



**HAL**  
open science

# Cartographie multi-échelle des ressources en eau sur les régions irriguées : une approche par télédétection multi-capteurs

Claire Pascal

► **To cite this version:**

Claire Pascal. Cartographie multi-échelle des ressources en eau sur les régions irriguées : une approche par télédétection multi-capteurs. Hydrologie. Université Paul Sabatier - Toulouse III, 2022. Français. NNT : 2022TOU30277 . tel-04139330

**HAL Id: tel-04139330**

**<https://theses.hal.science/tel-04139330>**

Submitted on 23 Jun 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



# THÈSE

**En vue de l'obtention du  
DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE  
Délivré par l'Université Toulouse 3 - Paul Sabatier**

---

**Présentée et soutenue par  
Claire PASCAL**

Le 18 octobre 2022

**Cartographie multi-échelle des ressources en eau sur les régions  
irriguées: une approche par télédétection multi-capteurs**

---

Ecole doctorale : **SDUZE - Sciences de l'Univers, de l'Environnement et de  
l'Espace**

Spécialité : **Surfaces et interfaces continentales, Hydrologie**

Unité de recherche :  
**CESBIO - Centre d'Etudes Spatiales de la BIOSphère**

Thèse dirigée par  
**Olivier MERLIN et Sylvain FERRANT**

Jury

**M. Laurent LONGUEVERNE, Rapporteur**  
**M. Thierry PELLARIN, Rapporteur**  
**M. Benoît DEWANDEL, Rapporteur**  
**M. Lionel JARLAN, Examinateur**  
**Mme Marie-Paule BONNET, Examinatrice**  
**Mme Valérie DEMAREZ, Examinatrice**  
**M. Olivier MERLIN, Directeur de thèse**  
**M. Sylvain FERRANT, Co-directeur de thèse**



# Abstract

Sustainable water resources management is a major issue on a global scale in a context of over-exploitation due in particular to a growing population. It is also a crucial lever to cope with the consequences of global warming which leads to an increase in climatic, hydrological and agricultural droughts, particularly in arid and semi-arid regions. Difficulties currently lie in the lack of knowledge of the resources available in the different hydrological compartments (surface water, unsaturated soil zone and groundwater), and this at multiple scales (surface reservoirs, irrigated areas, aquifer, region).

The objective of the thesis is to develop innovative methods to better quantify water resources from multi-sensor remote sensing. This work focuses on the state of Telangana (114,800 km<sup>2</sup>) in India, characterized by a granitic aquifer with low storage capacity, whose recharge during the monsoons is as variable as the annual pumping for irrigation. In this region, 70% of irrigation comes from groundwater (GWS). The remaining 30% comes from 1) a monsoon runoff harvesting system (RHS) in upstream areas, made up of thousands of small reservoirs (called “tanks”) and 2) large dams, canals and perennial rivers in downstream areas. We propose to make better use of multi-sensor/multi-resolution remote sensing to ensure multi-scale monitoring of the RHS and GWS and thus assess the impact of irrigation on the resource.

A first axis of this thesis evaluates the potential of GRACE gravity and SMOS soil moisture data for the qualitative monitoring of the evolution of resources in Telangana. GRACE has provided since 2002 maps of monthly total water stock anomalies at a resolution of 300 km. The groundwater stock (GWS) obtained by subtracting the contributions of surface and soil reservoirs is well correlated with aggregated piezometric measurements at the Telangana scale ( $R^2=0.70$ ). Concerning SMOS data, the observation of a humidity peak in the dry season reveals the impact of irrigation and its amplitude even constitutes an indicator of the quantity of rice cultivated ( $R=0.81$ ).

The second axis concerns the spatial downscaling of GRACE data for monitoring the GWS at a scale (50 km) more relevant for the management of territories. Several statistical and semi-empirical methods are proposed based on two hypotheses (GWS variations with or without interannual trend). In order to evaluate these methods, a new performance metric for downscaled products is developed. This validation approach estimates the spatial gain compared to non-downscaled data and here highlights the interest of empirical methods calibrated from GRACE data and auxiliary data (NDVI, soil moisture, precipitation).

The third and last axis of the thesis exploits the capabilities of high-resolution images for RHS monitoring through two applications : 1) RHS volume monitoring in near-real-time with Pléiades digital elevation models (DEM) at a 50 cm resolution and water maps derived from Sentinel-2 at 10 m, and 2) the quantification of the capacity of the RHS at large-scale and the assessment of its contribution to the total surface water stock. These applications demonstrate the interest of finely resolved DEM for the establishment of drought warning systems and a better understanding of the impact of the RHS on recharge.



# Résumé

La gestion durable des ressources en eau est un enjeu majeur à l'échelle mondiale dans un contexte de surexploitation liée notamment à une population croissante. C'est aussi un levier crucial pour faire face aux conséquences du réchauffement global qui entraîne une multiplication des sécheresses climatiques, hydrologiques et agricoles, particulièrement dans les régions arides et semi-arides. Des difficultés résident actuellement dans le manque de connaissance des ressources disponibles dans les différents compartiments hydrologiques (eaux de surface, zone du sol non saturée et eaux souterraines), et ceci à des échelles multiples (réservoirs de surface, périmètres irrigués, aquifères, régions).

L'objectif de la thèse est de développer des méthodes innovantes permettant de mieux quantifier la ressource en eau à partir de la télédétection multi-capteurs. Ces travaux se focalisent sur l'État du Telangana (114 800 km<sup>2</sup>) en Inde, caractérisé par un aquifère granitique de faible capacité de stockage, dont la recharge lors des moussons est aussi variable que le pompage annuel pour l'irrigation. Dans cette région, l'irrigation provient à 70% de l'eau souterraine (GWS). Les 30% restants proviennent 1) en amont d'un système de récupération des eaux de ruissellement de la mousson (RHS), composé de milliers de petits réservoirs (appelés "tanks") et 2) en aval de grands barrages, des canaux et rivières pérennes. Nous nous proposons de tirer le meilleur parti de la télédétection multi-capteurs/multi-résolutions pour assurer un suivi multi-échelle du RHS et du GWS et ainsi évaluer l'impact de l'irrigation sur la ressource.

Un premier axe de cette thèse évalue le potentiel des données gravimétriques GRACE et d'humidité du sol SMOS pour le suivi qualitatif de l'évolution des ressources sur le Telangana. GRACE fournit depuis 2002 des cartes d'anomalies mensuelles de stock d'eau total à une résolution de 300 km. Le stock d'eau souterraine (GWS) obtenu en soustrayant les contributions des réservoirs de surface et du sol est bien corrélé avec les mesures piézométriques agrégées à l'échelle du Telangana ( $R^2=0.70$ ). Concernant les données SMOS, l'observation d'un pic d'humidité en saison sèche révèle l'impact de l'irrigation et son amplitude constitue même un indicateur précoce de la quantité de riz cultivé ( $R=0.81$ ).

Le deuxième axe porte sur la désagrégation spatiale des données GRACE pour le suivi du GWS à une échelle (50 km) plus pertinente pour la gestion des territoires. Plusieurs méthodes statistiques et semi-empiriques sont proposées à partir de deux hypothèses (variations du GWS avec ou sans tendance interannuelle). Dans l'objectif d'évaluer ces méthodes, une nouvelle métrique de performance des produits désagrégés est développée. Cette approche de validation estime le gain spatial par rapport aux données non-désagrégées et met ici en évidence l'intérêt des méthodes empiriques calibrées à partir des données GRACE et de données auxiliaires (NDVI, humidité du sol, précipitations).

Le troisième et dernier axe de la thèse exploite la capacité des images haute résolution pour le suivi du RHS à travers deux applications : 1) le suivi volumétrique en quasi-temps réel du RHS avec des modèles numériques de terrain (MNT) Pléiades à 50 cm de résolution et des cartes en eau dérivées de Sentinel-2 à 10 m, et 2) la quantification de la capacité du RHS à large échelle et l'estimation de sa contribution au stock total d'eau de surface. Ces applications démontrent l'intérêt des MNT à résolution fine pour la mise en place de systèmes d'alerte de sécheresse et une meilleure compréhension de l'impact du RHS sur la recharge.



# Remerciements

On entend des histoires parfois inquiétantes sur les thèses, et l'idée d'en faire une ne me séduisait pas. Lorsque l'opportunité s'est présentée, je me suis tout de même lancée dans l'aventure en prenant certaines précautions, notamment la déconstruction de cette vision de la thèse comme l'aboutissement d'un travail d'une vie, synonyme de sacrifice, de douleur et d'épuisement. Mais considérer la thèse comme un CDD de trois ans avec à la fin un rapport de quelques centaines de pages à rédiger et un joli titre décerné ne suffit pas à la rendre facile, et ce fut tout de même trois ans mouvementés, que je n'aurais pas vécu de la même façon (certainement moins bien) sans les personnes qui m'ont entourées et soutenues pendant cette période.

Je souhaite d'abord remercier mes superviseurs pour m'avoir donné cette opportunité et qui m'ont toujours soutenu même dans mes plus grands moments de frustrations : d'abord Olivier pour sa formidable pédagogie, ses idées brillantes et son optimisme à tout épreuve (même celle de mon pessimisme), puis Sylvain, pour avoir manifesté tant d'enthousiasme à l'idée de travailler avec moi en stage puis en thèse, et pour son soutien professionnel et personnel tout au long de ces trois ans. Je pense que j'ai eu l'un des meilleurs encadrements de thèse dont on puisse rêver.

Par ailleurs, je tiens à remercier l'ensemble des membres de mon jury pour leur intérêt porté à ces travaux, et en particulier les trois rapporteurs de ce manuscrit - Benoît Dewandel, Laurent Longueverne et Thierry Pellarin, pour leur lecture attentive et leurs retours tant positifs que pertinents.

