. Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *J. Comput. Appl. Math.* **1987**, *20*, 53–65. [CrossRef]
71. Normandin, C.; Frappart, F.; Lubac, B.; Bélanger, S.; Marieu, V.; Blarel, F.; Robinet, A.; Guiastrennec-Faugas, L. Quantification of surface water volume changes in the Mackenzie Delta using satellite multi-mission data. *Hydrol. Earth Syst. Sci.* **2018**, *22*, 1543–1561. [CrossRef]
72. Becker, M.; Papa, F.; Frappart, F.; Alsdorf, D.; Calmant, S.; da Silva, J.S.; Prigent, C.; Seyler, F. Satellite-based estimates of surface water dynamics in the Congo River Basin. *Int. J. Appl. Earth Obs. Geoinf.* **2018**, *66*, 196–209. [CrossRef]
73. Fatras, C.; Borderies, P.; Frappart, F.; Mougin, E.; Blumstein, D.; Niño, F. Impact of surface soil moisture variations on radar altimetry echoes at Ku and Ka bands in semi-arid areas. *Remote Sens.* **2018**, *10*, 582. [CrossRef]
74. Bonnefond, P.; Verron, J.; Aublanc, J.; Babu, K.N.; Bergé-Nguyen, M.; Cancet, M.; Chaudhary, A.; Crétaux, J.-F.; Frappart, F.; Haines, B.; et al. The Benefits of the Ka-Band as Evidenced from the SARAL/AltiKa Altimetric Mission: Quality Assessment and Unique Characteristics of AltiKa Data. *Remote Sens.* **2018**, *10*, 83. [CrossRef]
75. Frappart, F.; Papa, F.; Güntner, A.; Tomasella, J.; Pfeffer, J.; Ramillien, G.; Emilio, T.; Schietti, J.; Seoane, L.; da Silva Carvalho, J.; et al. The spatio-temporal variability of groundwater storage in the Amazon River Basin. *Adv. Water Resour.* **2019**, *124*, 41–52. [CrossRef]

Bibliography

Abbott, B.W., Bishop, K., Zarnetske, J.P., Minaudo, C., Chapin, F.S., Krause, S., Hannah, D.M., Conner, L., Ellison, D., Godsey, S.E., Plont, S., Marçais, J., Kolbe, T., Huebner, A., Frei, R.J., Hampton, T., Gu, S., Buhman, M., Sara Sayedi, S., Ursache, O., Chapin, M., Henderson, K.D., Pinay, G., 2019. Human domination of the global water cycle absent from depictions and perceptions. Nat. Geosci. 12, 533–540. doi:10.1038/s41561-019-0374-y.

Abdalla, S., Kolahchi, A.A., Ablain, M., Adusumilli, S., Bhowmick, S.A., Alou-Font, E., Amarouche, L., Andersen, O.B., Antich, H., Aouf, L., et al., 2021. Altimetry for the future: Building on 25 years of progress. Advances in Space Research 68, 319–363.

Aires, F., Miolane, L., Prigent, C., Pham, B., Fluet-Chouinard, E., Lehner, B., Papa, F., 2017. A global dynamic long-term inundation extent dataset at high spatial resolution derived through downscaling of satellite observations. Journal of Hydrometeorology 18, 1305–1325.

Al-Khaldi, M.M., Johnson, J.T., Gleason, S., Chew, C.C., Gerlein-Safdi, C., Shah, R., Zuffada, C., 2021a. Inland water body mapping using cygnss coherence detection. IEEE Transactions on Geoscience and Remote Sensing 59, 7385–7394. doi:10.1109/TGRS.2020.3047075.

Al-Khaldi, M.M., Johnson, J.T., Gleason, S., Loria, E., O'Brien, A.J., Yi, Y., 2021b. An algorithm for detecting coherence in cyclone global navigation satellite system mission level-1 delay-doppler maps. IEEE Transactions on Geoscience and Remote Sensing 59, 4454–4463. doi:10.1109/TGRS.2020.3009784.

Al-Khaldi, M.M., Johnson, J.T., O'Brien, A.J., Balenzano, A., Mattia, F., 2019. Time-Series Retrieval of Soil Moisture Using CYGNSS. IEEE Trans. Geosci. Remote Sensing 57, 10.

Almazroui, M., Islam, M.N., Saeed, F., Saeed, S., Ismail, M., Ehsan, M.A., Diallo, I., O'Brien, E., Ashfaq, M., Martínez-Castro, D., Cavazos, T., Cerezo-Mota, R., Tippett, M.K., Gutowski, W.J., Alfaro, E.J., Hidalgo, H.G., Vichot-Llano, A., Campbell, J.D., Kamil, S., Rashid, I.U., Sylla, M.B., Stephenson, T., Taylor, M., Barlow, M., 2021. Projected Changes in Temperature and Precipitation Over the United States, Central America, and the Caribbean in CMIP6 GCMs. Earth Syst Environ 5, 1–24. doi:10.1007/s41748-021-00199-5.

Almazroui, M., Saeed, F., Saeed, S., Nazrul Islam, M., Ismail, M., Klutse, N.A.B., Siddiqui, M.H., 2020a. Projected Change in Temperature and Precipitation Over Africa from CMIP6. Earth Syst Environ 4, 455–475. doi:10.1007/s41748-020-00161-x.

Almazroui, M., Saeed, S., Saeed, F., Islam, M.N., Ismail, M., 2020b. Projections of Precipitation and Temperature over the South Asian Countries in CMIP6. Earth Syst Environ 4, 297–320. doi:10.1007/s41748-020-00157-7.

Alonso-Arroyo, A., Camps, A., Aguasca, A., Forte, G.F., Monerris, A., Rüdiger, C., Walker, J.P., Park, H., Pascual, D., Onrubia, R., 2014a. Dual-polarization gnss-r interference pattern technique for soil moisture mapping. IEEE journal of selected topics in applied earth observations and remote sensing 7, 1533–1544.

Alonso-Arroyo, A., Camps, A., Park, H., Pascual, D., Onrubia, R., Martín, F., 2014b. Retrieval of significant wave height and mean sea surface level using the gnss-r interference pattern technique: Results from a three-month field campaign. IEEE Transactions on Geoscience and Remote Sensing 53, 3198–3209.

Alonso-Arroyo, A., Camps, A., Park, H., Pascual, D., Onrubia, R., Martín, F., 2015. Retrieval of significant wave height and mean sea surface level using the gnss-r interference pattern technique: Results from a three-month field campaign. IEEE Transactions on Geoscience and Remote Sensing 53, 3198–3209. doi:10.1109/TGRS.2014.2371540.

Alonso-Arroyo, A., Zavorotny, V.U., Camps, A., 2017. Sea ice detection using uk tds-1 gnss-r data. IEEE Transactions on Geoscience and Remote Sensing 55, 4989–5001.

Alsdorf, D., Bates, P., Melack, J., Wilson, M., Dunne, T., 2007. Spatial and temporal complexity of the amazon flood measured from space. Geophysical research letters 34. doi:10.1029/2007GL029447.

Alsdorf, D., Birkett, C., Dunne, T., Melack, J., Hess, L., 2001a. Water level changes in a large amazon lake measured with spaceborne radar interferometry and altimetry. Geophysical Research Letters 28, 2671–2674. doi:10.1029/2001GL012962.

Alsdorf, D., Smith, L., Melack, J., 2001b. Amazon floodplain water level changes measured with interferometric sir-c radar. IEEE Transactions on Geoscience and Remote Sensing 39, 423–431. doi:10.1109/36.905250.

Alsdorf, D.E., Lettenmaier, D.P., 2003. Tracking fresh water from space. Science 301, 1491–1494. doi:10.1126/science.108980.

Alsdorf, D.E., Melack, J.M., Dunne, T., Mertes, L.A., Hess, L.L., Smith, L.C., 2000. Interferometric radar measurements of water level changes on the amazon flood plain. Nature 404, 174–177. doi:10.1038/35004560.

Anderson, K.D., 2000. Determination of water level and tides using interferometric observations of gps signals. Journal of Atmospheric and Oceanic Technology 17, 1118–1127.

Andrews, T., Forster, P.M., Boucher, O., Bellouin, N., Jones, A., 2010. Precipitation, radiative forcing and global temperature change: PRECIPITATION AND RADIATIVE FORCING. Geophys. Res. Lett. 37, n/a–n/a. doi:10.1029/2010GL043991.

Asoka, A., Gleeson, T., Wada, Y., Mishra, V., 2017. Relative contribution of monsoon precipitation and pumping to changes in groundwater storage in India. Nature Geosci 10, 109–117. doi:10.1038/ngeo2869.

Attema, E., Ulaby, F.T., 1978. Vegetation modeled as a water cloud. Radio science 13, 357–364.

Aubert, J.C., Bibaut, A., Rigal, J.M., 1994. Characterization of multipath on land and sea at gps frequencies, in: Proceedings of the 7th international technical meeting of the satellite division of the institute of navigation (ION GPS 1994), pp. 1155–1171.

Bai, Y., Wu, W., Yang, Z., Yu, J., Zhao, B., Liu, X., Yang, H., Mas, E., Koshimura, S., 2021. Enhancement of detecting permanent water and temporary water in flood disasters by fusing sentinel-1 and sentinel-2 imagery using deep learning algorithms: Demonstration of sen1floods11 benchmark datasets. Remote Sensing 13, 2220. doi:10.3390/rs13112220.

Bamler, R., Hartl, P., 1998. Synthetic aperture radar interferometry. Inverse problems 14, R1. doi:10.1088/0266-5611/14/4/001.

Bauer-Marschallinger, B., Cao, S., Tupas, M.E., Roth, F., Navacchi, C., Melzer, T., Freeman, V., Wagner, W., 2022. Satellite-based flood mapping through bayesian inference from a sentinel-1 sar datacube. Remote Sensing 14, 3673. doi:10.3390/rs14153673.

Becker, M., Papa, F., Frappart, F., Alsdorf, D., Calmant, S., da Silva, J.S., Prigent, C., Seyler, F., 2018. Satellite-based estimates of surface water dynamics in the congo river basin. International Journal of Applied Earth Observation and Geoinformation 66, 196–209. doi:10.1016/j.jag.2017.11.015.

Beckheinrich, J., Hirrle, A., Schön, S., Beyerle, G., Semmling, M., Wickert, J., 2014. Water level monitoring of the mekong delta using gnss reflectometry technique, in: 2014 IEEE Geoscience and Remote Sensing Symposium, IEEE. pp. 3798–3801.

Beckmann, P., Spizzichino, A., 1987. The scattering of electromagnetic waves from rough surfaces. Norwood .

Bennett, A.C., Dargie, G.C., Cuni-Sanchez, A., Tshibamba Mukendi, J., Hubau, W., Mukinzi, J.M., Phillips, O.L., Malhi, Y., Sullivan, M.J., Cooper, D.L., et al., 2021. Resistance of african tropical forests to an extreme climate anomaly. Proceedings of the National Academy of Sciences 118, e2003169118. doi:10.1073/pnas.20031691.

Berg, A., Findell, K., Lintner, B., Giannini, A., Seneviratne, S.I., van den Hurk, B., Lorenz, R., Pitman, A., Hagemann, S., Meier, A., Cheruy, F., Ducharne, A., Malyshev, S., Milly, P.C.D., 2016. Land–atmosphere feedbacks amplify aridity increase over land under global warming. Nature Clim Change 6, 869–874. doi:10.1038/nclimate3029.

Berg, A., Sheffield, J., Milly, P.C.D., 2017. Divergent surface and total soil moisture projections under global warming. Geophys. Res. Lett. 44, 236–244. doi:10.1002/2016GL071921.

Berghuijs, W.R., Woods, R.A., Hrachowitz, M., 2014. A precipitation shift from snow towards rain leads to a decrease in streamflow. Nature Clim Change 4, 583–586. doi:10.1038/nclimate2246.

Betbeder, J., Gond, V., Frappart, F., Baghdadi, N.N., Briant, G., Bartholome, E., 2014. Mapping of Central Africa Forested Wetlands Using Remote Sensing. IEEE J. Sel. Top. Appl. Earth Observations Remote Sensing 7, 531–542. doi:10.1109/JSTARS.2013.2269733.

Biancamaria, S., Lettenmaier, D.P., Pavelsky, T.M., 2016. The swot mission and its capabilities for land hydrology. Surveys in Geophysics 37, 307–337. doi:10.1007/s10712-015-9346-y.

Bierkens, M.F.P., Wada, Y., 2019. Non-renewable groundwater use and groundwater depletion: a review. Environ. Res. Lett. 14, 063002. doi:10.1088/1748-9326/ab1a5f.

Bilich, A., Larson, K.M., Axelrad, P., 2008. Modeling gps phase multipath with snr: Case study from the salar de uyuni, boliva. Journal of Geophysical Research: Solid Earth 113. doi:10.1029/2007JB005194.

Bioresita, F., Puissant, A., Stumpf, A., Malet, J.P., 2018. A method for automatic and rapid mapping of water surfaces from sentinel-1 imagery. Remote Sensing 10, 217. doi:10.3390/rs10020217.

Bolanos, S., Stiff, D., Brisco, B., Pietroniro, A., 2016. Operational surface water detection and monitoring using radarsat 2. Remote Sensing 8, 285. doi:doi:10.3390/rs8040285.

Bolch, T., Menounos, B., Wheate, R., 2010. Landsat-based inventory of glaciers in western Canada, 1985–2005. Remote Sensing of Environment 114, 127–137. doi:10.1016/j.rse.2009.08.015.

Bonneton, P., Bonneton, N., Parisot, J.P., Castelle, B., 2015. Tidal bore dynamics in funnel-shaped estuaries. Journal of Geophysical Research: Oceans 120, 923–941.

Boysen, L.R., Brovkin, V., Pongratz, J., Lawrence, D.M., Lawrence, P., Vuichard, N., Peylin, P., Liddicoat, S., Hajima, T., Zhang, Y., Rocher, M., Delire, C., Séférian, R., Arora, V.K., Nieradzik, L., Anthoni, P., Thiery, W., Laguë, M.M., Lawrence, D., Lo, M.H., 2020. Global climate response to idealized deforestation in CMIP6 models. Biogeosciences 17, 5615–5638. doi:10.5194/bg-17-5615-2020.