Je voudrais également remercier mes collègues du CESBIO qui ont contribué à la construction d'un environnement de travail agréable : Erwan, Maël, Roiya, Pierre, Luis, Gaëtan, Marine, Jordi, Laurent, Micaël, et mes collègues de bureau Sylvain et Vincent qui, eux aussi, ont dû supporter ma mauvaise humeur dans les pires moments et mon trop plein d'enthousiasme dans les meilleurs. En particulier, merci aux doctorants du CESBIO, de Toulouse et du monde entier avec qui les échanges me rappelaient que nous étions toute une communauté solidaire à faire face aux mêmes problèmes : Robin à Toulouse, Jérémy à Paris avec qui j'ai commencé l'aventure toulousaine, Alexis à Tours, toujours passionné et passionnant, Eliott au Japon, un génie incompris, la brillante Judith à Los Angeles, Bertille en Finlande avec qui je partagerai toujours les couleurs de la vie, et Olivia en Nouvelle-Zélande, toujours présente malgré la distance d'une demi-circonférence terrestre qui nous sépare.

Je dois également beaucoup à tous mes amis de Toulouse : Robin, Arthur, Coline, Thomas, Andy, Célia, Gabriel et Laure qui ont très clairement contribué à mon bonheur pendant ces trois ans. Merci également à tous mes camarades de danse, qui ne savent probablement pas de quoi parle ma thèse ni que j'en fais une, mais que j'étais contente d'avoir près de moi malgré tout !

J'ai aussi une pensée pour mes anciennes collègues de la Maison de la Télédétection à Montpellier qui ont vu mes premiers pas dans cette discipline, Hélène et Lou, que je ne vois pas souvent mais que je porte dans mon cœur !

Merci aussi à mes amis de prépa et d'école qui sont toujours là : Camille, Margot, Martial (pour ne citer qu'eux), et bien sûr Inès, toujours prête à échanger des bêtises avec moi et partir faire n'importe quoi, n'importe où, et n'importe quand.



Merci enfin à ma famille : d'abord à mes sœurs pour m'avoir hébergé tant de fois au pied levé à Paris pour jouer aux jeux de société en buvant du thé, et ensuite à ma maman, une personne exceptionnelle que j'estime énormément, et qui, à ma plus grande surprise, s'est toujours souvenue des thèmes de recherche de ma thèse alors même que j'en oubliais fréquemment l'intitulé. Si j'en suis là aujourd'hui, c'est parce qu'elle m'a toujours soutenue et toujours cru en moi.

Et parce que le contenu des remerciements varie énormément avec le temps qui passe et le moment où on les écrit, et que je ne veux oublier personne ni négliger l'importance que certaines personnes ont ou ont eu pour moi : merci à tous ceux que je n'ai pas cité mais qui m'ont soutenue et font partie de ma vie ou en ont fait partie au cours de ces 3 ans.

# Table des matières

<b>1</b>	<b>Introduction générale</b>	<b>3</b>
1.1	Besoins en eau . . . . .	3
1.1.1	Agriculture et ressources en eau en Inde . . . . .	4
1.1.2	Une pression croissante sur les ressources en eau . . . . .	6
1.2	Ressources en eau et leur gestion dans l'État du Telangana . . . . .	7
1.2.1	Des ressources en eau limitées par la géologie des aquifères . . . . .	7
1.2.2	Usages de l'eau pour l'irrigation . . . . .	9
1.2.3	Suivi local des ressources en eau . . . . .	10
1.3	Apport de la télédétection spatiale pour le suivi des ressources en eau . . . . .	11
1.3.1	Satellites d'observation de la Terre . . . . .	11
1.3.2	Estimation de variables environnementales à partir de données satellitaires . . . . .	12
1.4	Position du problème . . . . .	16
1.4.1	Incertitudes de la connaissance des stocks d'eau à toutes les échelles . . . . .	16
1.4.2	Capacités de la télédétection . . . . .	17
1.4.3	Objectifs et organisation du manuscrit . . . . .	18
<b>2</b>	<b>Données</b>	<b>21</b>
2.1	Introduction . . . . .	22
2.2	Données <i>in situ</i> . . . . .	23
2.2.1	Mesures piézométriques . . . . .	23
2.2.2	Suivi des petits réservoirs . . . . .	23
2.2.3	Occupation du sol . . . . .	25
2.3	Données satellitaires . . . . .	27
2.3.1	Variation du stock d'eau total (GRACE) . . . . .	27
2.3.2	Humidité du sol . . . . .	33
2.3.3	Indice de végétation et évapotranspiration MODIS . . . . .	34
2.3.4	Eau de surface . . . . .	34
2.3.5	Précipitations TRMM . . . . .	36
2.4	Modèles . . . . .	37
2.4.1	GLDAS . . . . .	37
2.4.2	GLEAM . . . . .	38
2.4.3	Évapotranspiration potentielle . . . . .	38
2.5	Ce qu'il faut retenir . . . . .	39
<b>3</b>	<b>Potentiel des données GRACE et SMOS pour le suivi des stocks d'eau souterraine</b>	<b>41</b>
3.1	Introduction . . . . .	41

3.2	Homogénéisation des données . . . . .	43
3.2.1	Résolution spatiale et temporelle, projection . . . . .	43
3.2.2	Unité . . . . .	43
3.2.3	Spatialisation des données piézométriques . . . . .	43
3.3	Évolution des ressources en eau à l'échelle du Telangana . . . . .	44
3.3.1	Interprétation qualitative des séries temporelles . . . . .	44
3.3.2	Tendances à long terme . . . . .	45
3.4	Déconvolution de GRACE à l'échelle du Telangana . . . . .	47
3.4.1	Équation générale de bilan d'eau . . . . .	47
3.4.2	Sensibilité du GWS estimé à différents jeux de données . . . . .	49
3.4.3	Discussion . . . . .	51
3.5	Potentiel de SMOS . . . . .	52
3.5.1	Un cycle annuel d'humidité bimodal . . . . .	52
3.5.2	Interprétation à l'échelle de SMOS . . . . .	54
3.5.3	Quantification de l'eau utilisée pour l'irrigation avec SMOS . . . . .	57
3.5.4	Indicator of flood-irrigated crops from SMOS and SMAP soil moisture products in southern India . . . . .	59
3.6	Ce qu'il faut retenir . . . . .	65
<b>4</b>	<b>Désagrégation des données GRACE</b>	<b>67</b>
4.1	Introduction . . . . .	68
4.2	Problématiques et applications sur le Telangana . . . . .	68
4.2.1	Problématiques . . . . .	68
4.2.2	Comparaison de différents modèles . . . . .	70
4.2.3	Limites et perspectives . . . . .	77
4.3	Evaluating downscaling methods of GRACE data : a case study over a fractured crystalline aquifer in South India . . . . .	79
4.4	Ce qu'il faut retenir . . . . .	98
<b>5</b>	<b>Cartographie des volumes d'eau de surface avec Pléiades et Sentinel-2</b>	<b>99</b>
5.1	Introduction . . . . .	100
5.2	Quantification de la capacité du RHS et suivi temporel avec des images haute résolution . . . . .	101
5.2.1	État de l'art . . . . .	101
5.2.2	Méthodes et application sur le RHS du Telangana . . . . .	102
5.2.3	Résultats . . . . .	107
5.3	Monitoring the volume of South Indian Rainwater Harvesting System : sensitivity of the bathymetry retrieval to DEM features . . . . .	110
5.4	Discussion et perspectives . . . . .	136
5.4.1	Précision du MNT . . . . .	136
5.4.2	Précision des surfaces en eau détectées sur Sentinel-2 . . . . .	136
5.4.3	Délimitation des tanks . . . . .	137
5.4.4	Futures applications . . . . .	138
5.5	Ce qu'il faut retenir . . . . .	139

<b>6 Conclusions et perspectives</b>	<b>141</b>
6.1 Potentiel des données GRACE et SMOS/SMAP pour le suivi multi-échelle des res- sources en eau . . . . .	141
6.2 Désagrégation des données GRACE pour le suivi de l'eau souterraine à haute résolution	142
6.3 Suivi du volume d'eau de surface avec des données très haute résolution . . . . .	144
6.4 Sur la synergie des données satellitaires multicapteurs . . . . .	145
<b>Acronymes</b>	<b>147</b>
<b>Bibliographie</b>	<b>169</b>



# Table des figures

1.1	Évolution des surfaces irriguées par source en Inde. . . . .	4
1.2	Distribution spatiale des puits d’observation ( <i>Groundwater Department</i> <a href="https://gwd.telangana.gov.in/">https://gwd.telangana.gov.in/</a> ) dans l’État du Telangana. . . . .	7
1.3	Distribution (médianes et quartiles) des précipitations (TRMM) par mois entre 2002 et 2020. . . . .	8
1.4	Évolution des surfaces irriguées par source dans le Telangana. . . . .	9
1.5	Exemples de variables environnementales extraites d’images satellitaires et séries temporelles à l’échelle du Telangana. . . . .	13
1.6	Détection d’eau sur une image Sentinel-2. . . . .	15
2.1	Localisation des tanks étudiés dans Perrin et al. (2012) et volumes et surfaces mesurés.	25
2.2	Emplacement des parcelles labellisées pour la saison humide (Kharif) 2019 et des tuiles Sentinel-2. . . . .	26
2.3	Rappel et précision des cartes d’occupation du sol. . . . .	27
2.4	Trace au sol de GRACE et compromis entre résolution spatiale et temporelle (Schmidt et al., 2008). . . . .	30
2.5	Localisation des MNT Pléiades. . . . .	36
3.1	Séries temporelles de pluie, TWS et GWS sur le Telangana. . . . .	45
3.2	Comparaison du TWS de GRACE dans l’Inde du Nord et du Sud. . . . .	47
3.3	Anomalies de RZSM (période 2015-2020) et précipitations. . . . .	50
3.4	GWSA déconvolué et GWSA <i>in situ</i> . . . . .	51
3.5	Humidité de surface des données SMOS, SMAP, CCI, GLDAS-NOAH et GLEAM .	53
3.6	Séries temporelles de NDVI, d’humidité de surface (SSM) de SMOS et de pluie sur la zone granitique du Telangana. . . . .	53
3.7	Humidité de surface de SMOS sur différentes régions d’Inde. . . . .	54
3.8	Grille SMOS et occupation du sol. . . . .	55
3.9	Séries temporelles de NDVI, SSM SMOS, GWL et occupation du sol sur deux pixels de $25 \times 25$ km <sup>2</sup> . . . . .	56
3.10	Proportion de riz en Rabi, amplitude du pic de NDVI en saison sèche, GWLR et GWLK en fonction du pic de SSM SMOS en saison sèche. . . . .	58
4.1	Logigramme de la désagrégation empirique statique. . . . .	72
4.2	Schéma explicatif du calcul des gains spatiaux et temporels. . . . .	74
4.3	Distribution des gains temporels et spatiaux. . . . .	74
4.4	Distribution des gains temporels et spatiaux de $\Delta TWS_{HR}$ . . . . .	76
4.5	Distribution des gains temporels et spatiaux de TWS pour le modèle dynamique. . .	77
4.6	Série temporelle du GWS désagrégé avec réinitialisation tous les 4 ans. . . . .	78

5.1	Étapes de l'estimation du volume avec Pléiades et Sentinel-2. . . . .	103
5.2	Localisation de la zone d'étude et inventaire des données. . . . .	104
5.3	Distribution du MWAE des tanks pour chaque emprise Pléiades et pour l'ensemble des emprises (abcd). . . . .	105
5.4	Courbes hypsométriques d'un tank délimité par Sentinel-2 et une isoligne Pléiades. .	106
5.5	Différence entre les MWAE dérivées de Landsat, Sentinel-2 et la délimitation par une isoligne Pléiades. . . . .	108
5.6	Quantification de la capacité théorique du RHS à l'échelle du Telangana. . . . .	108
5.7	Séries temporelles de volume d'eau à l'échelle des quatre emprises Pléiades. . . . .	109
5.8	Localisation et bathymétries des quatre tanks de Perrin et al. (2012). . . . .	137
5.9	Distribution des erreurs des surfaces en eau Sentinel-2 et volumes dérivés par rapport aux images optiques très haute résolution. . . . .	137

# Liste des tableaux

2.1	Récapitulatif des données utilisées. . . . .	24
3.1	Tendances annuelles calculées sur les séries de TWS de GRACE dans la littérature. . . . .	46
3.2	R et $R^2$ entre les données <i>in situ</i> et le GWSA estimé par déconvolution pour différents jeux de données. . . . .	50
3.3	Plages de valeurs de R et $R^2$ entre le GWS <i>in situ</i> et issu de la déconvolution de GRACE dans la littérature. . . . .	52
3.4	Corrélations de Pearson entre l’amplitude du pic d’humidité de SMOS en saison sèche, l’occupation du sol, le NDVI et les données piézométriques. . . . .	57
4.1	Sélection de variables et performances des modèles prédictifs de TWS à basse résolution. . . . .	73
4.2	Sélection de variables et performances des modèles prédictifs de $\Delta$ TWS à LR. . . . .	75
5.1	Étendue d’eau maximale observée (MWAE) et capacité correspondante dérivé de GSW et Sentinel-2. . . . .	107





# Chapitre 1

## Introduction générale

### Sommaire

---

<b>1.1 Besoins en eau</b>	<b>3</b>
1.1.1 Agriculture et ressources en eau en Inde	4
1.1.2 Une pression croissante sur les ressources en eau	6
<b>1.2 Ressources en eau et leur gestion dans l'État du Telangana</b>	<b>7</b>
1.2.1 Des ressources en eau limitées par la géologie des aquifères	7
1.2.2 Usages de l'eau pour l'irrigation	9
1.2.3 Suivi local des ressources en eau	10
<b>1.3 Apport de la télédétection spatiale pour le suivi des ressources en eau</b>	<b>11</b>
1.3.1 Satellites d'observation de la Terre	11
1.3.2 Estimation de variables environnementales à partir de données satellitaires	12
<b>1.4 Position du problème</b>	<b>16</b>
1.4.1 Incertitudes de la connaissance des stocks d'eau à toutes les échelles	16
1.4.2 Capacités de la télédétection	17
1.4.3 Objectifs et organisation du manuscrit	18

---

### 1.1 Besoins en eau

Beaucoup de régions du monde souffrent d'une pénurie d'eau croissante, à la fois pour l'irrigation des cultures et les usages domestiques en zones urbaines (Wada et al., 2011). Dans un contexte de changement global où la population ne cesse d'augmenter et les épisodes de sécheresses se multiplient et s'intensifient, la gestion de l'eau douce pour les activités humaines est essentielle pour subvenir aux besoins de la population mondiale. En particulier, l'irrigation est le secteur le plus consommateur en eau, particulièrement dans les climats arides et semi-arides où elle représente 70% des prélèvements d'eau douce et 90% des usages en eau (Siebert et al., 2010). Les besoins en eau ne font qu'augmenter depuis plus d'un demi-siècle, en particulier avec les changements profonds apportés à l'agriculture dans le cadre de la "révolution verte" ayant suivi les crises alimentaires des années 1960. Durant cette période, plusieurs pays asiatiques ont réformé leur politique agricole pour tendre vers une agriculture intensive leur permettant de faire face aux pénuries et assurer

leur autonomie alimentaire, avec l'adoption de variétés à fort rendement plus gourmandes en eau, l'utilisation de pesticides et la construction d'aménagements hydrauliques conséquents pour développer l'irrigation. En particulier, l'Inde est un cas d'école en terme de développement rapide de la demande en eau pour l'irrigation et de surexploitation des ressources.

### 1.1.1 Agriculture et ressources en eau en Inde

En Inde, les transformations de la révolution verte ont été particulièrement spectaculaires, puisque la proportion de surfaces irriguées sur le territoire indien a augmenté de 4% en 1951 à 20% en 2016 (d'après le rapport annuel 2020-2021 de la *Central Water Commission* - CWC, <http://cwc.gov.in>, dernier accès le 22 juin 2022). Les méthodes d'irrigation ont également drastiquement évolué, puisque l'utilisation de l'eau souterraine s'est développée jusqu'à supplanter l'irrigation par l'eau de surface (par des canaux reliant les grands lacs de barrage aux exploitations agricoles, ou par des petits réservoirs appelés "tanks"), avec une consommation passant de 10-20 km<sup>3</sup> en 1950 à 250 km<sup>3</sup>/an en 2010 (Siebert et al., 2010), faisant aujourd'hui de l'Inde le plus grand consommateur d'eau souterraine du monde (voir figure 1.1).

Les 1105 mm de précipitations annuelles constituent une ressource en eau potentielle de 1123 milliards de mètres cubes (BCM - *billion cubic meters*) utilisables, soit 342 mm, répartis de manière hétérogène sur le territoire indien, dont 690 BCM (201 mm) d'eau de surface et 433 BCM (132 mm) d'eau souterraine (source : CWC). Depuis des décennies, de nombreuses infrastructures de taille variée sont continuellement construites pour optimiser l'exploitation de cette ressource potentielle en surface et souterraine pour l'irrigation.

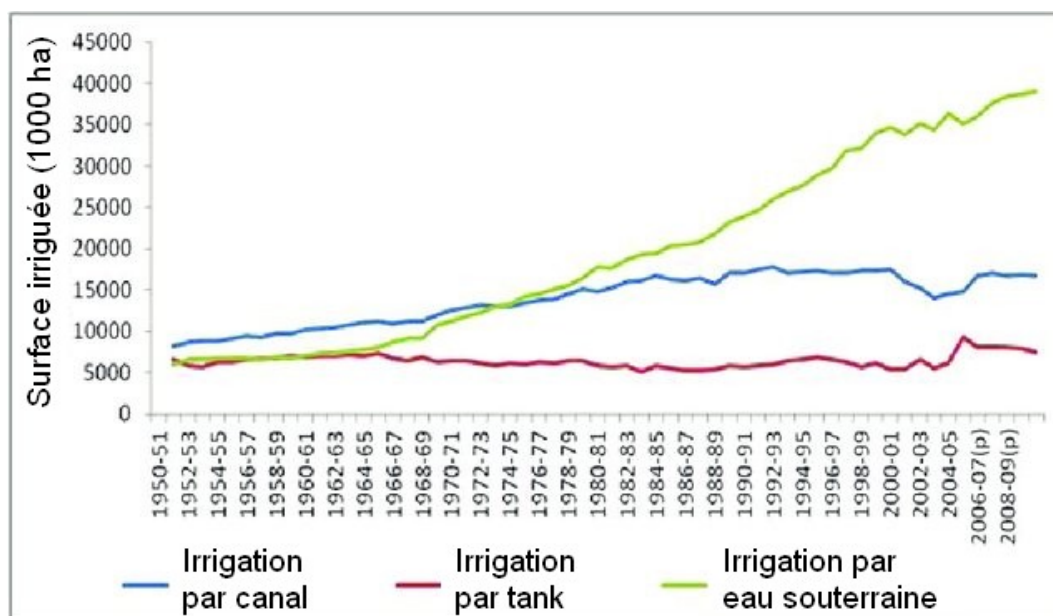


FIGURE 1.1 – Évolution des surfaces irriguées par source en Inde (modifié à partir de Ali (2012), source : *Agricultural Statistics, Government of India*).

#### 1.1.1.1 Eau souterraine

L'eau souterraine est une des plus importantes ressources d'eau douce, puisque contrairement à l'eau de surface, elle est mieux protégée de la pollution, moins sensible aux fluctuations saisonnières,

et distribuée plus uniformément sur des zones larges. Ces caractéristiques contribuent à en faire une ressource pouvant facilement être prélevée sur place à l'échelle individuelle, limitant les pertes par transport ainsi que la concurrence et les négociations liées à une ressource moins visible et moins comprise que l'eau de surface.