Bridgham, S.D., Cadillo-Quiroz, H., Keller, J.K., Zhuang, Q., 2013. Methane emissions from wetlands: biogeochemical, microbial, and modeling perspectives from local to global scales. Glob Change Biol 19, 1325–1346. doi:10.1111/gcb.12131.

Caballero, I., Ruiz, J., Navarro, G., 2019. Sentinel-2 satellites provide near-real time evaluation of catastrophic floods in the west mediterranean. Water 11, 2499. doi:10.3390/w11122499.

Camps, A., 2019. Spatial resolution in gnss-r under coherent scattering. IEEE Geoscience and Remote Sensing Letters 17, 32–36.

Camps, A., Golkar, A., Gutierrez, A., De Azua, J.R., Munoz-Martin, J.F., Fernandez, L., Diez, C., Aguilella, A., Briatore, S., Akhtyamov, R., et al., 2018. Fsscat, the 2017 copernicus masters'"esa sentinel small satellite challenge" winner: A federated polar and soil moisture tandem mission based on 6u cubesats, in: IGARSS 2018-2018 IEEE International Geoscience and Remote Sensing Symposium, IEEE. pp. 8285–8287.

Camps, A., Munoz-Martin, J.F., 2020. Analytical computation of the spatial resolution in gnss-r and experimental validation at l1 and l5. Remote Sensing 12, 3910.

Camps, A., Park, H., i Domènech, E.V., Pascual, D., Martin, F., Rius, A., Ribo, S., Benito, J., Andrés-Beivide, A., Saameno, P., et al., 2014. Optimization and performance analysis of interferometric gnss-r altimeters: Application to the paris iod mission. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 7, 1436–1451.

Camps, A., Park, H., Pablos, M., Foti, G., Gommenginger, C.P., Liu, P.W., Judge, J., 2016. Sensitivity of GNSS-R Spaceborne Observations to Soil Moisture and Vegetation. IEEE J. Sel. Top. Appl. Earth Observations Remote Sensing 9, 4730–4742. doi:10.1109/JSTARS.2016.2588467.

Canadell, J., Monteiro, P., Costa, M., Cotrim da Cunha, L., Cox, P., Eliseev, A., Henson, S., Ishii, M., Jaccard, S., Koven, C., Lohila, A., Patra, P., Piao, S., Rogelj, J., Syampungani, S., Zaehle, S., Zickfeld, K., 2021. Global carbon and other biogeochemical cycles and feedbacks, in: Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, p. 673–816. doi:10.1017/9781009157896.007.

Cao, L., Bala, G., Caldeira, K., 2012. Climate response to changes in atmospheric carbon dioxide and solar irradiance on the time scale of days to weeks. Environ. Res. Lett. 7, 034015. doi:10.1088/1748-9326/7/3/034015.

Cao, N., Lee, H., Jung, H.C., Yu, H., 2018. Estimation of water level changes of large-scale amazon wetlands using alos2 scansar differential interferometry. Remote Sensing 10, 966. doi:10.3390/rs10060966.

Cardellach, E., Ao, C., De la Torre Juárez, M., Hajj, G., 2004. Carrier phase delay altimetry with gps-reflection/occultation interferometry from low earth orbiters. Geophysical Research Letters 31.

Cardellach, E., Fabra, F., Nogués-Correig, O., Oliveras, S., Ribó, S., Rius, A., 2011. GNSS-R ground-based and airborne campaigns for ocean, land, ice, and snow techniques: Application to the GOLD-RTR data sets. Radio Science 46, 1–16. doi:10.1029/2011RS004683.

Cardellach, E., Ruffini, G., Pino, D., Rius, A., Komjathy, A., Garrison, J.L., 2003. Mediterranean balloon experiment: Ocean wind speed sensing from the stratosphere, using gps reflections. Remote Sensing of Environment 88, 351–362.

Carreno-Luengo, H., Camps, A., Via, P., Munoz, J.F., Cortiella, A., Vidal, D., Jané, J., Catarino, N., Hagenfeldt, M., Palomo, P., et al., 2016. 3cat-2—an experimental nanosatellite for gnss-r earth observation: Mission concept and analysis. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 9, 4540–4551.

Carreno-Luengo, H., Luzi, G., Crosetto, M., 2019. Sensitivity of CyGNSS Bistatic Reflectivity and SMAP Microwave Radiometry Brightness Temperature to Geophysical Parameters over Land Surfaces. IEEE J. Sel. Top. Appl. Earth Observations Remote Sensing 12, 107–122. doi:10.1109/JSTARS.2018.2856588. publisher: IEEE.

Carreno-Luengo, H., Luzi, G., Crosetto, M., 2020. Above-ground biomass retrieval over tropical forests: A novel gnss-r approach with cygnss. Remote Sensing 12, 1368.

Carroll, M.L., Townshend, J.R., DiMiceli, C.M., Noojipady, P., Sohlberg, R.A., 2009. A new global raster water mask at 250 m resolution. International Journal of Digital Earth 2, 291–308. doi:10.1080/17538940902951401.

Chahine, M.T., 1992. The hydrological cycle and its influence on climate. Nature 359, 373–380.

Chan, S.C., Kendon, E.J., Roberts, N.M., Fowler, H.J., Blenkinsop, S., 2016a. The characteristics of summer sub-hourly rainfall over the southern UK in a high-resolution convective permitting model. Environ. Res. Lett. 11, 094024. doi:10.1088/1748-9326/11/9/094024.

Chan, S.K., Bindlish, R., O'Neill, P., Jackson, T., Njoku, E., Dunbar, S., Chaubell, J., Piepmeier, J., Yueh, S., Entekhabi, D., et al., 2018. Development and assessment of the smap enhanced passive soil moisture product. Remote Sensing of Environment 204, 931–941.

Chan, S.K., Bindlish, R., O'Neill, P.E., Njoku, E., Jackson, T., Colliander, A., Chen, F., Burgin, M., Dunbar, S., Piepmeier, J., et al., 2016b. Assessment of the smap passive soil moisture product. IEEE Transactions on Geoscience and Remote Sensing 54, 4994–5007. doi:10.1109/TGRS.2016.2561938.

Chapman, B.D., Russo, I.M., Galdi, C., Morris, M., di Bisceglie, M., Zuffada, C., Downs, B., Lavalle, M., Loria, E., O'Brien, A.J., 2022. Comparison of sar and cygnss surface water extent metrics. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 15, 3235–3245. doi:10.1109/JSTARS.2022.3162764.

Chen, Y., Huang, C., Ticehurst, C., Merrin, L., Thew, P., 2013. An Evaluation of MODIS Daily and 8-day Composite Products for Floodplain and Wetland Inundation Mapping. Wetlands 33, 823–835. doi:10.1007/s13157-013-0439-4.

Cheng, D.K., 1993. Fundamentals of engineering electromagnetics .

Chew, C., 2021. Spatial interpolation based on previously-observed behavior: A framework for interpolating spaceborne gnss-r data from cygnss. Journal of Spatial Science , 1–14.

Chew, C., Reager, J.T., Small, E., 2018. CYGNSS data map flood inundation during the 2017 Atlantic hurricane season. Sci Rep 8, 9336. doi:10.1038/s41598-018-27673-x.

Chew, C., Shah, R., Zuffada, C., Hajj, G., Masters, D., Mannucci, A.J., 2016. Demonstrating soil moisture remote sensing with observations from the UK TechDemoSat-1 satellite mission. Geophysical Research Letters 43, 3317–3324. doi:10.1002/2016GL068189.

Chew, C., Small, E., 2020a. Description of the ucar/cu soil moisture product. Remote Sensing 12, 1558. doi:doi:10.3390/rs12101558.

Chew, C., Small, E., 2020b. Estimating inundation extent using CYGNSS data: A conceptual modeling study. Remote Sensing of Environment 246, 111869. doi:10.1016/j.rse.2020.111869.

Chew, C.C., Small, E.E., 2018. Soil Moisture Sensing Using Spaceborne GNSS Reflections: Comparison of CYGNSS Reflectivity to SMAP Soil Moisture. Geophysical Research Letters 45, 4049–4057. doi:10.1029/2018GL077905.

Chew, C.C., Small, E.E., Larson, K.M., Zavorotny, V.U., 2013. Effects of near-surface soil moisture on gps snr data: Development of a retrieval algorithm for soil moisture. IEEE Transactions on Geoscience and Remote Sensing 52, 537–543.

Chini, M., Hostache, R., Giustarini, L., Matgen, P., 2017. A hierarchical split-based approach for parametric thresholding of sar images: Flood inundation as a test case. IEEE Transactions on Geoscience and Remote Sensing 55, 6975–6988. doi:10.1109/TGRS.2017.2737664.

Ciais, P., Sabine, C., Bala, G., Bopp, L., Brovkin, V., Canadell, J., Chhabra, A., DeFries, R., Galloway, J., Heimann, M., Jones, C., Le Quéré, C., Myneni, R., S, P., Thornton, P., 2013. Carbon and other biogeochemical cycles, in: Climate Change 2013 – The Physical Science Basis: Working Group I Contribution to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press. doi:10.1017/CBO9781107415324.015.

Clarizia, M., Gommenginger, C., Gleason, S., Srokosz, M., Galdi, C., Di Bisceglie, M., 2009. Analysis of gnss-r delay-doppler maps from the uk-dmc satellite over the ocean. Geophysical Research Letters 36.

Clarizia, M.P., Pierdicca, N., Costantini, F., Floury, N., 2019. Analysis of CYGNSS Data for Soil Moisture Retrieval. IEEE J. Sel. Top. Appl. Earth Observations Remote Sensing 12, 9.

Clarizia, M.P., Ruf, C.S., 2016. Wind speed retrieval algorithm for the cyclone global navigation satellite system (cygnss) mission. IEEE Transactions on Geoscience and Remote Sensing 54, 4419–4432. doi:10.1109/TGRS.2016.2541343.

Clarizia, M.P., Ruf, C.S., Cipollini, P., Zuffada, C., 2016. First spaceborne observation of sea surface height using GPS-Reflectometry. Geophysical Research Letters 43, 767–774. doi:10.1002/2015GL066624.

Collett, I., Wang, Y., Shah, R., Morton, Y.J., 2022. Phase coherence of gps signal land reflections and its dependence on surface characteristics. IEEE Geoscience and Remote Sensing Letters 19, 1–5. doi:10.1109/LGRS.2021.3094407.

Cook, B.I., Mankin, J.S., Marvel, K., Williams, A.P., Smerdon, J.E., Anchukaitis, K.J., 2020. Twenty-first century drought projections in the cmip6 forcing scenarios. Earth's Future 8. doi:10.1029/2019EF001461.

Covey, K.R., Megonigal, J.P., 2019. Methane production and emissions in trees and forests. New Phytol 222, 35–51. doi:10.1111/nph.15624.

Cox, P.M., Pearson, D., Booth, B.B., Friedlingstein, P., Huntingford, C., Jones, C.D., Luke, C.M., 2013. Sensitivity of tropical carbon to climate change constrained by carbon dioxide variability. Nature 494, 341–344. doi:10.1038/nature11882.

Cretaux, J.F., Nielsen, K., Frappart, F., Papa, F., Calmant, S., Benveniste, J., 2017. Hydrological applications of satellite altimetryrivers, lakes, man-made reservoirs, inundated areas, in: Satellite altimetry over oceans and land surfaces. CRC Press, pp. 459–504.

Crétaux, J.F., Arsen, A., Calmant, S., Kouraev, A., Vuglinski, V., Bergé-Nguyen, M., Gennero, M.C., Nino, F., Rio, R.A.D., Cazenave, A., Maisongrande, P., 2011. SOLS: A lake database to monitor in the Near Real Time water level and storage variations from remote sensing data. Advances in Space Research 47, 1497–1507. doi:https://doi.org/10.1016/j.asr.2011.01.004.

Cuthbert, M.O., Gleeson, T., Moosdorf, N., Befus, K.M., Schneider, A., Hartmann, J., Lehner, B., 2019. Global patterns and dynamics of climate–groundwater interactions. Nature Clim Change 9, 137–141. doi:10.1038/s41558-018-0386-4.

CYGNSS, 2020. CYGNSS level 1 science data record version 3.0. ver. 3.0. PO.DAAC, CA, USA. URL: https://doi.org/10.5067/CYGNS-L1X30. accessed: 2021-10-01.

Dagan, G., Stier, P., 2020. Constraint on precipitation response to climate change by combination of atmospheric energy and water budgets. npj Clim Atmos Sci 3, 34. doi:10.1038/s41612-020-00137-8.

Dagan, G., Stier, P., Watson-Parris, D., 2019. Analysis of the Atmospheric Water Budget for Elucidating the Spatial Scale of Precipitation Changes Under Climate Change. Geophys. Res. Lett. 46, 10504–10511. doi:10.1029/2019GL084173.

DAHITI, . Water height time series available online:. URL: https://dahiti.dgfi.tum.de/en. (accessed on 31 August 2022).

Darrozes, J., Roussel, N., Zribi, M., 2016. The reflected global navigation satellite system (gnss-r): From theory to practice, in: Microwave remote sensing of land surface. Elsevier, pp. 303–355.

De Kauwe, M.G., Medlyn, B.E., Zaehle, S., Walker, A.P., Dietze, M.C., Hickler, T., Jain, A.K., Luo, Y., Parton, W.J., Prentice, I.C., Smith, B., Thornton, P.E., Wang, S., Wang, Y.P., Wårlind, D., Weng, E., Crous, K.Y., Ellsworth, D.S., Hanson, P.J., Seok Kim, H., Warren, J.M., Oren, R., Norby, R.J., 2013. Forest water use and water use efficiency at elevated CO $_2$: a model-data intercomparison at two contrasting temperate forest FACE sites. Glob Change Biol 19, 1759–1779. doi:10.1111/gcb.12164.

De Roo, R.D., Ulaby, F.T., 1994. Bistatic specular scattering from rough dielectric surfaces. IEEE Transactions on Antennas and Propagation 42, 220–231.

DeVries, B., Huang, C., Armston, J., Huang, W., Jones, J.W., Lang, M.W., 2020. Rapid and robust monitoring of flood events using sentinel-1 and landsat data on the google earth engine. Remote Sensing of Environment 240, 111664. doi:10.1016/j.rse.2020.111664.

Di Vittorio, C.A., Georgakakos, A.P., 2018. Land cover classification and wetland inundation mapping using MODIS. Remote Sensing of Environment 204, 1–17. doi:10.1016/j.rse.2017.11.001.

D'Odorico, P., Davis, K.F., Rosa, L., Carr, J.A., Chiarelli, D., Dell'Angelo, J., Gephart, J., MacDonald, G.K., Seekell, D.A., Suweis, S., Rulli, M.C., 2018. The global food-energy-water nexus. Reviews of Geophysics 56, 456–531. doi:10.1029/2017RG000591.

Donchyts, G., Baart, F., Winsemius, H., Gorelick, N., Kwadijk, J., van de Giesen, N., 2016. Earth's surface water change over the past 30 years. Nature Clim Change 6, 810–813. doi:10.1038/nclimate3111.

Dong, B., Dai, A., 2017. The uncertainties and causes of the recent changes in global evapotranspiration from 1982 to 2010. Clim Dyn 49, 279–296. doi:10.1007/s00382-016-3342-x.

Douville, H., Raghavan, K., Renwick, J., Allan, R., Arias, P., Barlow, M., Cerezo-Mota, R., Cherchi, A., Gan, T., Gergis, J., Jiang, D., Khan, A., Pokam Mba, W., Rosenfeld, D., Tierney, J., Zolina, O., 2021. Water cycle changes, in: Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, p. 1055–1210. doi:10.1017/9781009157896.010.

Downing, J.A., 2010. Emerging global role of small lakes and ponds: little things mean a lot. Limnetica 29, 0009–24.

Downing, J.A., Prairie, Y., Cole, J., Duarte, C., Tranvik, L., Striegl, R.G., McDowell, W., Kortelainen, P., Caraco, N., Melack, J., et al., 2006. The global abundance and size distribution of lakes, ponds, and impoundments. Limnology and Oceanography 51, 2388–2397.

Dudley, R., Hodgkins, G., McHale, M., Kolian, M., Renard, B., 2017. Trends in snowmelt-related streamflow timing in the conterminous United States. Journal of Hydrology 547, 208–221. doi:10.1016/j.jhydrol.2017.01.051.

Döll, P., Müller Schmied, H., Schuh, C., Portmann, F.T., Eicker, A., 2014. Global-scale assessment of groundwater depletion and related groundwater abstractions: Combining hydrological modeling with information from well observations and grace satellites. Water Resources Research 50, 5698–5720. doi:10.1002/2014WR015595.

Eekhout, J.P.C., Hunink, J.E., Terink, W., de Vente, J., 2018. Why increased extreme precipitation under climate change negatively affects water security. Hydrol. Earth Syst. Sci. 22, 5935–5946. doi:10.5194/hess-22-5935-2018.

Egido, A., Paloscia, S., Motte, E., Guerriero, L., Pierdicca, N., Caparrini, M., Santi, E., Fontanelli, G., Floury, N., 2014. Airborne GNSS-R polarimetric measurements for soil moisture and above-ground biomass estimation. IEEE J. Sel. Top. Appl. Earth Observations Remote Sensing 7, 1522–1532. doi:10.1109/JSTARS.2014.2322854. publisher: IEEE.

Enge, P.K., 1994. The global positioning system: Signals, measurements, and performance. International Journal of Wireless Information Networks 1, 83–105.

Entekhabi, D., Njoku, E.G., O'Neill, P.E., Kellogg, K.H., Crow, W.T., Edelstein, W.N., Entin, J.K., Goodman, S.D., Jackson, T.J., Johnson, J., et al., 2010. The soil moisture

active passive (smap) mission. Proceedings of the IEEE 98, 704–716. doi:10.1109/JPROC.2010.2043918.

Eroglu, O., Kurum, M., Boyd, D., Gurbuz, A.C., 2019. High spatio-temporal resolution cygnss soil moisture estimates using artificial neural networks. Remote Sensing 11. doi:10.3390/rs11192272.

ESA, 2017. Land cover CCI product user guide version 2. tech. rep. (2017). URL: maps.elie.ucl.ac.be/CCI/viewer/download/ESACCI-LC-Ph2-PUGv2_2.0.pdf.

Etminan, M., Myhre, G., Highwood, E., Shine, K., 2016. Radiative forcing of carbon dioxide, methane, and nitrous oxide: A significant revision of the methane radiative forcing. Geophysical Research Letters 43, 12–614. doi:10.1002/2016GL071930.

Famiglietti, J.S., 2014. The global groundwater crisis. Nature Climate Change 4, 945–948.

Fang, Y., Michalak, A.M., Schwalm, C.R., Huntzinger, D.N., Berry, J.A., Ciais, P., Piao, S., Poulter, B., Fisher, J.B., Cook, R.B., Hayes, D., Huang, M., Ito, A., Jain, A., Lei, H., Lu, C., Mao, J., Parazoo, N.C., Peng, S., Ricciuto, D.M., Shi, X., Tao, B., Tian, H., Wang, W., Wei, Y., Yang, J., 2017. Global land carbon sink response to temperature and precipitation varies with ENSO phase. Environ. Res. Lett. 12, 064007. doi:10.1088/1748-9326/aa6e8e.

Fatras, C., Frappart, F., Mougin, E., Frison, P.L., Faye, G., Borderies, P., Jarlan, L., 2015. Spaceborne altimetry and scatterometry backscattering signatures at c-and ku-bands over west africa. Remote Sensing of Environment 159, 117–133.

Feng, M., Sexton, J.O., Channan, S., Townshend, J.R., 2016. A global, high-resolution (30-m) inland water body dataset for 2000: First results of a topographic–spectral classification algorithm. International Journal of Digital Earth 9, 113–133. doi:10.1080/17538947.2015.1026420.

Fily, M., Royer, A., Goıta, K., Prigent, C., 2003. A simple retrieval method for land surface temperature and fraction of water surface determination from satellite microwave brightness temperatures in sub-arctic areas. Remote Sensing of Environment 85, 328–338.

Fjørtoft, R., Gaudin, J.M., Pourthie, N., Lalaurie, J.C., Mallet, A., Nouvel, J.F., Martinot-Lagarde, J., Oriot, H., Borderies, P., Ruiz, C., et al., 2013. Karin on swot: Characteristics of near-nadir ka-band interferometric sar imagery. IEEE Transactions on Geoscience and Remote Sensing 52, 2172–2185. doi:10.1109/TGRS.2013.2258402.

Fleischmann, A., Siqueira, V., Paris, A., Collischonn, W., Paiva, R., Pontes, P., Crétaux, J.F., Bergé-Nguyen, M., Biancamaria, S., Gosset, M., et al., 2018. Modelling hydrologic and hydrodynamic processes in basins with large semi-arid wetlands. Journal of Hydrology 561, 943–959.

Fluet-Chouinard, E., Lehner, B., Rebelo, L.M., Papa, F., Hamilton, S.K., 2015. Development of a global inundation map at high spatial resolution from topographic downscaling of coarse-scale remote sensing data. Remote Sensing of Environment 158, 348–361. doi:https://doi.org/10.1016/j.rse.2014.10.015.

Fossey, M., Rousseau, A.N., Savary, S., 2016. Assessment of the impact of spatio-temporal attributes of wetlands on stream flows using a hydrological modelling framework: a theoretical case study of a watershed under temperate climatic conditions. Hydrological Processes 30, 1768–1781.

Foster, J.L., Sun, C., Walker, J.P., Kelly, R., Chang, A., Dong, J., Powell, H., 2005. Quantifying the uncertainty in passive microwave snow water equivalent observations. Remote Sensing of environment 94, 187–203.

Foti, G., Gommenginger, C., Jales, P., Unwin, M., Shaw, A., Robertson, C., Roselló, J., 2015. Spaceborne GNSS reflectometry for ocean winds: First results from the UK TechDemoSat-1 mission. Geophys. Res. Lett. 42, 5435–5441. doi:10.1002/2015GL064204.

Fox-Kemper, B., Hewitt, H., Xiao, C., Aðalgeirsdóttir, G., Drijfhout, S., Edwards, T., Golledge, N., Hemer, M., Kopp, R., Krinner, G., Mix, A., Notz, D., Nowicki, S., Nurhati, I., Ruiz, L., Sallée, J.B., Slangen, A., Yu, Y., 2021. 2021: Ocean, cryosphere and sea level change, in: Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, p. 1211–1362. doi:10.1017/9781009157896.011.

Frank, D.C., Poulter, B., Saurer, M., Esper, J., Huntingford, C., Helle, G., Treydte, K., Zimmermann, N.E., Schleser, G., Ahlström, A., Ciais, P., Friedlingstein, P., Levis, S., Lomas, M., Sitch, S., Viovy, N., Andreu-Hayles, L., Bednarz, Z., Berninger, F., Boettger, T., D'Alessandro, C.M., Daux, V., Filot, M., Grabner, M., Gutierrez, E., Haupt, M., Hilasvuori, E., Jungner, H., Kalela-Brundin, M., Krapiec, M., Leuenberger, M., Loader, N.J., Marah, H., Masson-Delmotte, V., Pazdur, A., Pawelczyk, S., Pierre, M., Planells, O., Pukiene, R., Reynolds-Henne, C.E., Rinne, K.T., Saracino, A., Sonninen, E., Stievenard, M., Switsur, V.R., Szczepanek, M., Szychowska-Krapiec, E., Todaro, L., Waterhouse, J., Weigl, M., 2015. Water-use efficiency and transpiration across European forests during the Anthropocene. Nature Clim Change 5, 579–583. doi:10.1038/nclimate2614.

Frappart, F., Biancamaria, S., Normandin, C., Blarel, F., Bourrel, L., Aumont, M., Azemar, P., Vu, P.L., Le Toan, T., Lubac, B., Darrozes, J., 2018. Influence of recent climatic events on the surface water storage of the Tonle Sap Lake. Science of The Total Environment 636, 1520–1533. doi:10.1016/j.scitotenv.2018.04.326.

Frappart, F., Blarel, F., Papa, F., Prigent, C., Mougin, É., Paillou, P., Baup, F., Zeiger, P., Salameh, E., Darrozes, J., et al., 2021a. Backscattering signatures at ka, ku, c and s bands from low resolution radar altimetry over land. Advances in Space Research 68, 989–1012.

Frappart, F., Minh, K.D., l'Hermitte, J., Cazenave, A., Ramillien, G., Le Toan, T., Mognard-Campbell, N., 2006. Water volume change in the lower mekong from satellite altimetry and imagery data. Geophysical Journal International 167, 570–584. doi:10.1111/j.1365-246X.2006.03184.x.

Frappart, F., Papa, F., Famiglietti, J.S., Prigent, C., Rossow, W.B., Seyler, F., 2008. Interannual variations of river water storage from a multiple satellite approach: A case

study for the rio negro river basin. Journal of Geophysical Research: Atmospheres 113. doi:10.1029/2007JD009438.

Frappart, F., Papa, F., Güntner, A., Werth, S., Da Silva, J.S., Tomasella, J., Seyler, F., Prigent, C., Rossow, W.B., Calmant, S., et al., 2011. Satellite-based estimates of groundwater storage variations in large drainage basins with extensive floodplains. Remote Sensing of Environment 115, 1588–1594. doi:10.1016/j.rse.2011.02.003.

Frappart, F., Papa, F., Malbeteau, Y., León, J.G., Ramillien, G., Prigent, C., Seoane, L., Seyler, F., Calmant, S., 2014. Surface freshwater storage variations in the orinoco floodplains using multi-satellite observations. Remote Sensing 7, 89–110. doi:10.3390/rs70100089.

Frappart, F., Papa, F., da Silva, J.S., Ramillien, G., Prigent, C., Seyler, F., Calmant, S., 2012. Surface freshwater storage and dynamics in the amazon basin during the 2005 exceptional drought. Environmental Research Letters 7, 044010. doi:10.1088/1748-9326/7/4/044010.

Frappart, F., Roussel, N., Darrozes, J., Bonneton, P., Bonneton, N., Detandt, G., Perosanz, F., Loyer, S., 2016. High rate gnss measurements for detecting non-hydrostatic surface wave. application to tidal borein the garonne river. European Journal of Remote Sensing 49, 917–932.

Frappart, F., Seyler, F., Martinez, J.M., León, J.G., Cazenave, A., 2005. Floodplain water storage in the negro river basin estimated from microwave remote sensing of inundation area and water levels. Remote Sensing of Environment 99, 387–399. doi:10.1016/j.rse.2005.08.016.

Frappart, F., Wigneron, J.P., Li, X., Liu, X., Al-Yaari, A., Fan, L., Wang, M., Moisy, C., Le Masson, E., Aoulad Lafkih, Z., et al., 2020. Global monitoring of the vegetation dynamics from the vegetation optical depth (vod): A review. Remote Sensing 12, 2915. doi:10.3390/rs12182915.

Frappart, F., Zeiger, P., Betbeder, J., Gond, V., Bellot, R., Baghdadi, N., Blarel, F., Darrozes, J., Bourrel, L., Seyler, F., 2021b. Automatic Detection of Inland Water Bodies along Altimetry Tracks for Estimating Surface Water Storage Variations in the Congo Basin. Remote Sensing 13, 22.