Faciliter l'accès aux ressources en eau souterraine par l'électrification des zones rurales a été un aspect phare des politiques agricoles dans de nombreux États indiens, encourageant les agriculteurs à forer des puits et à les équiper de pompes pour prélever l'eau souterraine à proximité de leurs exploitations agricoles. Dans les années 2000 en Inde, 45% de l'eau souterraine était utilisée pour l'irrigation et 9% pour les usages domestiques (Glendenning et al., 2012), et son exploitation s'est intensifiée ces dernières décennies avec l'émergence de politiques instaurant l'électricité gratuite pour les agriculteurs, notamment dans les États du Tamil Nadu, du Punjab et du Telangana (Aggarwal et al., 2020). Ces politiques ont été fortement promues dans l'État du Telangana, devenant un argument majeur lors des campagnes électorales, et n'ont cessé de progresser : 6 heures d'électricité par jour sont offertes aux agriculteurs en 2004, 7 heures en 2014, 9 heures en 2016, et en 2018 le Telangana est le premier État indien à fournir gratuitement l'électricité 24h/24 aux agriculteurs (Fosli et al., 2021). Cette disponibilité croissante de l'électricité a soutenu la surexploitation des aquifères, menant à l'adoption de cultures plus gourmandes en eau, notamment le riz (Badiani and Jessoe, 2013) et la mise en place de deux, voire trois saisons culturales par an. Aujourd'hui, l'exploitation des eaux souterraines se fait à l'échelle individuelle de manière beaucoup moins structurée que la gestion de l'eau de surface et s'organise généralement en dehors de toute gestion institutionnalisée.

### 1.1.1.2 Eau de surface

L'eau de surface demeure une composante importante de l'irrigation en Inde. De grandes infrastructures d'irrigation sont continuellement construites pour optimiser l'exploitation de la ressource potentielle en surface par le stockage des pluies de moussons dans des réservoirs de taille variée et leur transport pour répartir au mieux les ressources sur le territoire.

De grands barrages sont construits le long des rivières pour stocker les eaux de surface, qui sont acheminées par canaux à proximité des exploitations agricoles. En 2021, le territoire indien comptait plus de 5200 grands réservoirs (ayant une hauteur supérieure à 15 m) répartis entre les différents États et plus de 400 en construction (source : *National Register of Large Dams*, CWC, <http://cwc.gov.in/national-register-large-dams> ; dernier accès le 22 juin 2022). La capacité cumulée des infrastructures stockant l'eau en surface serait de 253 BCM (77 mm), avec 155 BCM (47 mm) supplémentaires potentiels pour les infrastructures en construction (source : CWC).

Une autre stratégie adoptée par les États indiens pour lutter contre la sécheresse est la construction d'un grand nombre de petits réservoirs (appelés "tanks") pour collecter l'eau des précipitations, formant un système de récupération des eaux de pluie (RHS, *Rainwater Harvesting System*). La multiplication de ce type de petits réservoirs, en particulier en amont des rivières, est une stratégie commune pour lutter contre la sécheresse (Giordano and Villholth, 2007; Krol et al., 2011). Ces tanks, en plus de maintenir une source d'eau accessible pour le bétail et de participer à l'irrigation des cultures dans une moindre mesure (figure 1.1), sont promus en tant que tanks de percolation permettant d'augmenter la recharge de l'aquifère. Cependant, l'efficacité de ce système de gestion de la recharge de l'aquifère (MAR, *Managed Aquifer Recharge*) est remise en question car les pertes associées à l'évaporation pourraient être importantes (Boisson et al., 2014; Massuel et al., 2014; Perrin et al., 2012), et son impact pourrait être minime par rapport à la recharge par la pluie et

les flux de retour de l'irrigation (Brauns et al., 2022).

Le RHS indien est composé de plusieurs centaines de milliers de tanks répartis le long d'un réseau de drainage non pérenne. Le nombre de tanks fonctionnels en Inde est estimé entre 200 000 (Palanisami et al., 2010; Sharma, 2003) et 500 000 (Reddy, 2015). Plusieurs programmes de réhabilitation de ces tanks sur le territoire indien visent à améliorer la recharge en retirant le dépôt de sédiments, montrant l'importance socio-politique du RHS face aux besoins grandissants en eau. Pourtant, les tanks du RHS demeurent encore très peu documentés. Contrairement aux grands lacs de barrage inventoriés et suivis par le CWC, ils ne font pas l'objet d'un inventaire exhaustif ni d'un suivi volumétrique du fait de leur grand nombre et de leur petite taille, et leurs impacts dans les études hydrologiques à large échelle sont souvent négligés ou grossièrement estimés, alors qu'ils pourraient être significatifs (Habets et al., 2018).

### 1.1.2 Une pression croissante sur les ressources en eau

D'importants aménagements sont continuellement mis en place pour répartir l'eau de surface sur le territoire et augmenter le potentiel d'irrigation, comme le témoignent les 400 grands réservoirs en construction. Cependant, ils engendrent des coûts économiques et humains importants, ainsi qu'une compétition entre les régions plus ou moins pourvues et des négociations complexes avec les autres États situés en aval. En particulier, c'est dans le Telangana qu'a été lancé le projet d'irrigation le plus ambitieux du monde, le projet Kaleshwaram. Devant être achevé en 2026, il vise à alimenter plus de 20 barrages avec l'eau prélevée dans la rivière Godavari sur plus de 300 km, afin d'irriguer le centre de l'État et d'alimenter en eau potable Hyderabad, capitale du Telangana. Cette infrastructure gigantesque a nécessité la délocalisation de milliers de personnes, pour un coût total estimé à 12 milliards d'euros.

Concernant les ressources souterraines, les prélèvements sur les aquifères sont la somme de milliers de décisions individuelles, dont l'absence d'organisation entrave leur gestion. Ceci est particulièrement problématique dans le contexte indien où les précipitations n'ont lieu que pendant les quelques mois de la mousson et doivent être gérées de manière à soutenir les besoins en eau jusqu'à la saison des pluies suivante.

Dans plusieurs régions du monde, l'infiltration de la pluie dans les roches poreuses du sol ne suffit plus à compenser les prélèvements anthropiques, ce qui se traduit par une baisse chronique du niveau des nappes, en particulier dans les aquifères sédimentaires profonds du nord de l'Inde, dont le déclin, estimé entre 2 et 3 cm/an, est le plus spectaculaire observé au monde (Asoka et al., 2017; Rodell et al., 2018; Tiwari et al., 2009). Cependant, certaines régions du monde, comme les deux tiers sud de l'Inde, reposent sur des roches cristallines à faible porosité ne constituant pas de réservoir d'eau massif comme dans le nord de l'Inde. La gestion des usages de l'eau dans le sud de l'Inde est d'autant plus essentielle, car les besoins en eau non satisfaits par les précipitations annuelles ne peuvent pas être compensés par les ressources anciennes. C'est notamment le cas du Telangana où la ferveur des politiques agricoles en faveur des agriculteurs ont entraîné une surexploitation massive des ressources en eau souterraine, si bien que les puits s'assèchent progressivement et les coupures d'électricité liées à une consommation croissante due au pompage de l'eau à des profondeurs toujours plus grandes se multiplient.

## 1.2 Ressources en eau et leur gestion dans l'État du Telangana

### 1.2.1 Des ressources en eau limitées par la géologie des aquifères

Le Telangana est un État indien de 114 800 km<sup>2</sup> densément peuplé (environ 335 habitants par km<sup>2</sup> selon l'UIDAI (Unique Identification Authority of India (UIDAI)) et fortement irrigué, situé dans la partie sud de l'Inde (figure 1.2). Il est dominé par un climat tropical sec/semi-aride où une saison humide, aussi appelée Kharif, alterne avec une saison sèche, aussi appelée Rabi. Les précipitations de la mousson ont lieu entre juillet et octobre (comme le montre la distribution des pluies par mois sur la figure 1.3) et peuvent représenter de 540 à 1300 mm par an, avec une moyenne de 879 mm (*Indian Meteorological Department*, <https://mausam.imd.gov.in/>, dernier accès le 13 juin 2022).

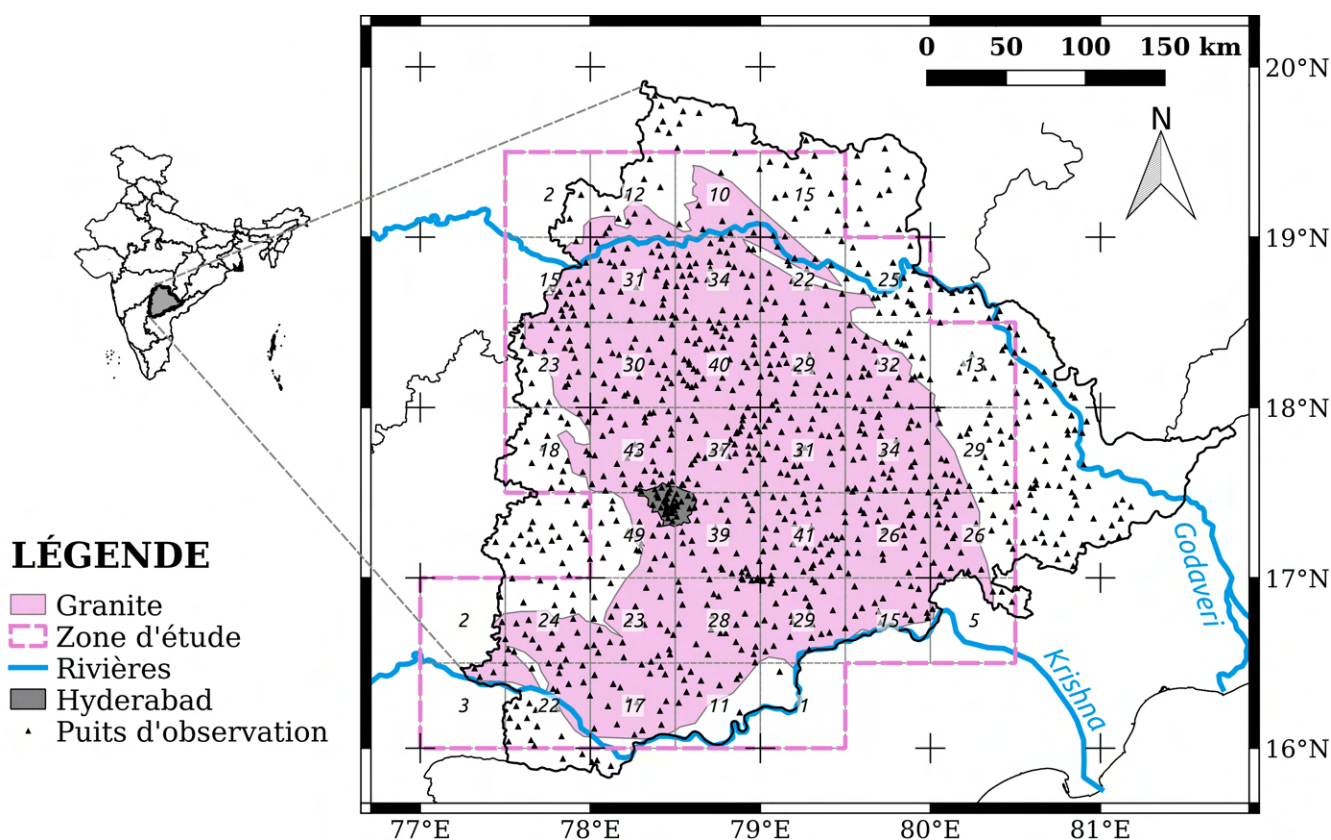


FIGURE 1.2 – Distribution spatiale des puits d'observation (*Groundwater Department* <https://gwd.telangana.gov.in/>) dans l'État du Telangana.

L'État du Telangana est composé d'une formation granitique sur 67 000 km<sup>2</sup>, soit 60% de sa superficie. La gestion et la conservation de l'eau souterraine est particulièrement difficile dans ce type d'environnement où les faibles porosité et conductivité hydraulique des roches cristallines empêchent un stockage important dans le sol. En effet, ces roches cristallines (aussi qualifiées de plutoniques) sont formées à partir d'un magma refroidi dans des conditions de température et de pression permettant une cristallisation complète de la roche. Du fait de la disposition contiguë de leurs cristaux, les roches cristallines possèdent très peu d'espace intergranulaire, et sont associées à des porosités primaires et perméabilités négligeables ne leur permettant pas de contribuer au

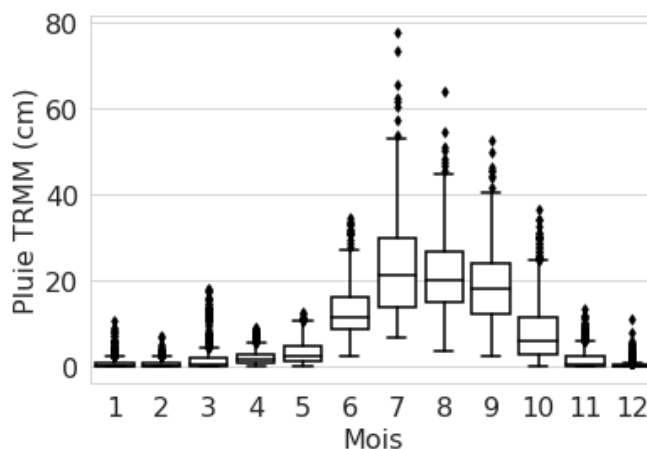


FIGURE 1.3 – Distribution (médianes et quartiles) des précipitations (TRMM) par mois entre 2002 et 2020.

stockage de l'eau souterraine. Le stockage et la circulation de l'eau n'est possible que dans les roches cristallines altérées (altérites) ou fracturées par des altérations physiques ou chimiques. Le stockage de l'eau dans ces fractures (désignant les discontinuités dans la roche) de dimensions variées (moins d'un millimètre à plusieurs centaines de kilomètres) est associé à une porosité "secondaire" de la roche (car non intrinsèque) (Nicolas, 2019), formant des aquifères cristallins.

Dans les milieux tectoniquement stables, l'altération de la roche cristalline exposée pendant une très longue période induit sa stratification en plusieurs couches aux propriétés hydrauliques variées en fonction de leur stade d'altération. Sur la zone granitique du Telangana, le profil d'altération type est composé de trois niveaux principaux, de la plus superficielle à la plus profonde (Dewandel et al., 2006) :

- une couche de saprolite argilo-sableuse ou sablo-argileuse issue de la décomposition du socle granitique, épaisse d'une dizaine de mètres et de forte porosité efficace ( $Sy$  - *Specific yield* estimé à 10%) mais faiblement perméable (Dewandel et al., 2006; Maréchal et al., 2004). La porosité efficace d'une roche indique sa capacité de stockage de l'eau : elle est définie par le rapport du volume d'eau que peut stocker une roche puis libérer sous l'effet d'un drainage complet par rapport à son volume total,
- une couche de granite fissuré de faible porosité ( $Sy \simeq 1\%$ ) dont la densité des fractures diminue avec la profondeur (Dewandel et al., 2006; Maréchal et al., 2004),
- le socle non fracturé constitue la base de cet aquifère multicouche et n'est uniquement perméable localement qu'à l'emplacement des fractures tectoniques.

Les aquifères cristallins, en particulier les aquifères granitiques du Telangana, sont donc discontinus et hétérogènes. Du fait de leur structure verticale particulière, les dynamiques de flux souterrains varient en fonction des conditions de niveau d'eau. Lorsque le niveau d'eau est haut, l'eau circule à l'échelle du bassin versant dans la roche altérée (saprolite) et la couche supérieure de granite (la plus fracturée). En revanche, à faible niveau d'eau, les aquifères sont caractérisés par une compartimentation latérale du fait de la diminution du nombre de fractures dans le granite (Guihéneuf et al., 2014).

La complexité de ces aquifères, leur faible capacité de stockage, et leur faible recharge par les précipitations les rend particulièrement vulnérables à leur surexploitation. En effet, du fait de

leur faible porosité, les aquifères sont caractérisés par des fluctuations annuelles piézométriques très fortes avec l’alternance d’une saison humide aux précipitations intenses qui reconstituent plus ou moins les ressources, et d’une saison sèche où l’aridité et les prélèvements pour l’irrigation par pompage contribuent à les vider rapidement (Maréchal et al., 2006). Les aquifères du sud de l’Inde constitue donc une ressource en eau fragile, et les ressources annuelles allouées par la mousson doivent être précautionneusement gérées sans la possibilité de puiser dans les stocks anciens, comme cela peut-être fait dans les aquifères sédimentaires profonds du nord de l’Inde (Chidambaram et al., 2009; Singh et al., 2005).

## 1.2.2 Usages de l’eau pour l’irrigation

La demande en eau est très importante dans l’État du Telangana, tant pour les usages domestiques que pour l’agriculture. Les ressources en eau fournies par la mousson doivent généralement soutenir deux saisons culturales par an, de juin à octobre (Kharif) et de février à mai (Rabi). Comme pour le sous-continent indien (voir figure 1.1), l’irrigation sur l’État du Telangana s’appuie sur trois sources principales avec des proportions plus ou moins importantes ayant évolué ces dernières décennies : l’irrigation par l’eau souterraine, par l’eau des grands réservoirs acheminée à proximité des exploitations agricoles par des canaux et par les tanks du RHS (figure 1.4).

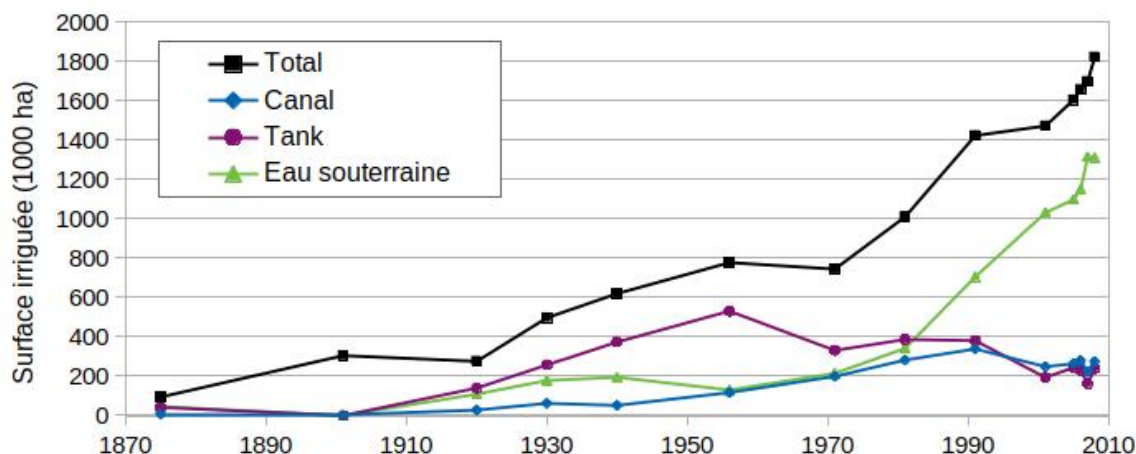


FIGURE 1.4 – Évolution des surfaces irriguées par source dans le Telangana (d’après Pingle (2011)).

La majeure partie des besoins en eau pour l’irrigation est assurée par le pompage de l’eau souterraine résultant de l’infiltration des pluies de mousson dans les aquifères, opéré de manière individuelle par les agriculteurs de manière déstructurée (Selles and Maréchal, 2022). L’exploitation de l’eau souterraine a considérablement augmenté depuis les années 1970, et alimentait en 2010 72% des terres irriguées soit 11% de la surface du Telangana (Pingle, 2011).