Friedlingstein, P., Jones, M.W., O'Sullivan, M., Andrew, R.M., Bakker, D.C.E., Hauck, J., Le Quéré, C., Peters, G.P., Peters, W., Pongratz, J., Sitch, S., Canadell, J.G., Ciais, P., Jackson, R.B., Alin, S.R., Anthoni, P., Bates, N.R., Becker, M., Bellouin, N., Bopp, L., Chau, T.T.T., Chevallier, F., Chini, L.P., Cronin, M., Currie, K.I., Decharme, B., Djeutchouang, L.M., Dou, X., Evans, W., Feely, R.A., Feng, L., Gasser, T., Gilfillan, D., Gkritzalis, T., Grassi, G., Gregor, L., Gruber, N., Gürses, O., Harris, I., Houghton, R.A., Hurtt, G.C., Iida, Y., Ilyina, T., Luijkx, I.T., Jain, A., Jones, S.D., Kato, E., Kennedy, D., Klein Goldewijk, K., Knauer, J., Korsbakken, J.I., Körtzinger, A., Landschützer, P., Lauvset, S.K., Lefèvre, N., Lienert, S., Liu, J., Marland, G., McGuire, P.C., Melton, J.R., Munro, D.R., Nabel, J.E.M.S., Nakaoka, S.I., Niwa, Y., Ono, T., Pierrot, D., Poulter, B., Rehder, G., Resplandy, L., Robertson, E., Rödenbeck, C., Rosan, T.M., Schwinger, J.,

Schwingshackl, C., Séférian, R., Sutton, A.J., Sweeney, C., Tanhua, T., Tans, P.P., Tian, H., Tilbrook, B., Tubiello, F., van der Werf, G.R., Vuichard, N., Wada, C., Wanninkhof, R., Watson, A.J., Willis, D., Wiltshire, A.J., Yuan, W., Yue, C., Yue, X., Zaehle, S., Zeng, J., 2022. Global Carbon Budget 2021. Earth Syst. Sci. Data 14, 1917–2005. doi:10.5194/essd-14-1917-2022.

Friedlingstein, P., O'Sullivan, M., Jones, M.W., Andrew, R.M., Hauck, J., Olsen, A., Peters, G.P., Peters, W., Pongratz, J., Sitch, S., Le Quéré, C., Canadell, J.G., Ciais, P., Jackson, R.B., Alin, S., Aragão, L.E.O.C., Arneth, A., Arora, V., Bates, N.R., Becker, M., Benoit-Cattin, A., Bittig, H.C., Bopp, L., Bultan, S., Chandra, N., Chevallier, F., Chini, L.P., Evans, W., Florentie, L., Forster, P.M., Gasser, T., Gehlen, M., Gilfillan, D., Gkritzalis, T., Gregor, L., Gruber, N., Harris, I., Hartung, K., Haverd, V., Houghton, R.A., Ilyina, T., Jain, A.K., Joetzjer, E., Kadono, K., Kato, E., Kitidis, V., Korsbakken, J.I., Landschützer, P., Lefèvre, N., Lenton, A., Lienert, S., Liu, Z., Lombardozzi, D., Marland, G., Metzl, N., Munro, D.R., Nabel, J.E.M.S., Nakaoka, S.I., Niwa, Y., O'Brien, K., Ono, T., Palmer, P.I., Pierrot, D., Poulter, B., Resplandy, L., Robertson, E., Rödenbeck, C., Schwinger, J., Séférian, R., Skjelvan, I., Smith, A.J.P., Sutton, A.J., Tanhua, T., Tans, P.P., Tian, H., Tilbrook, B., van der Werf, G., Vuichard, N., Walker, A.P., Wanninkhof, R., Watson, A.J., Willis, D., Wiltshire, A.J., Yuan, W., Yue, X., Zaehle, S., 2020. Global Carbon Budget 2020. Earth Syst. Sci. Data 12, 3269–3340. doi:10.5194/essd-12-3269-2020.

Gallant, A.L., 2015. The challenges of remote monitoring of wetlands. doi:10.3390/rs70810938.

Gariano, S.L., Guzzetti, F., 2016. Landslides in a changing climate. Earth-Science Reviews 162, 227–252. doi:10.1016/j.earscirev.2016.08.011.

Garrison, J.L., Katzberg, S.J., 1997. Detection of ocean reflected gps signals: Theory and experiment, in: Proceedings IEEE SOUTHEASTCON'97.'Engineering the New Century', IEEE. pp. 290–294.

Garrison, J.L., Katzberg, S.J., 2000. The application of reflected gps signals to ocean remote sensing. Remote Sensing of Environment 73, 175–187.

Garrison, J.L., Komjathy, A., Zavorotny, V.U., Katzberg, S.J., 2002. Wind speed measurement using forward scattered gps signals. IEEE Transactions on Geoscience and Remote Sensing 40, 50–65.

Ge, M., Gendt, G., Rothacher, M.a., Shi, C., Liu, J., 2008. Resolution of gps carrier-phase ambiguities in precise point positioning (ppp) with daily observations. Journal of geodesy 82, 389–399.

Georgiadou, Y., Kleusberg, A., 1988. On carrier signal multipath effects in relative gps positioning. Manuscripta geodaetica 13, 172–179.

Geremia-Nievinski, F., Hobiger, T., Haas, R., Liu, W., Strandberg, J., Tabibi, S., Vey, S., Wickert, J., Williams, S., 2020. Snr-based gnss reflectometry for coastal sea-level altimetry: results from the first iag inter-comparison campaign. Journal of geodesy 94, 1–15.

Gerlein-Safdi, C., Bloom, A.A., Plant, G., Kort, E.A., Ruf, C.S., 2021. Improving representation of tropical wetland methane emissions with cygnss inundation maps. Global Biogeochemical Cycles 35, e2020GB006890. doi:https://doi.org/10.1029/2020GB006890.

Gerlein-Safdi, C., Ruf, C.S., 2019. A CYGNSS-Based Algorithm for the Detection of Inland Waterbodies. Geophysical Research Letters 46, 12065–12072. doi:10.1029/2019GL085134.

Ghasemigoudarzi, P., Huang, W., De Silva, O., Yan, Q., Power, D.T., 2020. Flash Flood Detection From CYGNSS Data Using the RUSBoost Algorithm. IEEE Access 8, 171864–171881. doi:10.1109/ACCESS.2020.3025302.

Gleason, S., Adjrad, M., Unwin, M., 2005. Sensing ocean, ice and land reflected signals from space: results from the uk-dmc gps reflectometry experiment, in: Proceedings of the 18th International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GNSS 2005), pp. 1679–1685.

Gleason, S., O'Brien, A., Russel, A., Al-Khaldi, M.M., Johnson, J.T., 2020. Geolocation, calibration and surface resolution of cygnss gnss-r land observations. Remote Sensing 12. doi:10.3390/rs12081317.

Gleason, S., Ruf, C.S., O'Brien, A.J., McKague, D.S., 2019. The CYGNSS Level 1 Calibration Algorithm and Error Analysis Based on On-Orbit Measurements. IEEE J. Sel. Top. Appl. Earth Observations Remote Sensing 12, 37–49. doi:10.1109/JSTARS.2018.2832981.

Goffi, A., Stroppiana, D., Brivio, P.A., Bordogna, G., Boschetti, M., 2020. Towards an automated approach to map flooded areas from sentinel-2 msi data and soft integration of water spectral features. International Journal of Applied Earth Observation and Geoinformation 84, 101951. doi:10.1016/j.jag.2019.101951.

Goodwin, P., Williams, R.G., Ridgwell, A., 2015. Sensitivity of climate to cumulative carbon emissions due to compensation of ocean heat and carbon uptake. Nature Geoscience 8, 29–34.

Green, J.K., Seneviratne, S.I., Berg, A.M., Findell, K.L., Hagemann, S., Lawrence, D.M., Gentine, P., 2019. Large influence of soil moisture on long-term terrestrial carbon uptake. Nature 565, 476–479. doi:10.1038/s41586-018-0848-x.

Guzha, A., Rufino, M., Okoth, S., Jacobs, S., Nóbrega, R., 2018. Impacts of land use and land cover change on surface runoff, discharge and low flows: Evidence from East Africa. Journal of Hydrology: Regional Studies 15, 49–67. doi:10.1016/j.ejrh.2017.11.005.

Hall, C., Cordey, R., 1988. Multistatic scatterometry, in: International Geoscience and Remote Sensing Symposium,'Remote Sensing: Moving Toward the 21st Century'., IEEE. pp. 561–562.

Hamilton, S.E., Casey, D., 2016. Creation of a high spatio-temporal resolution global database of continuous mangrove forest cover for the 21st century (cgmfc-21). Global Ecology and Biogeography 25, 729–738.

Hatfield, J.L., Dold, C., 2019. Water-use efficiency: Advances and challenges in a changing climate. Frontiers in Plant Science 10. doi:10.3389/fpls.2019.00103.

Hein, G.W., Avila-Rodriguez, J.A., Wallner, S., 2006. The galileo code and others. Inside GNSS 1, 62–74.

Hess, L.L., Melack, J.M., Affonso, A.G., Barbosa, C., Gastil-Buhl, M., Novo, E.M., 2015. Wetlands of the lowland amazon basin: Extent, vegetative cover, and dual-season inundated area as mapped with jers-1 synthetic aperture radar. Wetlands 35, 745–756. doi:10.1007/s13157-015-0666-y.

Hess, L.L., Melack, J.M., Novo, E.M., Barbosa, C.C., Gastil, M., 2003. Dual-season mapping of wetland inundation and vegetation for the central Amazon basin. Remote Sensing of Environment 87, 404–428. doi:10.1016/j.rse.2003.04.001.

Hirabayashi, Y., Mahendran, R., Koirala, S., Konoshima, L., Yamazaki, D., Watanabe, S., Kim, H., Kanae, S., 2013. Global flood risk under climate change. Nature climate change 3, 816–821. doi:10.1038/nclimate1911.

Hirabayashi, Y., Tanoue, M., Sasaki, O., Zhou, X., Yamazaki, D., 2021. Global exposure to flooding from the new cmip6 climate model projections. Scientific reports 11, 1–7. doi:10.1038/s41598-021-83279-w.

Hoegh-Guldberg, O., Jacob, D., Taylor, M., Bindi, M., Brown, S., Camilloni, I., Diedhiou, A., Djalante, R., Ebi, K., Engelbrecht, F., Guiot, J., Hijioka, Y., Mehrotra, S., Payne, A., Seneviratne, S., Thomas, A., Warren, R., Zhou, G., 2018. Impacts of 1.5ºc global warming on natural and human systems, in: Global Warming of 1.5°C. An IPCC Special Report on the impacts of global warming of 1.5°C above pre-industrial levels and related global greenhouse gas emission pathways, in the context of strengthening the global response to the threat of climate change, sustainable development, and efforts to eradicate poverty. Cambridge University Press, pp. 175–312. doi:10.1017/9781009157940.005.

Holden, L.D., Larson, K.M., 2021. Ten years of lake taupō surface height estimates using the gnss interferometric reflectometry. Journal of Geodesy 95, 1–12.

Holgerson, M.A., Raymond, P.A., 2016. Large contribution to inland water co2 and ch4 emissions from very small ponds. Nature Geoscience 9, 222–226. doi:10.1038/ngeo2654.

Hong, S.H., Wdowinski, S., Kim, S.W., 2010. Evaluation of terrasar-x observations for wetland insar application. IEEE Transactions on Geoscience and Remote Sensing 48, 864–873. doi:10.1109/TGRS.2009.2026895.

Hristov, H.D., 2000. Fresnal Zones in Wireless Links, Zone Plate Lenses and Antennas. Artech House, Inc.

Huang, M., Jin, S., 2020. Rapid flood mapping and evaluation with a supervised classifier and change detection in shouguang using sentinel-1 sar and sentinel-2 optical data. Remote Sensing 12, 2073. doi:10.3390/rs12132073.

Huete, A., Liu, H., Batchily, K., van Leeuwen, W., 1997. A comparison of vegetation indices over a global set of TM images for EOS-MODIS. Remote Sensing of Environment 59, 440–451. doi:10.1016/S0034-4257(96)00112-5.

Hugonnet, R., McNabb, R., Berthier, E., Menounos, B., Nuth, C., Girod, L., Farinotti, D., Huss, M., Dussaillant, I., Brun, F., Kääb, A., 2021. Accelerated global glacier mass loss in the early twenty-first century. Nature 592, 726–731. doi:10.1038/s41586-021-03436-z.

Humphrey, V., Zscheischler, J., Ciais, P., Gudmundsson, L., Sitch, S., Seneviratne, S.I., 2018. Sensitivity of atmospheric CO2 growth rate to observed changes in terrestrial water storage. Nature 560, 628–631. doi:10.1038/s41586-018-0424-4.

Huss, M., Hock, R., 2018. Global-scale hydrological response to future glacier mass loss. Nature Clim Change 8, 135–140. doi:10.1038/s41558-017-0049-x.

Hydroweb, . Water height time series available online:. URL: http://hydroweb.theia-land.fr. (accessed on 31 August 2022).

Jensen, K., Mcdonald, K., 2019. Surface water microwave product series version 3: a near-real time and 25-year historical global inundated area fraction time series from active and passive microwave remote sensing. IEEE Geoscience and remote sensing letters 16, 1402–1406. doi:10.1109/LGRS.2019.2898779.

Jensen, K., McDonald, K., Podest, E., Rodriguez-Alvarez, N., Horna, V., Steiner, N., 2018. Assessing L-Band GNSS-reflectometry and imaging radar for detecting sub-canopy inundation dynamics in a tropicalwetlands complex. Remote Sensing 10. doi:10.3390/rs10091431.

Jones, C.D., Friedlingstein, P., 2020. Quantifying process-level uncertainty contributions to TCRE and carbon budgets for meeting Paris Agreement climate targets. Environ. Res. Lett. 15, 074019. doi:10.1088/1748-9326/ab858a.

Jung, M., Reichstein, M., Schwalm, C.R., Huntingford, C., Sitch, S., Ahlström, A., Arneth, A., Camps-Valls, G., Ciais, P., Friedlingstein, P., Gans, F., Ichii, K., Jain, A.K., Kato, E., Papale, D., Poulter, B., Raduly, B., Rödenbeck, C., Tramontana, G., Viovy, N., Wang, Y.P., Weber, U., Zaehle, S., Zeng, N., 2017. Compensatory water effects link yearly global land CO2 sink changes to temperature. Nature 541, 516–520. doi:10.1038/nature20780.

Junk, W.J., An, S., Finlayson, C., Gopal, B., Květ, J., Mitchell, S.A., Mitsch, W.J., Robarts, R.D., 2013. Current state of knowledge regarding the world's wetlands and their future under global climate change: a synthesis. Aquatic sciences 75, 151–167.