Dans une moindre mesure, l’eau stockée en surface joue également un rôle dans l’irrigation. La plupart des rivières ne sont pas pérennes et ne sont actives que quelques mois pendant et après la mousson, à l’exception des fleuves majeurs de la Godavari et la Krishna, et l’État du Telangana compte 184 grands barrages d’une capacité cumulée de 13 BCM, soit 113 mm (source : NRLD). Le Telangana abrite aussi le projet Kaleshwaram, un projet ambitieux qui vise à acheminer l’eau de la rivière Godavari dans 20 barrages sur plus de 300 km pour l’irrigation dans le centre de l’État et près de la capitale. La surface irriguée par canaux reliés aux grands réservoirs de barrage a connu une augmentation à partir des années 1950 avec la construction de nouveaux



réservoirs et d'infrastructures d'irrigation, et a légèrement régressé à partir des années 1990 avec la démocratisation de l'irrigation par l'eau souterraine, pour atteindre en 2010 15% des terres irriguées et 2% de la surface du Telangana.

Le Telangana dispose également d'un RHS composé de milliers de tanks. Un tank typique dans cet État est une dépression plate d'un kilomètre de long, avec une digue de 3 à 10 mètres de haut et de 20 mètres d'épaisseur, qui se vide presque entièrement pendant la saison sèche. L'irrigation par tank était majoritaire dans les années 1950, mais a rapidement décliné avec l'utilisation massive d'eau souterraine (figure 1.4), et n'était employée en 2010 que pour 13% des surfaces irriguées soit moins de 2% de la surface de l'État). Le RHS demeure cependant entretenu, comme le montre le programme Kakatiya lancé en 2015 initiant la construction de nouveaux tanks et la réhabilitation de 46 000 tanks afin d'en retirer le dépôt de sédiments pour favoriser la recharge. Comme à l'échelle de l'Inde, les tanks du RHS au sein de l'État du Telangana demeurent encore très peu documentés.

### 1.2.3 Suivi local des ressources en eau

C'est dans ce contexte de pression croissante sur la ressource en eau qu'ont été mis en place des réseaux de surveillance. Le CWC (*Central Water Commission*, <http://cwc.gov.in>, dernier accès le 22 juin 2022), créé en 1970 à partir d'un organisme existant depuis 1945, a pour mission la conservation et l'utilisation des ressources en eau à l'échelle nationale pour la gestion de l'irrigation, des inondations, de l'approvisionnement en eau potable et la production d'énergie hydraulique. Entre autre, le CWC possède un réseau de suivi des phénomènes hydro-météorologiques (ruissellement, précipitations, température, etc.) et réalise un suivi des projets d'irrigation grâce à un inventaire national de plus de 5200 lacs de barrage majeurs (NRLD - *National Registry of Large Dams*) et le suivi volumétrique de plus d'une centaine de grands réservoirs.

Les études sur la quantité et la qualité de l'eau souterraine à l'échelle nationale sont réalisées depuis les années 1970 par un organisme gouvernemental, le CGWB (*Central GroundWater Board*, <http://cgwb.gov.in/>), qui assure le suivi saisonnier du niveau des nappes sur un grand nombre (plus de 70 000) de puits d'observation répartis sur tout le territoire indien, dont 2700 sur le Telangana (<https://indiawris.gov.in/>, dernier accès le 13 juin 2022). Son équivalent télanganais, le *Groundwater Department* (<https://gwd.telangana.gov.in/>, dernier accès le 23 juin 2022) possède un réseau de plus de 1000 puits sur le Telangana relevés mensuellement.

En 1999, une collaboration franco-indienne a vu le jour entre le BRGM (Bureau de recherches géologiques et minières) et l'Institut Indien de Recherche Géophysique (NGRI - *National Geophysical Research Institute*) dans l'objectif de mieux comprendre les processus hydrologiques dans les roches cristallines fracturées. Les études réalisées dans le cadre du Centre Franco-Indien de Recherche sur les Eaux Souterraines (CEFIRES ou IFCGR - *Indo-French Center for Groundwater Research*), situé à Hyderabad, la capitale du Telangana, ont contribué à la compréhension des propriétés hydrauliques des aquifères sur plusieurs observatoires équipés (Dewandel et al., 2006, 2010, 2012, 2011, 2017a,b; Maréchal et al., 2004, 2006, 2018; Nicolas, 2019) ainsi qu'à celles du RHS et de son impact sur la recharge (Boisson et al., 2014, 2015; Massuel et al., 2014, 2008; Nicolas et al., 2019) dans des contextes naturels ou fortement impactés par les prélèvements d'eau souterraine pour l'irrigation.

Les mesures *in situ* (mesure directe à l'endroit où le phénomène se produit) du niveau d'eau de surface ou souterraine donnent accès à une donnée d'une grande précision dans l'espace et dans le temps. Cependant, les réseaux d'observation sont coûteux à mettre en place et à maintenir, et les mesures ponctuelles d'un réseau clairsemé sont peu informatives à large échelle. Aussi, les mesures piézométriques (mesure du niveau d'eau souterraine au niveau des puits d'observation)

sont fréquemment interpolées pour combler les zones non échantillonnées par le réseau d'observation dans une optique de compréhension globale de l'évolution des ressources en eau. Le krigeage est l'approche géostatistique la plus utilisée pour interpoler les mesures piézométriques selon un modèle spatial entre les observations (défini par un variogramme) pour prédire les valeurs d'une variable à des positions où aucun échantillon n'est disponible. Le krigeage tente de comprendre les relations existantes entre les observations séparées par des distances différentes dans l'espace, mais dépend entre autres d'une bonne représentativité des données dans l'espace et des biais de tendance de topographie. Les mesures piézométriques peuvent être interpolées avec des méthodes plus simples sans modélisation comme l'interpolation par triangulation irrégulière (TIN) ou par pondération inverse à la distance (IDW - *Inverse Distance Weighting*).

## 1.3 Apport de la télédétection spatiale pour le suivi des ressources en eau

Le suivi des ressources en eau et la compréhension de l'impact des prélèvements anthropiques sur le cycle naturel sont essentiels pour la gestion raisonnée des ressources. Cependant elle est rendue difficile dans un contexte déstructuré où les prélèvements d'eau dans l'aquifère se font à l'échelle individuelle en dehors de toute gestion institutionnalisée et où les nombreux tanks du RHS ne sont ni recensés, ni suivis et où leur rôle est mal compris à l'échelle régionale.

Si les mesures *in situ* permettent de suivre les ressources en eau avec une grande précision sur des points discrets dans l'espace et dans le temps, elles peuvent être sujettes aux défaillances des instruments, à l'interruption du flux de données, aux erreurs de mesures et aux questions de représentativité spatiale, d'autant plus pour les réseaux d'observation peu denses. Dans ce contexte, la télédétection spatiale permet de cartographier l'ensemble de la terre fréquemment et de fournir une information globale spatialisée, et de suivre sur une longue période de temps une diversité de phénomènes environnementaux.

### 1.3.1 Satellites d'observation de la Terre

Les milliers de satellites orbitant autour de la Terre jouent un rôle essentiel sur les plans économique (télécommunications, positionnement, prévision météorologique), militaire (renseignement) et scientifique (observation astronomique, observation de la Terre, océanographie, altimétrie).

Les satellites scientifiques chargés de l'observation de la Terre peuvent être équipés de divers types de capteurs : des capteurs optiques leur permettant d'acquérir des images de la Terre dans le spectre du visible ou au delà, des radars permettant diverses applications comme l'altimétrie ou la production d'images obtenues par la rétrodiffusion de lumière polarisée, des radiomètres mesurant le rayonnement dans divers domaines spectraux (courtes longueurs d'ondes optiques, thermique et micro-ondes), ou des systèmes de télémétrie micro-ondes mesurant la distance entre les satellites d'une constellation permettant de rendre compte des variations locales du champ gravitationnel de la Terre.

La plupart des acquisitions par un satellite ou une constellation de satellites sont généralement réalisées à l'échelle globale et mises à disposition sous forme d'images géoréférencées selon un système de projection qui leur attribue des coordonnées géographiques. Ces images sont principalement caractérisées par leur résolution spatiale, définie par la distance minimale entre deux objets adjacents pouvant être distingués, mais qui désigne couramment la taille des pixels de l'image. La résolution spatiale peut varier de moins d'un mètre à plusieurs centaines de kilomètres selon les

caractéristiques des détecteurs et l’altitude du satellite. La résolution spatiale des capteurs n’est pas forcément celle à laquelle les images sont fournies : souvent, les images sont rééchantillonnées à une résolution plus élevée (par exemple, les images optiques de Pléiades à 70 m sont fournies à 50 m et les images SMOS à 40 km de résolution sont fournies sur une grille de 25 km). La figure 1.5a illustre les différences de résolution entre trois images satellitaires des missions GRACE, SMOS et MODIS.

Les acquisitions d’une même scène à différentes dates constituent une série temporelle d’images satellitaires, caractérisée par sa résolution temporelle définie par l’écart entre les différentes acquisitions. Cette résolution temporelle dépend du temps de revisite du satellite, c’est-à-dire du temps séparant deux passages consécutifs du satellite observant le même point, et du nombre de satellites dans la constellation. Les longues séries temporelles d’images satellitaires sont précieuses et informatives en hydrologie, puisque le suivi homogène d’un même phénomène sur plusieurs années permet d’en étudier les dynamiques cycliques et les évolutions à long terme. La figure 1.5b présente les séries temporelles des données de stock d’eau total, d’humidité du sol de surface et d’indice de végétation NDVI issues des missions GRACE, SMOS et MODIS sur le Telangana, mettant en évidence les dynamiques annuelles cyclique des trois variables.