Junk, W.J., Brown, M., Campbell, I.C., Finlayson, M., Gopal, B., Ramberg, L., Warner, B.G., 2006. The comparative biodiversity of seven globally important wetlands: a synthesis. Aquat. Sci. 68, 400–414. doi:10.1007/s00027-006-0856-z.

Kang, D.H., Gao, H., Shi, X., Islam, S.u., Déry, S.J., 2016. Impacts of a Rapidly Declining Mountain Snowpack on Streamflow Timing in Canada's Fraser River Basin. Sci Rep 6, 19299. doi:10.1038/srep19299.

Kaplan, E.D., Hegarty, C., 1996. Understanding GPS/GNSS: Principles and applications. Artech house.

Kaplan, E.D., Hegarty, C., 2017. Understanding GPS/GNSS: Principles and applications. Artech house.

Katzberg, S.J., Garrison, J.L., 1996. Utilizing GPS to determine ionospheric delay over the ocean. Technical Report.

Kavak, A., Vogel, W., et al., 1998. Using gps to measure ground complex permittivity. Electronics Letters 34, 254–255.

Kerr, Y.H., Waldteufel, P., Wigneron, J.P., Martinuzzi, J., Font, J., Berger, M., 2001. Soil moisture retrieval from space: The soil moisture and ocean salinity (smos) mission. IEEE transactions on Geoscience and remote sensing 39, 1729–1735. doi:10.1109/36.942551.

Kim, H., Lakshmi, V., 2018. Use of Cyclone Global Navigation Satellite System (CyGNSS) Observations for Estimation of Soil Moisture. Geophysical Research Letters 45, 11.

Kitambo, B., Papa, F., Paris, A., Tshimanga, R.M., Calmant, S., Fleischmann, A.S., Frappart, F., Becker, M., Tourian, M.J., Prigent, C., et al., 2022a. A combined use of in situ and satellite-derived observations to characterize surface hydrology and its variability in the congo river basin. Hydrology and Earth System Sciences 26, 1857–1882. doi:10.5194/hess-26-1857-2022.

Kitambo, B., Papa, F., Paris, A., Tshimanga, R.M., Frappart, F., Calmant, S., Elmi, O., Fleischmann, A.S., Becker, M., Tourian, M.J., Prigent, C., et al., 2022b. A long-term monthly surface water storage dataset for the congo basin from 1992-2015. in prep. .

Konings, A.G., Piles, M., Das, N., Entekhabi, D., 2017. L-band vegetation optical depth and effective scattering albedo estimation from smap. Remote Sensing of Environment 198, 460–470. doi:10.1016/j.rse.2017.06.037.

Kouadio, I.K., Aljunid, S., Kamigaki, T., Hammad, K., Oshitani, H., 2012. Infectious diseases following natural disasters: prevention and control measures. Expert Review of Anti-infective Therapy 10, 95–104. doi:10.1586/eri.11.155.

Kreibich, H., Van Loon, A.F., Schröter, K., Ward, P.J., Mazzoleni, M., Sairam, N., Abeshu, G.W., Agafonova, S., AghaKouchak, A., Aksoy, H., et al., 2022. The challenge of unprecedented floods and droughts in risk management. Nature , 1–7doi:10.1038/s41586-022-04917-5.

Langley, R.B., 1997. Glonass: Review and update. GPS world .

Larson, K.M., Braun, J.J., Small, E.E., Zavorotny, V.U., Gutmann, E.D., Bilich, A.L., 2009a. Gps multipath and its relation to near-surface soil moisture content. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 3, 91–99.

Larson, K.M., Gutmann, E.D., Zavorotny, V.U., Braun, J.J., Williams, M.W., Nievinski, F.G., 2009b. Can we measure snow depth with gps receivers? Geophysical research letters 36.

Larson, K.M., Löfgren, J.S., Haas, R., 2013a. Coastal sea level measurements using a single geodetic gps receiver. Advances in Space Research 51, 1301–1310. doi:10.1016/j.asr.2012.04.017.

Larson, K.M., Nievinski, F.G., 2013. Gps snow sensing: results from the earthscope plate boundary observatory. GPS solutions 17, 41–52.

Larson, K.M., Ray, R.D., Nievinski, F.G., Freymueller, J.T., 2013b. The accidental tide gauge: a gps reflection case study from kachemak bay, alaska. IEEE Geoscience and Remote Sensing Letters 10, 1200–1204.

Larson, K.M., Small, E.E., Gutmann, E., Bilich, A., Axelrad, P., Braun, J., 2008a. Using gps multipath to measure soil moisture fluctuations: initial results. GPS solutions 12, 173–177. doi:10.1007/s10291-007-0076-6.

Larson, K.M., Small, E.E., Gutmann, E.D., Bilich, A.L., Braun, J.J., Zavorotny, V.U., 2008b. Use of gps receivers as a soil moisture network for water cycle studies. Geophysical Research Letters 35.

Lee, H., Yuan, T., Jung, H.C., Beighley, E., 2015. Mapping wetland water depths over the central congo basin using palsar scansar, envisat altimetry, and modis vcf data. Remote Sensing of Environment 159, 70–79. doi:10.1016/j.rse.2014.11.030.

Lehner, B., Döll, P., 2004. Development and validation of a global database of lakes, reservoirs and wetlands. Journal of Hydrology 296, 1–22. doi:10.1016/j.jhydrol.2004.03.028.

Levy, M.C., Lopes, A.V., Cohn, A., Larsen, L.G., Thompson, S.E., 2018. Land Use Change Increases Streamflow Across the Arc of Deforestation in Brazil. Geophys. Res. Lett. 45, 3520–3530. doi:10.1002/2017GL076526.

Li, W., Cardellach, E., Fabra, F., Ribo, S., Rius, A., 2018. Lake level and surface topography measured with spaceborne gnss-reflectometry from cygnss mission: Example for the lake qinghai. Geophysical Research Letters 45, 13–332.

Li, W., Cardellach, E., Fabra, F., Rius, A., Ribó, S., Martín-Neira, M., 2017. First spaceborne phase altimetry over sea ice using TechDemoSat-1 GNSS-R signals. Geophys. Res. Lett. 44, 8369–8376. doi:10.1002/2017GL074513.

Li, W., CARDELLACH, E., RIBÓ, S., RIUS, A., ZHOU, B., 2021. First spaceborne demonstration of beidou-3 signals for gnss reflectometry from cygnss constellation. Chinese Journal of Aeronautics 34, 1–10. doi:https://doi.org/10.1016/j.cja.2020.11.016.

Li, W., Cardellach, E., Ribó, S., Oliveras, S., Rius, A., 2022. Exploration of multi-mission spaceborne gnss-r raw if data sets: Processing, data products and potential applications. Remote Sensing 14. doi:10.3390/rs14061344.

Lian, X., Piao, S., Chen, A., Huntingford, C., Fu, B., Li, L.Z.X., Huang, J., Sheffield, J., Berg, A.M., Keenan, T.F., McVicar, T.R., Wada, Y., Wang, X., Wang, T., Yang, Y., Roderick, M.L., 2021. Multifaceted characteristics of dryland aridity changes in a warming world. Nat Rev Earth Environ 2, 232–250. doi:10.1038/s43017-021-00144-0.

Llaveria, D., Munoz-Martin, J.F., Herbert, C., Pablos, M., Park, H., Camps, A., 2021. Sea ice concentration and sea ice extent mapping with l-band microwave radiometry and gnss-r data from the ffscat mission using neural networks. Remote Sensing 13, 1139.

Löfgren, J.S., Haas, R., 2014. Sea level measurements using multi-frequency gps and glonass observations. EURASIP Journal on Advances in Signal Processing 2014, 1–13.

Loria, E., O'Brien, A., Zavorotny, V., Downs, B., Zuffada, C., 2020. Analysis of scattering characteristics from inland bodies of water observed by cygnss. Remote Sensing of Environment 245, 111825. doi:https://doi.org/10.1016/j.rse.2020.111825.

MacDougall, A.H., 2016. The transient response to cumulative co2 emissions: a review. Current Climate Change Reports 2, 39–47. doi:10.1007/s40641-015-0030-6.

MacDougall, A.H., Friedlingstein, P., 2015. The origin and limits of the near proportionality between climate warming and cumulative co2 emissions. Journal of Climate 28, 4217–4230. doi:10.1175/JCLI-D-14-00036.1.

MacDougall, A.H., Swart, N.C., Knutti, R., 2017. The uncertainty in the transient climate response to cumulative co2 emissions arising from the uncertainty in physical climate parameters. Journal of Climate 30, 813–827. doi:10.1175/JCLI-D-16-0205.1.

MacDougall, A.H., Zickfeld, K., Knutti, R., Matthews, H.D., 2015. Sensitivity of carbon budgets to permafrost carbon feedbacks and non-CO_2 forcings. Environmental Research Letters 10, 125003. doi:10.1088/1748-9326/10/12/125003.

Martin-Neira, M., et al., 1993. A passive reflectometry and interferometry system (paris): Application to ocean altimetry. ESA journal 17, 331–355.

Martinis, S., Groth, S., Wieland, M., Knopp, L., Rättich, M., 2022. Towards a global seasonal and permanent reference water product from sentinel-1/2 data for improved flood mapping. Remote Sensing of Environment 278, 113077. doi:10.1016/j.rse.2022.113077.

Martinis, S., Kersten, J., Twele, A., 2015a. A fully automated terrasar-x based flood service. ISPRS Journal of Photogrammetry and Remote Sensing 104, 203–212. doi:10.1016/j.isprsjprs.2014.07.014.

Martinis, S., Kuenzer, C., Wendleder, A., Huth, J., Twele, A., Roth, A., Dech, S., 2015b. Comparing four operational sar-based water and flood detection approaches. International Journal of Remote Sensing 36, 3519–3543. doi:10.1080/01431161.2015.1060647.

Martinis, S., Twele, A., Voigt, S., 2009. Towards operational near real-time flood detection using a split-based automatic thresholding procedure on high resolution terrasar-x data. Natural Hazards and Earth System Sciences 9, 303–314.

Martins, K., Bonneton, P., Frappart, F., Detandt, G., Bonneton, N., Blenkinsopp, C.E., 2017. High frequency field measurements of an undular bore using a 2d lidar scanner. Remote Sensing 9, 462.

Marty, C., Tilg, A.M., Jonas, T., 2017. Recent Evidence of Large-Scale Receding Snow Water Equivalents in the European Alps. Journal of Hydrometeorology 18, 1021–1031. doi:10.1175/JHM-D-16-0188.1.

Masson-Delmotte, V., Zhai, P., Pirani, A., Connors, S.L., Péan, C., Berger, S., Caud, N., Chen, Y., Goldfarb, L., Gomis, M., et al., 2021. Climate change 2021: the physical science basis. Contribution of working group I to the sixth assessment report of the intergovernmental panel on climate change , 2.

McGraw, G.A., Braasch, M.S., 1999. Gnss multipath mitigation using gated and high resolution correlator concepts, in: Proceedings of the 1999 national technical meeting of the institute of navigation, pp. 333–342.

Mekonnen, M.M., Hoekstra, A.Y., 2016. Four billion people facing severe water scarcity. Sci. Adv. 2, e1500323. doi:10.1126/sciadv.1500323.

Melton, J., Wania, R., Hodson, E., Poulter, B., Ringeval, B., Spahni, R., Bohn, T., Avis, C., Beerling, D., Chen, G., et al., 2013. Present state of global wetland extent and wetland methane modelling: conclusions from a model inter-comparison project (wetchimp). Biogeosciences 10, 753–788.

Mervart, L., 1995. Ambiguity resolution techniques in geodetic and geodynamic applications of the Global Positioning System. volume 53. Schweizerische Geodätische Kommission.

Miralles, D.G., van den Berg, M.J., Gash, J.H., Parinussa, R.M., de Jeu, R.A.M., Beck, H.E., Holmes, T.R.H., Jiménez, C., Verhoest, N.E.C., Dorigo, W.A., Teuling, A.J., Johannes Dolman, A., 2014. El Niño–La Niña cycle and recent trends in continental evaporation. Nature Clim Change 4, 122–126. doi:10.1038/nclimate2068.

Morris, M., Chew, C., Reager, J.T., Shah, R., Zuffada, C., 2019. A novel approach to monitoring wetland dynamics using CYGNSS: Everglades case study. Remote Sensing of Environment 233, 111417. doi:10.1016/j.rse.2019.111417.

Munoz-Martin, J.F., Camps, A., 2021. Sea surface salinity and wind speed retrievals using gnss-r and l-band microwave radiometry data from fmpl-2 onboard the fsscat mission. Remote Sensing 13, 3224.

Munoz-Martin, J.F., Capon, L.F., Ruiz-de Azua, J.A., Camps, A., 2020a. The flexible microwave payload-2: A sdr-based gnss-reflectometer and l-band radiometer for cubesats. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 13, 1298–1311. doi:10.1109/JSTARS.2020.2977959.

Munoz-Martin, J.F., Fernandez, L., Perez, A., Ruiz-de Azua, J.A., Park, H., Camps, A., Domínguez, B.C., Pastena, M., 2020b. In-orbit validation of the fmpl-2 instrument—the gnss-r and l-band microwave radiometer payload of the fsscat mission. Remote Sensing 13, 121.

Munoz-Martin, J.F., Llaveria, D., Herbert, C., Pablos, M., Park, H., Camps, A., 2021. Soil moisture estimation synergy using gnss-r and l-band microwave radiometry data from fsscat/fmpl-2. Remote Sensing 13, 994.

Munoz-Martin, J.F., Miguelez, N., Castella, R., Fernandez, L., Solanellas, A., Via, P., Camps, A., 2018. 3cat-4: combined gnss-r, l-band radiometer with rfi mitigation, and ais receiver for a i-unit cubesat based on software defined radio, in: IGARSS 2018-2018 IEEE International Geoscience and Remote Sensing Symposium, IEEE. pp. 1063–1066.

Munoz-Martin, J.F., Onrubia, R., Pascual, D., Park, H., Camps, A., Rüdiger, C., Walker, J., Monerris, A., 2020c. Untangling the incoherent and coherent scattering components in gnss-r and novel applications. Remote Sensing 12, 1208.

Munoz-Martin, J.F., Pascual, D., Onrubia, R., Park, H., Camps, A., Rüdiger, C., Walker, J.P., Monerris, A., 2022. Vegetation canopy height retrieval using l1 and l5 airborne gnss-r. IEEE Geoscience and Remote Sensing Letters 19, 1–5. doi:10.1109/LGRS.2021.3131263.

Munoz-Martin, J.F., Perez, A., Camps, A., Ribó, S., Cardellach, E., Stroeve, J., Nandan, V., Itkin, P., Tonboe, R., Hendricks, S., et al., 2020d. Snow and ice thickness retrievals using gnss-r: Preliminary results of the mosaic experiment. Remote Sensing 12, 4038.

Musa, Z., Popescu, I., Mynett, A., 2015. A review of applications of satellite sar, optical, altimetry and dem data for surface water modelling, mapping and parameter estimation. Hydrology and Earth System Sciences 19, 3755–3769.

Nakmuenwai, P., Yamazaki, F., Liu, W., 2017. Automated extraction of inundated areas from multi-temporal dual-polarization radarsat-2 images of the 2011 central thailand flood. Remote Sensing 9, 78. doi:doi:10.3390/rs9010078.

Neal, J., Schumann, G., Bates, P., 2012. A subgrid channel model for simulating river hydraulics and floodplain inundation over large and data sparse areas. Water Resources Research 48.

Nievinski, F.G., Larson, K.M., 2014. Inverse modeling of gps multipath for snow depth estimation—part ii: Application and validation. IEEE transactions on geoscience and remote sensing 52, 6564–6573.

Normandin, C., Frappart, F., Lubac, B., Bélanger, S., Marieu, V., Blarel, F., Robinet, A., Guiastrennec-Faugas, L., 2018. Quantification of surface water volume changes in the Mackenzie Delta using satellite multi-mission data. Hydrol. Earth Syst. Sci. 22, 1543–1561. doi:10.5194/hess-22-1543-2018.

Pangala, S.R., Enrich-Prast, A., Basso, L.S., Peixoto, R.B., Bastviken, D., Hornibrook, E.R.C., Gatti, L.V., Marotta, H., Calazans, L.S.B., Sakuragui, C.M., Bastos, W.R., Malm, O., Gloor, E., Miller, J.B., Gauci, V., 2017. Large emissions from floodplain trees close the Amazon methane budget. Nature 552, 230–234. doi:10.1038/nature24639.

Papa, F., Frappart, F., 2021. Surface water storage in rivers and wetlands derived from satellite observations: a review of current advances and future opportunities for hydrological sciences. Remote Sensing 13, 4162. doi:10.3390/rs13204162.

Papa, F., Frappart, F., Malbeteau, Y., Shamsudduha, M., Vuruputur, V., Sekhar, M., Ramillien, G., Prigent, C., Aires, F., Pandey, R.K., et al., 2015. Satellite-derived surface and sub-surface water storage in the ganges–brahmaputra river basin. Journal of Hydrology: Regional Studies 4, 15–35. doi:10.1016/j.ejrh.2015.03.004.

Parrens, M., Al Bitar, A., Frappart, F., Papa, F., Calmant, S., Crétaux, J.F., Wigneron, J.P., Kerr, Y., 2017. Mapping Dynamic Water Fraction under the Tropical Rain Forests of the Amazonian Basin from SMOS Brightness Temperatures. Water 9, 350. doi:10.3390/w9050350.

Peixoto, J.P., Oort, A.H., 1992. Physics of climate. American Institute of Physics, New York.

Pekel, J.F., Cottam, A., Gorelick, N., Belward, A.S., 2016. High-resolution mapping of global surface water and its long-term changes. Nature 540, 418–422. doi:10.1038/nature20584.

Pendergrass, A.G., Lehner, F., Sanderson, B.M., Xu, Y., 2015. Does extreme precipitation intensity depend on the emissions scenario? Geophys. Res. Lett. 42, 8767–8774. doi:10.1002/2015GL065854.

Pham-Duc, B., Papa, F., Prigent, C., Aires, F., Biancamaria, S., Frappart, F., 2019. Variations of surface and subsurface water storage in the lower mekong basin (vietnam and cambodia) from multisatellite observations. Water 11, 75. doi:10.3390/w11010075.

Pham-Duc, B., Prigent, C., Aires, F., Papa, F., 2017. Comparisons of global terrestrial surface water datasets over 15 years. Journal of Hydrometeorology 18, 993–1007. doi:10.1175/JHM-D-16-0206.1.

Pham-Duc, B., Sylvestre, F., Papa, F., Frappart, F., Bouchez, C., Crétaux, J.F., 2020. The lake chad hydrology under current climate change. Scientific reports 10, 1–10. doi:10.1038/s41598-020-62417-w.

Phillips, O.L., Aragão, L.E., Lewis, S.L., Fisher, J.B., Lloyd, J., López-González, G., Malhi, Y., Monteagudo, A., Peacock, J., Quesada, C.A., et al., 2009. Drought sensitivity of the amazon rainforest. Science 323, 1344–1347. doi:10.1126/science.1164033.

Piao, S., Wang, X., Wang, K., Li, X., Bastos, A., Canadell, J.G., Ciais, P., Friedlingstein, P., Sitch, S., 2020. Interannual variation of terrestrial carbon cycle: Issues and perspectives. Global Change Biology 26, 300–318. doi:10.1111/gcb.14884.

Pierdicca, N., Comite, D., Camps, A., Carreno-Luengo, H., Cenci, L., Clarizia, M.P., Costantini, F., Dente, L., Guerriero, L., Mollfulleda, A., et al., 2021. The potential of spaceborne gnss reflectometry for soil moisture, biomass, and freeze–thaw monitoring: Summary of a european space agency-funded study. IEEE geoscience and remote sensing magazine 10, 8–38.

Pierdicca, N., Pulvirenti, L., Chini, M., Guerriero, L., Candela, L., 2013. Observing floods from space: Experience gained from cosmo-skymed observations. Acta Astronautica 84, 122–133. doi:10.1016/j.actaastro.2012.10.034.

Polade, S.D., Pierce, D.W., Cayan, D.R., Gershunov, A., Dettinger, M.D., 2015. The key role of dry days in changing regional climate and precipitation regimes. Sci Rep 4, 4364. doi:10.1038/srep04364.

Poulter, B., Bousquet, P., Canadell, J.G., Ciais, P., Peregon, A., Saunois, M., Arora, V.K., Beerling, D.J., Brovkin, V., Jones, C.D., et al., 2017. Global wetland contribution to 2000–2012 atmospheric methane growth rate dynamics. Environmental Research Letters 12, 094013.

Prather, M.J., Hsu, J., DeLuca, N.M., Jackman, C.H., Oman, L.D., Douglass, A.R., Fleming, E.L., Strahan, S.E., Steenrod, S.D., Søvde, O.A., Isaksen, I.S.A., Froidevaux, L., Funke, B., 2015. Measuring and modeling the lifetime of nitrous oxide including its variability. JGR Atmospheres 120, 5693–5705. doi:10.1002/2015JD023267.

Prigent, C., Jimenez, C., Bousquet, P., 2020. Satellite-Derived Global Surface Water Extent and Dynamics Over the Last 25 Years (GIEMS-2). J. Geophys. Res. Atmos. 125. doi:10.1029/2019JD030711.

Prigent, C., Matthews, E., Aires, F., Rossow, W.B., 2001. Remote sensing of global wetland dynamics with multiple satellite data sets. Geophys. Res. Lett. 28, 4631–4634. doi:10.1029/2001GL013263.

Prigent, C., Papa, F., Aires, F., Rossow, W.B., Matthews, E., 2007. Global inundation dynamics inferred from multiple satellite observations, 1993–2000. J. Geophys. Res. 112, D12107. doi:10.1029/2006JD007847.

Pulvirenti, L., Chini, M., Pierdicca, N., Guerriero, L., Ferrazzoli, P., 2011. Flood monitoring using multi-temporal cosmo-skymed data: Image segmentation and signature interpretation. Remote Sensing of Environment 115, 990–1002. doi:10.1016/j.rse.2010.12.002.

Purnell, D., Gomez, N., Chan, N.H., Strandberg, J., Holland, D.M., Hobiger, T., 2020. Quantifying the uncertainty in ground-based gnss-reflectometry sea level measurements. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 13, 4419–4428.

Rajabi, M., Nahavandchi, H., Hoseini, M., 2020. Evaluation of CYGNSS Observations for Flood Detection and Mapping during Sistan and Baluchestan Torrential Rain in 2020. Water 12, 2047. doi:10.3390/w12072047.

Ramillien, G., Frappart, F., Seoane, L., 2014. Application of the regional water mass variations from grace satellite gravimetry to large-scale water management in africa. Remote Sensing 6, 7379–7405.

Refice, A., Capolongo, D., Pasquariello, G., D'Addabbo, A., Bovenga, F., Nutricato, R., Lovergine, F.P., Pietranera, L., 2014. Sar and insar for flood monitoring: Examples with cosmo-skymed data. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 7, 2711–2722. doi:10.1109/JSTARS.2014.2305165.

Richards, J.A., Woodgate, P.W., Skidmore, A.K., 1987. An explanation of enhanced radar backscattering from flooded forests. International Journal of Remote Sensing 8, 1093–1100. doi:10.1080/01431168708954756.

Rodríguez-Álvarez, N., 2011. Contributions to earth observation using gnss-r opportunity signals. Ph.D. thesis. Universitat Politècnica de Catalunya (UPC). URL: http://hdl.handle.net/10803/53636.

Rodriguez-Alvarez, N., Aguasca, A., Valencia, E., Bosch-Lluis, X., Camps, A., Ramos-Perez, I., Park, H., Vall-Llossera, M., 2012. Snow thickness monitoring using gnss measurements. IEEE Geoscience and Remote Sensing Letters 9, 1109–1113.

Rodriguez-Alvarez, N., Bosch-Lluis, X., Camps, A., Vall-Llossera, M., Valencia, E., Marchan-Hernandez, J.F., Ramos-Perez, I., 2009. Soil moisture retrieval using gnss-r techniques: Experimental results over a bare soil field. IEEE transactions on geoscience and remote sensing 47, 3616–3624.

Rodriguez-Alvarez, N., Camps, A., Vall-Llossera, M., Bosch-Lluis, X., Monerris, A., Ramos-Perez, I., Valencia, E., Marchan-Hernandez, J.F., Martinez-Fernandez, J., Baroncini-Turricchia, G., et al., 2010. Land geophysical parameters retrieval using the interference pattern gnss-r technique. IEEE Transactions on Geoscience and Remote Sensing 49, 71–84.

Rodriguez-Alvarez, N., Podest, E., Jensen, K., McDonald, K.C., 2019. Classifying Inundation in a Tropical Wetlands Complex with GNSS-R. Remote Sensing 11, 1053. doi:10.3390/rs11091053.

Roussel, N., 2015. Application de la réflectométrie GNSS à l'étude des redistributions des masses d'eau à la surface de la terre. Ph.D. thesis. Université Paul Sabatier-Toulouse III.

Roussel, N., Frappart, F., Ramillien, G., Darrozes, J., Baup, F., Lestarquit, L., Ha, M.C., 2016. Detection of soil moisture variations using gps and glonass snr data for elevation angles ranging from 2 to 70. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 9, 4781–4794.

Roussel, N., Ramillien, G., Frappart, F., Darrozes, J., Gay, A., Biancale, R., Striebig, N., Hanquiez, V., Bertin, X., Allain, D., 2015. Sea level monitoring and sea state estimate using a single geodetic receiver. Remote sensing of Environment 171, 261–277. doi:10.1016/j.rse.2015.10.011.

Ruf, C.S., Atlas, R., Chang, P.S., Clarizia, M.P., Garrison, J.L., Gleason, S., Katzberg, S.J., Jelenak, Z., Johnson, J.T., Majumdar, S.J., O'Brien, A., Posselt, D.J., Ridley, A.J., Rose, R.J., Zavorotny, V.U., 2016. New ocean winds satellite mission to probe hurricanes and tropical convection. Bulletin of the American Meteorological Society 97, 385–395. doi:10.1175/BAMS-D-14-00218.1.

Ruffini, G., Soulat, F., Caparrini, M., Germain, O., Martín-Neira, M., 2004. The Eddy Experiment: Accurate GNSS-R ocean altimetry from low altitude aircraft. Geophys. Res. Lett. 31. doi:10.1029/2004GL019994.

Russo, I.M., Bisceglie, M.d., Galdi, C., Lavalle, M., Zuffada, C., 2022. Entropy-based coherence metric for land applications of gnss-r. IEEE Transactions on Geoscience and Remote Sensing 60, 1–13. doi:10.1109/TGRS.2021.3125858.

Sakamoto, T., Van Nguyen, N., Kotera, A., Ohno, H., Ishitsuka, N., Yokozawa, M., 2007. Detecting temporal changes in the extent of annual flooding within the Cambodia and the Vietnamese Mekong Delta from MODIS time-series imagery. Remote Sensing of Environment 109, 295–313. doi:10.1016/j.rse.2007.01.011.

Salameh, E., Frappart, F., Papa, F., Güntner, A., Venugopal, V., Getirana, A., Prigent, C., Aires, F., Labat, D., Laignel, B., 2017. Fifteen years (1993–2007) of surface freshwater storage variability in the ganges-brahmaputra river basin using multi-satellite observations. Water 9, 245. doi:10.3390/w9040245.

Salamon, P., Mctlormick, N., Reimer, C., Clarke, T., Bauer-Marschallinger, B., Wagner, W., Martinis, S., Chow, C., Böhnke, C., Matgen, P., et al., 2021. The new, systematic global flood monitoring product of the copernicus emergency management service, in: 2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS, IEEE. pp. 1053–1056. doi:10.1109/IGARSS47720.2021.9554214.

Sandvik, M.I., Sorteberg, A., Rasmussen, R., 2018. Sensitivity of historical orographically enhanced extreme precipitation events to idealized temperature perturbations. Clim Dyn 50, 143–157. doi:10.1007/s00382-017-3593-1.

Santoro, M., 2018. GlobBiomass - global datasets of forest biomass. doi:10.1594/PANGAEA.894711.