### 1.3.2 Estimation de variables environnementales à partir de données satellitaires

Les divers capteurs dont peuvent être équipés les satellites leur permettent de mesurer différentes grandeurs physiques. Si ces capteurs ne permettent pas de mesure directe comme certains capteurs *in situ*, les mesures des capteurs satellitaires peuvent être utilisées pour estimer les variables d’intérêt avec des modèles de transfert radiatifs ou physiques, de méthodes de classification ou des modèles hydrologiques. Cette partie présente de manière non exhaustive les possibilités qu’offre la diversité des données satellitaires.

#### 1.3.2.1 Interprétations physiques des mesures des capteurs satellitaires

Les mesures physiques d’une grande diversité de capteurs et de méthodes d’acquisition (réflectance, température de brillance, gravimétrie, stéréoscopie) permettent d’estimer des variables hydrologiques à l’aide de modèles de transfert radiatif ou physiques (stock d’eau total avec des mesures gravimétriques, humidité du sol de surface et précipitations avec des radiomètres), et de développer des indicateurs (de végétation, de sécheresse).

La mesure de la réflectance de différentes longueurs d’ondes permet d’obtenir des images hyperspectrales de la Terre (comme celles acquises par Sentinel-2, Landsat, PlanetScope ou les capteurs MODIS). Les valeurs de réflectances permettent de calculer des indicateurs qui permettent de caractériser la présence et la densité de végétation (comme le NDVI - *Normalized Difference Vegetation Index*) ou la présence d’eau (comme le NDWI - *Normalized Difference Water Index*). De telles images peuvent aussi être obtenues par radar en mesurant le coefficient de rétrodiffusion d’une lumière polarisée envoyée à la surface de la Terre par le satellite (Sentinel-1).

La température de brillance dérivée des radiomètres est exploitée pour l’estimation des précipitations (TRMM) et de l’humidité dans les premiers centimètres de sol (SMOS, SMAP). L’humidité du sol est une composante importante du cycle hydrologique (Corradini, 2014; Seneviratne et al., 2010; Wagner et al., 2007) et est essentielle pour l’estimation des différentes composantes du cycle de l’eau et de l’énergie. L’humidité du sol joue un rôle important pour l’évaluation et la prédiction de sécheresses (Enenkel et al., 2016; Sánchez et al., 2016) et d’inondations (Koster et al., 2010),

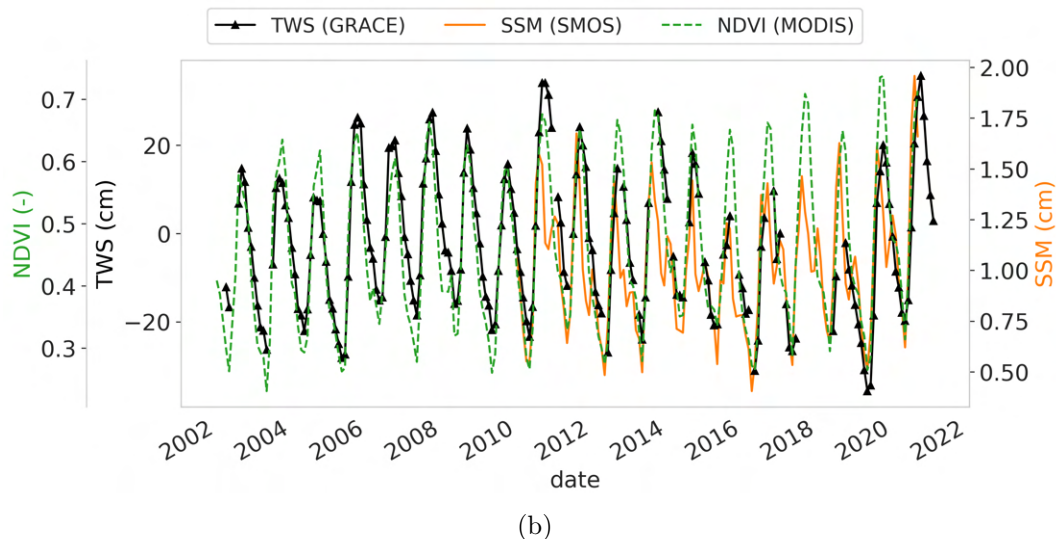
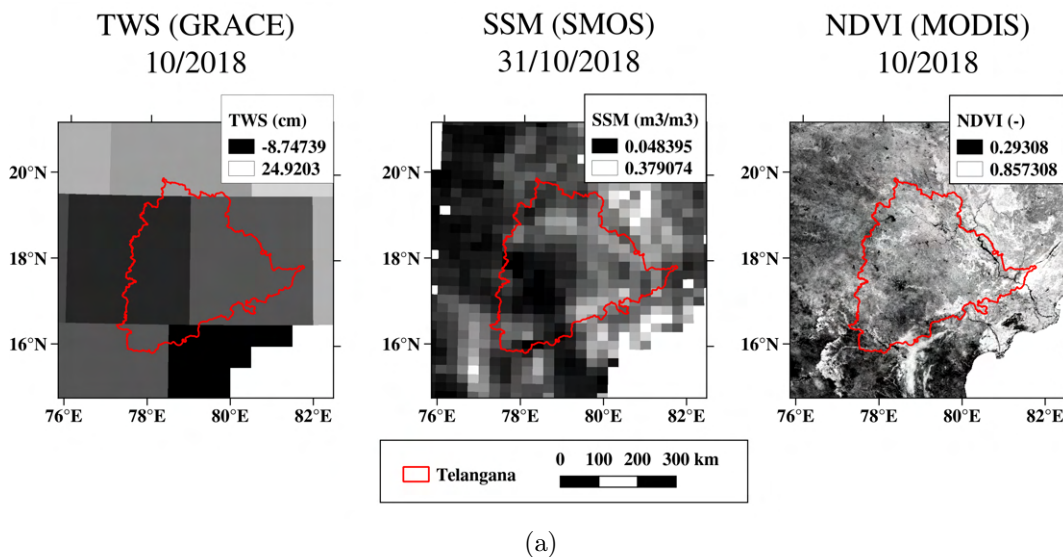


FIGURE 1.5 – Exemples de variables environnementales extraites d’images satellitaires et séries temporelles à l’échelle du Telangana. (a) Anomalie mensuelle du stock d’eau total (TWS) en octobre 2018 (GRACE) à  $3^\circ$  (environ 300 km) de résolution spatiale, humidité du sol de surface (SSM) du 31 octobre 2018 (SMOS) rééchantillonnée à 25 km de résolution, et l’indice de végétation NDVI d’octobre 2018 (MODIS) à 1 km de résolution. (b) Séries temporelles à l’échelle mensuelle de TWS, SSM et NDVI.

l’amélioration des estimations de pluie (Brocca et al., 2016; Crow et al., 2011, 2009; Pellarin et al., 2013), de ruissellement (Brocca et al., 2010) et d’évapotranspiration (Martens et al., 2017; Miralles et al., 2011). En particulier, l’humidité du sol constitue la ressource en eau utilisable par les plantes, et sa connaissance est précieuse pour la gestion de l’eau pour l’irrigation (Brocca et al., 2018; Rodríguez-Iturbe and Porporato, 2007).

Différentes techniques permettent de mesurer l’altitude des surfaces terrestres comme les altimètres radar (TOPEX/Poséidon (Fu et al., 1994), Jason-2 (Lambin et al., 2010)) ou de produire des modèles numériques de terrain (MNT) qui représentent la topographie terrestre : interférométrie radar (TanDEM-X (Krieger et al., 2013), RA-2 (Benveniste et al., 2001)), radiomètres (ASTER

(Tachikawa et al., 2011)) ou stéréoscopie (Pléiades (Gleyzes et al., 2012), ALOS (Tadono et al., 2014)). Les mesures altimétriques et MNT sont notamment utilisées pour étudier le volume des retenues d'eau (Avisse et al., 2017; Baup et al., 2014; Bonnema and Hossain, 2017; Li et al., 2019; Vanthof and Kelly, 2019).

L'étude des forces gravitationnelles s'exerçant sur les satellites peut aussi être interprétée comme des anomalies de champ de pesanteur. C'est le cas des satellites des constellations GRACE et GRACE Follow-On dont les variations de trajectoire sont interprétées comme des fluctuations de champ gravitationnel dues aux mouvements d'eau dans les enveloppes fluides de la Terre. Ces données sont particulièrement intéressantes en hydrologie car elles évaluent le stock d'eau total intégré verticalement (TWS - *Total Water Storage*), et renferment donc une information sur l'évolution du stock d'eau souterraine. Les données GRACE sont donc largement utilisées en hydrologie pour étudier l'évolution à long terme du stock d'eau souterraine en soustrayant à l'anomalie de TWS les stocks de surface et d'humidité du sol, à des échelles régionales ou globale (Breña-Naranjo et al., 2014; Cao and Roy, 2020; Frappart et al., 2019; Frappart and Ramillien, 2018; Papa et al., 2015; Rodell et al., 2018; Rzepecka and Birylo, 2020; Tiwari et al., 2009; Zhang et al., 2020).

### 1.3.2.2 Classification à partir d'images satellitaires

La grande diversité des variables fournies par la télédétection peut être utilisée à des fins de cartographie d'occupation du sol réalisées par classification automatique. Ces techniques sont notamment utilisées pour la détection des étendues d'eau en surface sur des images optiques ou radar (Bie et al., 2020; Bioresita et al., 2018; Caballero et al., 2019; Cordeiro et al., 2021; Evans et al., 2010; Hardy et al., 2019; Peña-Luque et al., 2021; Pekel et al., 2016; Tulbure et al., 2016; Vanthof and Kelly, 2019) ou la cartographie de l'occupation du sol plus ou moins détaillée : couvert végétal ou urbain, densité de couvert, culture irriguée ou non, type de culture, etc. (Ofori-Ampofo et al., 2021; Pageot et al., 2020; Pelletier et al., 2017; Wang et al., 2020). Deux grands types d'apprentissages permettent de réaliser ce genre de classification : l'apprentissage supervisé et non supervisé.