Santoro, M., Cartus, O., Carvalhais, N., Rozendaal, D.M.A., Avitabile, V., Araza, A., de Bruin, S., Herold, M., Quegan, S., Rodríguez-Veiga, P., Balzter, H., Carreiras, J., Schepaschenko, D., Korets, M., Shimada, M., Itoh, T., Moreno Martínez, A., Cavlovic, J., Cazzolla Gatti, R., da Conceição Bispo, P., Dewnath, N., Labrière, N., Liang, J., Lindsell, J., Mitchard, E.T.A., Morel, A., Pacheco Pascagaza, A.M., Ryan, C.M., Slik, F., Vaglio Laurin, G., Verbeeck, H., Wijaya, A., Willcock, S., 2021. The global forest aboveground biomass pool for 2010 estimated from high-resolution satellite observations. Earth System Science Data 13, 3927–3950. doi:10.5194/essd-13-3927-2021.

Santoro, M., Wegmüller, U., Lamarche, C., Bontemps, S., Defourny, P., Arino, O., 2015. Strengths and weaknesses of multi-year envisat asar backscatter measurements to map permanent open water bodies at global scale. Remote Sensing of Environment 171, 185–201. doi:10.1016/j.rse.2015.10.031.

Saunois, M., Bousquet, P., Poulter, B., Peregon, A., Ciais, P., Canadell, J.G., Dlugokencky, E.J., Etiope, G., Bastviken, D., Houweling, S., et al., 2017. Variability and quasi-decadal changes in the methane budget over the period 2000–2012. Atmospheric Chemistry and Physics 17, 11135–11161.

Saunois, M., Stavert, A.R., Poulter, B., Bousquet, P., Canadell, J.G., Jackson, R.B., Raymond, P.A., Dlugokencky, E.J., Houweling, S., Patra, P.K., Ciais, P., Arora, V.K.,

Bastviken, D., Bergamaschi, P., Blake, D.R., Brailsford, G., Bruhwiler, L., Carlson, K.M., Carrol, M., Castaldi, S., Chandra, N., Crevoisier, C., Crill, P.M., Covey, K., Curry, C.L., Etiope, G., Frankenberg, C., Gedney, N., Hegglin, M.I., Höglund-Isaksson, L., Hugelius, G., Ishizawa, M., Ito, A., Janssens-Maenhout, G., Jensen, K.M., Joos, F., Kleinen, T., Krummel, P.B., Langenfelds, R.L., Laruelle, G.G., Liu, L., Machida, T., Maksyutov, S., McDonald, K.C., McNorton, J., Miller, P.A., Melton, J.R., Morino, I., Müller, J., Murguia-Flores, F., Naik, V., Niwa, Y., Noce, S., O'Doherty, S., Parker, R.J., Peng, C., Peng, S., Peters, G.P., Prigent, C., Prinn, R., Ramonet, M., Regnier, P., Riley, W.J., Rosentreter, J.A., Segers, A., Simpson, I.J., Shi, H., Smith, S.J., Steele, L.P., Thornton, B.F., Tian, H., Tohjima, Y., Tubiello, F.N., Tsuruta, A., Viovy, N., Voulgarakis, A., Weber, T.S., van Weele, M., van der Werf, G.R., Weiss, R.F., Worthy, D., Wunch, D., Yin, Y., Yoshida, Y., Zhang, W., Zhang, Z., Zhao, Y., Zheng, B., Zhu, Q., Zhu, Q., Zhuang, Q., 2020. The Global Methane Budget 2000–2017. Earth Syst. Sci. Data 12, 1561–1623. doi:10.5194/essd-12-1561-2020.

Scheff, J., Frierson, D.M.W., 2014. Scaling Potential Evapotranspiration with Greenhouse Warming. Journal of Climate 27, 1539–1558. doi:10.1175/JCLI-D-13-00233.1.

Scheff, J., Frierson, D.M.W., 2015. Terrestrial Aridity and Its Response to Greenhouse Warming across CMIP5 Climate Models. Journal of Climate 28, 5583–5600. doi:10.1175/JCLI-D-14-00480.1.

Schewe, J., Heinke, J., Gerten, D., Haddeland, I., Arnell, N.W., Clark, D.B., Dankers, R., Eisner, S., Fekete, B.M., Colón-González, F.J., Gosling, S.N., Kim, H., Liu, X., Masaki, Y., Portmann, F.T., Satoh, Y., Stacke, T., Tang, Q., Wada, Y., Wisser, D., Albrecht, T., Frieler, K., Piontek, F., Warszawski, L., Kabat, P., 2014. Multimodel assessment of water scarcity under climate change. Proc. Natl. Acad. Sci. U.S.A. 111, 3245–3250. doi:10.1073/pnas.1222460110.

Schiavulli, D., Frappart, F., Ramillien, G., Darrozes, J., Nunziata, F., Migliaccio, M., 2017. Observing sea/ice transition using radar images generated from techdemosat-1 delay doppler maps. IEEE Geoscience and Remote Sensing Letters 14, 734–738.

Schroeder, R., McDonald, K., Chapman, B., Jensen, K., Podest, E., Tessler, Z., Bohn, T., Zimmermann, R., 2015. Development and Evaluation of a Multi-Year Fractional Surface Water Data Set Derived from Active/Passive Microwave Remote Sensing Data. Remote Sensing 7, 16688–16732. doi:10.3390/rs71215843.

Schwatke, C., Dettmering, D., Bosch, W., Seitz, F., 2015. Dahiti–an innovative approach for estimating water level time series over inland waters using multi-mission satellite altimetry. Hydrology and Earth System Sciences 19, 4345–4364.

Senyurek, V., Lei, F., Boyd, D., Kurum, M., Gurbuz, A.C., Moorhead, R., 2020. Machine learning-based CYGNSS soil moisture estimates over ISMN sites in CONUS. Remote Sensing 12, 1–24. doi:10.3390/rs12071168.

Shiklomanov, A.I., Lammers, R.B., Vörösmarty, C.J., 2002. Widespread decline in hydrological monitoring threatens pan-arctic research. Eos, Transactions American Geophysical Union 83, 13–17. doi:10.1029/2002EO000007.

da Silva, S.S., Fearnside, P.M., de Alencastro Graça, P.M.L., Brown, I.F., Alencar, A., de Melo, A.W.F., 2018. Dynamics of forest fires in the southwestern amazon. Forest Ecology and Management 424, 312–322. doi:10.1016/j.foreco.2018.04.041.

Singha, M., Dong, J., Sarmah, S., You, N., Zhou, Y., Zhang, G., Doughty, R., Xiao, X., 2020. Identifying floods and flood-affected paddy rice fields in bangladesh based on sentinel-1 imagery and google earth engine. ISPRS Journal of Photogrammetry and Remote Sensing 166, 278–293. doi:10.1016/j.isprsjprs.2020.06.011.

Sippel, S.J., Hamilton, S.K., Melack, J.M., Choudhury, B.J., 1994. Determination of inundation area in the Amazon River floodplain using the SMMR 37 GHz polarization difference. Remote Sensing of Environment 48, 70–76. doi:10.1016/0034-4257(94)90115-5.

Springer, T., Dach, R., 2010. Gps, glonass, and more. GPS world 21, 48–58.

Stilla, D., Zribi, M., Pierdicca, N., Baghdadi, N., Huc, M., 2020. Desert roughness retrieval using cygnss gnss-r data. Remote Sensing 12, 743.

Strandberg, J., Hobiger, T., Haas, R., 2016. Improving gnss-r sea level determination through inverse modeling of snr data. Radio science 51, 1286–1296.

Suk, J.E., Vaughan, E.C., Cook, R.G., Semenza, J.C., 2020. Natural disasters and infectious disease in Europe: a literature review to identify cascading risk pathways. European Journal of Public Health 30, 928–935. doi:10.1093/eurpub/ckz111.

Szopa, S., Naik, V., Adhikary, B., Artaxo, P., Berntsen, T., Collins, W., Fuzzi, S., Gallardo, L., Kiendler-Scharr, A., Klimont, Z., Liao, H., Unger, N., Zanis, P., 2021. Short-lived climate forcers, in: Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, p. 817–922. doi:10.1017/9781009157896.008.

Tabibi, S., Francis, O., 2020. Can gnss-r detect abrupt water level changes? Remote sensing 12, 3614.

Tabibi, S., Geremia-Nievinski, F., Francis, O., van Dam, T., 2020. Tidal analysis of gnss reflectometry applied for coastal sea level sensing in antarctica and greenland. Remote sensing of environment 248, 111959.

Tapley, B.D., Bettadpur, S., Ries, J.C., Thompson, P.F., Watkins, M.M., 2004. Grace measurements of mass variability in the earth system. science 305, 503–505.

Tapley, B.D., Watkins, M.M., Flechtner, F., Reigber, C., Bettadpur, S., Rodell, M., Sasgen, I., Famiglietti, J.S., Landerer, F.W., Chambers, D.P., et al., 2019. Contributions of grace to understanding climate change. Nature climate change 9, 358–369.

Tarpanelli, A., Mondini, A.C., Camici, S., 2022. Effectiveness of sentinel-1 and sentinel-2 for flood detection assessment in europe. Natural Hazards and Earth System Sciences 22, 2473–2489. doi:10.5194/nhess-22-2473-2022.

Taylor, R.G., Scanlon, B., Döll, P., Rodell, M., van Beek, R., Wada, Y., Longuevergne, L., Leblanc, M., Famiglietti, J.S., Edmunds, M., Konikow, L., Green, T.R., Chen, J., Taniguchi, M., Bierkens, M.F.P., MacDonald, A., Fan, Y., Maxwell, R.M., Yechieli, Y., Gurdak, J.J., Allen, D.M., Shamsudduha, M., Hiscock, K., Yeh, P.J.F., Holman, I., Treidel, H., 2013a. Ground water and climate change. Nature Clim Change 3, 322–329. doi:10.1038/nclimate1744.

Taylor, R.G., Todd, M.C., Kongola, L., Maurice, L., Nahozya, E., Sanga, H., MacDonald, A.M., 2013b. Evidence of the dependence of groundwater resources on extreme rainfall in East Africa. Nature Clim Change 3, 374–378. doi:10.1038/nclimate1731.

Teunissen, P.J., 1996. Gps carrier phase ambiguity fixing concepts. GPS for Geodesy , 263–335.

Tian, H., Xu, R., Canadell, J.G., Thompson, R.L., Winiwarter, W., Suntharalingam, P., Davidson, E.A., Ciais, P., Jackson, R.B., Janssens-Maenhout, G., Prather, M.J., Regnier, P., Pan, N., Pan, S., Peters, G.P., Shi, H., Tubiello, F.N., Zaehle, S., Zhou, F., Arneth, A., Battaglia, G., Berthet, S., Bopp, L., Bouwman, A.F., Buitenhuis, E.T., Chang, J., Chipperfield, M.P., Dangal, S.R.S., Dlugokencky, E., Elkins, J.W., Eyre, B.D., Fu, B., Hall, B., Ito, A., Joos, F., Krummel, P.B., Landolfi, A., Laruelle, G.G., Lauerwald, R., Li, W., Lienert, S., Maavara, T., MacLeod, M., Millet, D.B., Olin, S., Patra, P.K., Prinn, R.G., Raymond, P.A., Ruiz, D.J., van der Werf, G.R., Vuichard, N., Wang, J., Weiss, R.F., Wells, K.C., Wilson, C., Yang, J., Yao, Y., 2020. A comprehensive quantification of global nitrous oxide sources and sinks. Nature 586, 248–256. doi:10.1038/s41586-020-2780-0.

Tong, X., Luo, X., Liu, S., Xie, H., Chao, W., Liu, S., Liu, S., Makhinov, A., Makhinova, A., Jiang, Y., 2018. An approach for flood monitoring by the combined use of landsat 8 optical imagery and cosmo-skymed radar imagery. ISPRS journal of photogrammetry and remote sensing 136, 144–153. doi:10.1016/j.isprsjprs.2017.11.006.

Tootchi, A., Jost, A., Ducharne, A., 2019. Multi-source global wetland maps combining surface water imagery and groundwater constraints. Earth Syst. Sci. Data 11, 189–220.

Torrence, C., Compo, G.P., 1998. A practical guide to wavelet analysis. Bulletin of the American Meteorological society 79, 61–78.

Tortini, R., Noujdina, N., Yeo, S., Ricko, M., Birkett, C.M., Khandelwal, A., Kumar, V., Marlier, M.E., Lettenmaier, D.P., 2020. Satellite-based remote sensing data set of global surface water storage change from 1992 to 2018. Earth System Science Data 12, 1141–1151. doi:10.5194/essd-12-1141-2020.

Tourian, M., Reager, J., Sneeuw, N., 2018. The total drainable water storage of the amazon river basin: A first estimate using grace. Water Resources Research 54, 3290–3312. doi:10.1029/2017WR021674.

Tsyganskaya, V., Martinis, S., Marzahn, P., Ludwig, R., 2018. Sar-based detection of flooded vegetation–a review of characteristics and approaches. International journal of remote sensing 39, 2255–2293. doi:10.1080/01431161.2017.1420938.

Turetsky, M.R., Abbott, B.W., Jones, M.C., Anthony, K.W., Olefeldt, D., Schuur, E.A., Grosse, G., Kuhry, P., Hugelius, G., Koven, C., et al., 2020. Carbon release through abrupt permafrost thaw. Nature Geoscience 13, 138–143. doi:10.1038/s41561-019-0526-0.

Twele, A., Cao, W., Plank, S., Martinis, S., 2016. Sentinel-1-based flood mapping: a fully automated processing chain. International Journal of Remote Sensing 37, 2990–3004. doi:10.1080/01431161.2016.1192304.

Uddin, K., Matin, M.A., Meyer, F.J., 2019. Operational flood mapping using multi-temporal sentinel-1 sar images: A case study from bangladesh. Remote Sensing 11, 1581. doi:10.3390/rs11131581.

Ukkola, A.M., De Kauwe, M.G., Roderick, M.L., Abramowitz, G., Pitman, A.J., 2020. Robust Future Changes in Meteorological Drought in CMIP6 Projections Despite Uncertainty in Precipitation. Geophys. Res. Lett. 47. doi:10.1029/2020GL087820.

Ukkola, A.M., Prentice, I., Keenan, T.F., van Dijk, A.I., Viney, N.R., Myneni, R., Bi, J., 2016. Reduced streamflow in water-stressed climates consistent with CO2 effects on vegetation. Nature Clim Change 6, 75–78. doi:10.1038/nclimate2831.

Ulaby, F.T., Long, D.G., Blackwell, W.J., Elachi, C., Fung, A.K., Ruf, C., Sarabandi, K., Zebker, H.A., Van Zyl, J., 2014. Microwave radar and radiometric remote sensing. volume 4. University of Michigan Press Ann Arbor, MI, USA.

Unwin, M.J., Pierdicca, N., Cardellach, E., Rautiainen, K., Foti, G., Blunt, P., Guerriero, L., Santi, E., Tossaint, M., 2021. An introduction to the hydrognss gnss reflectometry remote sensing mission. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 14, 6987–6999. doi:10.1109/JSTARS.2021.3089550.

Urlichich, Y., Subbotin, V., Stupak, G., Dvorkin, V., Povalyaev, A., Karutin, S., 2011. Innovation-glonass-developing strategies for the future. GPS world 22, 42.

Verpoorter, C., Kutser, T., Seekell, D.A., Tranvik, L.J., 2014. A global inventory of lakes based on high-resolution satellite imagery. Geophysical Research Letters 41, 6396–6402. doi:10.1002/2014GL060641.

Vicente-Serrano, S.M., Miralles, D.G., Domínguez-Castro, F., Azorin-Molina, C., El Kenawy, A., McVicar, T.R., Tomás-Burguera, M., Beguería, S., Maneta, M., Peña-Gallardo, M., 2018. Global Assessment of the Standardized Evapotranspiration Deficit Index (SEDI) for Drought Analysis and Monitoring. J. Climate 31, 5371–5393. doi:10.1175/JCLI-D-17-0775.1.

Voronovich, A.G., Zavorotny, V.U., 2017. The transition from weak to strong diffuse radar bistatic scattering from rough ocean surface. IEEE Transactions on Antennas and Propagation 65, 6029–6034. doi:10.1109/TAP.2017.2752219.

Voronovich, A.G., Zavorotny, V.U., 2018. Bistatic radar equation for signals of opportunity revisited. IEEE Transactions on Geoscience and Remote Sensing 56, 1959–1968. doi:10.1109/TGRS.2017.2771253.

Vu, P.L., Ha, M.C., Frappart, F., Darrozes, J., Ramillien, G., Dufrechou, G., Gegout, P., Morichon, D., Bonneton, P., 2019. Identifying 2010 xynthia storm signature in gnss-r-based tide records. Remote Sensing 11, 782.

Wada, Y., Bierkens, M.F.P., 2014. Sustainability of global water use: past reconstruction and future projections. Environ. Res. Lett. 9, 104003. doi:10.1088/1748-9326/9/10/104003.

Walker, A.P., De Kauwe, M.G., Bastos, A., Belmecheri, S., Georgiou, K., Keeling, R.F., McMahon, S.M., Medlyn, B.E., Moore, D.J.P., Norby, R.J., Zaehle, S., Anderson-Teixeira, K.J., Battipaglia, G., Brienen, R.J.W., Cabugao, K.G., Cailleret, M., Campbell, E., Canadell, J.G., Ciais, P., Craig, M.E., Ellsworth, D.S., Farquhar, G.D., Fatichi, S., Fisher, J.B., Frank, D.C., Graven, H., Gu, L., Haverd, V., Heilman, K., Heimann, M., Hungate, B.A., Iversen, C.M., Joos, F., Jiang, M., Keenan, T.F., Knauer, J., Körner, C., Leshyk, V.O., Leuzinger, S., Liu, Y., MacBean, N., Malhi, Y., McVicar, T.R., Penuelas, J., Pongratz, J., Powell, A.S., Riutta, T., Sabot, M.E.B., Schleucher, J., Sitch, S., Smith, W.K., Sulman, B., Taylor, B., Terrer, C., Torn, M.S., Treseder, K.K., Trugman, A.T., Trumbore, S.E., Mantgem, P.J., Voelker, S.L., Whelan, M.E., Zuidema, P.A., 2021. Integrating the evidence for a terrestrial carbon sink caused by increasing atmospheric CO $_2$. New Phytologist 229, 2413–2445. doi:10.1111/nph.16866.

Walvoord, M.A., Kurylyk, B.L., 2016. Hydrologic Impacts of Thawing Permafrost-A Review. Vadose Zone Journal 15, vzj2016.01.0010. doi:10.2136/vzj2016.01.0010.

Wan, W., Liu, B., Zeng, Z., Chen, X., Wu, G., Xu, L., Chen, X., Hong, Y., 2019. Using CYGNSS Data to Monitor China's Flood Inundation during Typhoon and Extreme Precipitation Events in 2017. Remote Sensing 11, 854. doi:10.3390/rs11070854.

Wan, W., Zhang, J., Dai, L., Liang, H., Yang, T., Liu, B., Guo, Z., Hu, H., Zhao, L., 2022. A new snow depth data set over northern china derived using gnss interferometric reflectometry from a continuously operating network (gsnow-china v1. 0, 2013–2022). Earth System Science Data 14, 3549–3571.

Wang, J., Hu, Y., Li, Z., 2022. A new coherence detection method for mapping inland water bodies using cygnss data. Remote Sensing 14, 3195.

Wdowinski, S., Kim, S.W., Amelung, F., Dixon, T.H., Miralles-Wilhelm, F., Sonenshein, R., 2008. Space-based detection of wetlands' surface water level changes from l-band sar interferometry. Remote Sensing of Environment 112, 681–696. doi:10.1016/j.rse.2007.06.008.

Wei, Z., Yoshimura, K., Wang, L., Miralles, D.G., Jasechko, S., Lee, X., 2017. Revisiting the contribution of transpiration to global terrestrial evapotranspiration: Revisiting Global ET Partitioning. Geophys. Res. Lett. 44, 2792–2801. doi:10.1002/2016GL072235.

Westerhoff, R., Kleuskens, M., Winsemius, H., Huizinga, H., Brakenridge, G., Bishop, C., 2013. Automated global water mapping based on wide-swath orbital synthetic-aperture radar. Hydrology and Earth System Sciences 17, 651–663.

Wickert, J., Cardellach, E., Martín-Neira, M., Bandeiras, J., Bertino, L., Andersen, O.B., Camps, A., Catarino, N., Chapron, B., Fabra, F., et al., 2016. Geros-iss: Gnss reflectometry, radio occultation, and scatterometry onboard the international space station. IEEE Journal of selected topics in applied Earth observations and Remote Sensing 9, 4552–4581.

Wigneron, J.P., Li, X., Frappart, F., Fan, L., Al-Yaari, A., De Lannoy, G., Liu, X., Wang, M., Le Masson, E., Moisy, C., 2021. Smos-ic data record of soil moisture and l-vod: Historical development, applications and perspectives. Remote Sensing of Environment 254, 112238. doi:10.1016/j.rse.2020.112238.

Williams, R.G., Goodwin, P., Roussenov, V.M., Bopp, L., 2016. A framework to understand the transient climate response to emissions. Environmental Research Letters 11, 015003. doi:10.1088/1748-9326/11/1/015003.

Williams, R.G., Katavouta, A., Goodwin, P., 2019. Carbon-Cycle Feedbacks Operating in the Climate System. Curr Clim Change Rep 5, 282–295. doi:10.1007/s40641-019-00144-9.

Woodcock, C.E., Allen, R., Anderson, M., Belward, A., Bindschadler, R., Cohen, W., Gao, F., Goward, S.N., Helder, D., Helmer, E., et al., 2008. Free access to landsat imagery. SCIENCE VOL 320: 1011 .

Wulder, M.A., White, J.C., Loveland, T.R., Woodcock, C.E., Belward, A.S., Cohen, W.B., Fosnight, E.A., Shaw, J., Masek, J.G., Roy, D.P., 2016. The global landsat archive: Status, consolidation, and direction. Remote Sensing of Environment 185, 271–283. doi:10.1016/j.rse.2015.11.032.

Xiao, X., Boles, S., Liu, J., Zhuang, D., Frolking, S., Li, C., Salas, W., Moore, B., 2005. Mapping paddy rice agriculture in southern China using multi-temporal MODIS images. Remote Sensing of Environment 95, 480–492. doi:10.1016/j.rse.2004.12.009.

Yamazaki, D., Kanae, S., Kim, H., Oki, T., 2011. A physically based description of floodplain inundation dynamics in a global river routing model. Water Resources Research 47. doi:10.1029/2010WR009726.

Yamazaki, D., Trigg, M.A., Ikeshima, D., 2015. Development of a global˜ 90 m water body map using multi-temporal landsat images. Remote Sensing of Environment 171, 337–351. doi:10.1016/j.rse.2015.10.014.

Yan, Q., Huang, W., 2016. Spaceborne gnss-r sea ice detection using delay-doppler maps: First results from the uk techdemosat-1 mission. IEEE journal of selected topics in applied earth observations and remote sensing 9, 4795–4801.

Yan, Q., Huang, W., 2020. Sea ice thickness measurement using spaceborne gnss-r: First results with techdemosat-1 data. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 13, 577–587.

Yan, Q., Huang, W., Jin, S., Jia, Y., 2020. Pan-tropical soil moisture mapping based on a three-layer model from CYGNSS GNSS-R data. Remote Sensing of Environment 247, 111944. doi:10.1016/j.rse.2020.111944.

Yan, Q., Huang, W., Moloney, C., 2017. Neural networks based sea ice detection and concentration retrieval from gnss-r delay-doppler maps. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 10, 3789–3798.

Yang, Y., Gao, W., Guo, S., Mao, Y., Yang, Y., 2019. Introduction to beidou-3 navigation satellite system. Navigation 66, 7–18.

Yin, J., Gentine, P., Zhou, S., Sullivan, S.C., Wang, R., Zhang, Y., Guo, S., 2018. Large increase in global storm runoff extremes driven by climate and anthropogenic changes. Nat Commun 9, 4389. doi:10.1038/s41467-018-06765-2.

Yuan, T., Lee, H., Jung, H.C., Aierken, A., Beighley, E., Alsdorf, D.E., Tshimanga, R.M., Kim, D., 2017. Absolute water storages in the congo river floodplains from integration of insar and satellite radar altimetry. Remote Sensing of Environment 201, 57–72. doi:10.1016/j.rse.2017.09.003.

Zavorotny, V.U., Gleason, S., Cardellach, E., Camps, A., 2014. Tutorial on remote sensing using GNSS bistatic radar of opportunity. IEEE Geoscience and Remote Sensing Magazine 2, 8–45. doi:10.1109/MGRS.2014.2374220. publisher: IEEE.

Zavorotny, V.U., Voronovich, A.G., 2000. Scattering of gps signals from the ocean with wind remote sensing application. IEEE Transactions on Geoscience and Remote Sensing 38, 951–964.

Zeiger, P., Frappart, F., Darrozes, J., Prigent, C., Jiménez, C., 2022. Analysis of cygnss coherent reflectivity over land for the characterization of pan-tropical inundation dynamics. Remote Sensing of Environment 282, 113278. doi:https://doi.org/10.1016/j.rse.2022.113278.

Zeiger, P., Frappart, F., Darrozes, J., Roussel, N., Bonneton, P., Bonneton, N., Detandt, G., 2021. Snr-based water height retrieval in rivers: Application to high amplitude asymmetric tides in the garonne river. Remote Sensing 13, 1856.

Zemp, M., Huss, M., Thibert, E., Eckert, N., McNabb, R., Huber, J., Barandun, M., Machguth, H., Nussbaumer, S.U., Gärtner-Roer, I., Thomson, L., Paul, F., Maussion, F., Kutuzov, S., Cogley, J.G., 2019. Global glacier mass changes and their contributions to sea-level rise from 1961 to 2016. Nature 568, 382–386. doi:10.1038/s41586-019-1071-0.

Zeng, Z., Peng, L., Piao, S., 2018. Response of terrestrial evapotranspiration to Earth's greening. Current Opinion in Environmental Sustainability 33, 9–25. doi:10.1016/j.cosust.2018.03.001.

Zhang, K., Kimball, J.S., Nemani, R.R., Running, S.W., Hong, Y., Gourley, J.J., Yu, Z., 2015. Vegetation Greening and Climate Change Promote Multidecadal Rises of Global Land Evapotranspiration. Sci Rep 5, 15956. doi:10.1038/srep15956.

Zhang, S., Calvet, J.C., Darrozes, J., Roussel, N., Frappart, F., Bouhours, G., 2018. Deriving surface soil moisture from reflected gnss signal observations from a grassland site in southwestern france. Hydrology and Earth System Sciences 22, 1931–1946.

Zhang, S., Roussel, N., Boniface, K., Ha, M.C., Frappart, F., Darrozes, J., Baup, F., Calvet, J.C., 2017. Use of reflected gnss snr data to retrieve either soil moisture or vegetation height from a wheat crop. Hydrology and Earth System Sciences 21, 4767–4784.

Zhang, Y., Peña-Arancibia, J.L., McVicar, T.R., Chiew, F.H.S., Vaze, J., Liu, C., Lu, X., Zheng, H., Wang, Y., Liu, Y.Y., Miralles, D.G., Pan, M., 2016. Multi-decadal trends in global terrestrial evapotranspiration and its components. Sci Rep 6, 19124. doi:10.1038/srep19124.

Zhu, Z., Piao, S., Myneni, R.B., Huang, M., Zeng, Z., Canadell, J.G., Ciais, P., Sitch, S., Friedlingstein, P., Arneth, A., Cao, C., Cheng, L., Kato, E., Koven, C., Li, Y., Lian, X., Liu, Y., Liu, R., Mao, J., Pan, Y., Peng, S., Peñuelas, J., Poulter, B., Pugh, T.A.M., Stocker, B.D., Viovy, N., Wang, X., Wang, Y., Xiao, Z., Yang, H., Zaehle, S., Zeng, N., 2016. Greening of the Earth and its drivers. Nature Clim Change 6, 791–795. doi:10.1038/nclimate3004.