**Apprentissage supervisé.** L'apprentissage supervisé établit des règles de décision pour prédire une classe en fonction de variables (réflectances dans le cas d'images optiques) à partir d'individus (pixels ou groupe de pixels) étiquetés par l'utilisateur, et applique ces règles pour prédire la classe des individus non étiquetés. Parmi les algorithmes de classification supervisés se situent les arbres de décision, l'algorithme *Random Forest* (Breiman, 2001), le SVM (séparateur à vaste marge), ou les réseaux de neurones. En particulier, l'algorithme *Random Forest* est très utilisé pour les classifications d'occupation du sol ou pour la détection de l'eau sur des images optiques ou radar (Bangira et al., 2019; Feng et al., 2015; Ko et al., 2015; Pageot et al., 2020; Pelletier et al., 2017; Tulbure et al., 2016) du fait de sa rapidité, sa simplicité et sa résistance au bruit dans les étiquettes des échantillons (Pelletier et al., 2017). Cette méthode d'ensemble combine les prédictions de plusieurs arbres de décision indépendants apprenant chacun sur un sous-ensemble d'attributs, lui procurant sa robustesse.

**Apprentissage non supervisé.** L'apprentissage non-supervisé discrimine automatiquement les individus en plusieurs classes selon leurs caractéristiques sans nécessiter d'individus étiquetés. Ce principe est appliqué notamment par les algorithmes K-means ou EM (*Expectation-Maximisation*).

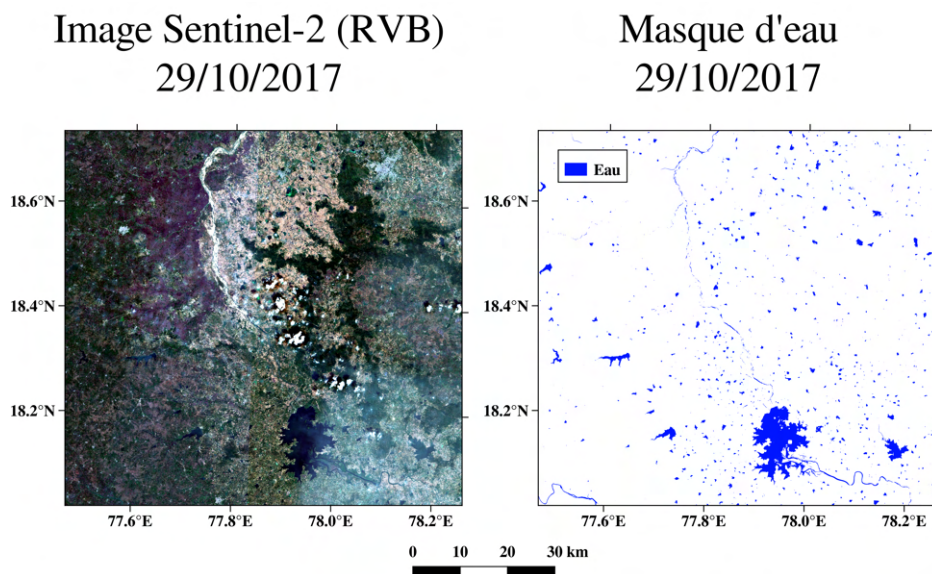


FIGURE 1.6 – Détection d’eau sur une image Sentinel-2 dans l’Ouest du Telangana avec un algorithme de classification supervisé *Random Forest* (la production des cartes d’eau est détaillée en section 2.3.4.3). Le masque d’eau (image de droite) est produit à partir des réflectances de l’image Sentinel-2 (image de gauche).

### 1.3.2.3 Modèles hydrologiques

Certaines variables et flux hydrologiques ne sont pas directement accessibles ou facilement dérivables à partir de mesures satellitaires. Ces variables peuvent être estimées à partir de modèles hydrologiques, comme les modèles de surface terrestre (LSM - *Land Surface Model*) décrivant les cycles d’eau et d’énergie à la surface de la Terre par des équations physiques, ou des modèles semi-empiriques plus conceptuels (Liu et al., 2018; Poulouvassilis et al., 2001; Ringerud et al., 2015).

Les LSM peuvent modéliser les processus à la surface de la Terre de manière plus ou moins complexe (Overgaard et al., 2006). Les modèles simples “monocouches” comme celui de Penman-Monteith (Monteith, 1965) considèrent la surface terrestre comme homogène et ne différencient pas la transpiration des plantes et l’évaporation du sol, mais peuvent décrire l’évapotranspiration (somme de l’évaporation et de la transpiration) avec une précision satisfaisante (Overgaard et al., 2006). Les modèles à deux couches ou plus permettent de modéliser les transferts d’énergie entre le sol et la canopée, décrite par un nombre de strates plus ou moins grand.

Les modèles font appel à un certain nombre de paramètres fixes déterminés par un processus de calibration minimisant l’écart entre les prédictions du modèle et des observations (mesures *in situ* ou données satellitaires par exemple). Les modèles sont généralement forcés par des observations climatiques (comme les précipitations, la température de l’air ou la pression atmosphérique) à partir desquelles le modèle génère des prédictions.

Les biais de forçage et de paramétrisation peuvent s’accumuler, rendant les prédictions des modèles imprécises. Des techniques d’assimilation visent à fusionner les prédictions du modèle à des observations de manière dynamique et continue, soit par ajustement des sorties, des paramètres ou des données de forçage. Deux grands types d’algorithmes sont utilisés pour l’assimilation de données dans les modèles : les filtres (*filters*) et les “lisseurs” *smoothers*, les plus utilisés étant ceux de Kalman (*ensemble Kalman filter* et *ensemble Kalman smoother*). Les filtres intègrent

des observations dès qu’elles sont disponibles et modifient seulement les états les plus récents du modèle. Ils sont par conséquent plus adaptés à la prédiction, étant donné qu’ils ne nécessitent que l’état initial du modèle. Les *smoothers* utilisent une série d’observations pour modifier les états du modèle passés et futurs sur une fenêtre de temps définie, et sont plus adaptés pour les réanalyses, mais sont plus coûteux en temps de calcul (Tian et al., 2017; Zaitchik et al., 2008).

La grande disponibilité des données satellitaires, combinées aux données *in situ*, permettent de calibrer des modèles et de faire des prédictions de manière distribuée sur tout le globe, améliorées par l’assimilation des données satellitaires. Certains modèles estiment un grand nombre de variables comme GLDAS (*Global Land Data Assimilation System*) (Rodell et al., 2004) qui génère des champs d’évapotranspiration, de ruissellement, de quantité de neige, d’humidité du sol dans les différentes couches du sol ou de stock d’eau total (TWS). D’autres modèles sont dédiés à l’estimation d’un petit nombre de variables, comme GLEAM qui génère uniquement des champs d’évapotranspiration et d’humidité du sol à partir du modèle d’évaporation simplifié de Priestley et Taylor (Martens et al., 2017; Miralles et al., 2011).

Malgré l’assimilation de données, la précision des estimations des modèles est limitée par certains facteurs comme leurs hypothèses simplificatrices des processus réels, leur paramétrisation, l’incertitude des données de forçage météorologique, ou l’absence de prise en compte dans le modèle des activités anthropiques telles que l’irrigation (Long et al., 2013). Ce dernier facteur est parfois considéré, comme dans le modèle WGHW (*WaterGap Hydrology model*, <http://www.watergap.de/>, dernier accès le 27 juin 2022) qui modélise les stocks et flux d’eau douce à l’échelle globale en incluant les prélèvements à destination de la consommation humaine, afin d’étudier l’impact de ces prélèvements sur les ressources.

## 1.4 Position du problème

### 1.4.1 Incertitudes de la connaissance des stocks d’eau à toutes les échelles

Dans l’État du Telangana, les systèmes de suivi de ressources en eau ne permettent pas de rendre compte de manière exhaustive de l’ensemble des ressources souterraines et de surface et de ses usages pour l’irrigation à fine ou moyenne échelle, ou à l’échelle de l’État à une résolution temporelle suffisante. Des estimations annuelles de fluctuations d’eau souterraine ou d’eau de surface utilisable sont fournies par des organismes gouvernementaux comme le CWC, mais sans grande flexibilité sur leur résolution spatiale et temporelle. En particulier, des incertitudes demeurent notamment sur la connaissance des stocks et des flux d’eau, notamment la connaissance du stock d’eau souterraine, du volume du RHS et de la quantité d’eau utilisée pour l’irrigation. Or, des systèmes de suivi régulier des ressources en eau à plusieurs échelles bénéficieraient à la compréhension fine des usages de l’eau, nécessaire pour la gestion des ressources à l’échelle locale ou régionale.

Le Telangana possède un réseau important de plus de 3700 piézomètres répartis sur 114 800 km<sup>2</sup>. Cependant, en plus d’être incertaines, les mesures piézométriques mensuelle ou saisonnière et ponctuelle peuvent aussi être impactée par des effets à court ou long terme comme le pompage d’eau à proximité au moment de la mesure, qui ne peut être ni détecté ou rectifié avec de telles fréquences d’acquisition. La connaissance du RHS est également limitée en raison de l’absence d’inventaire ou de système de suivi des tanks, du fait de leur grand nombre et de leur taille. Enfin, le volume d’eau utilisé pour l’irrigation est difficilement quantifiable, du fait de la variabilité des pratiques agricoles, de l’exploitation individuelle de l’aquifère par les agriculteurs, de l’absence de suivi du RHS et du coût des enquêtes terrain.

## 1.4.2 Capacités de la télédétection

La télédétection peut observer une grande variété de variables environnementales de manière spatialisée à différentes résolutions spatiales et temporelles. La synergie et de ces différentes données satellitaires représente un grand potentiel pour apporter une meilleure compréhension des processus hydrologiques, ainsi qu'un suivi quantitatif des ressources souterraines, du volume du RHS et de l'irrigation.

La mise en cohérence des séries temporelles de données satellitaires de pluie, d'humidité du sol, gravimétriques, et des classifications d'occupation du sol réalisées à partir d'imagerie optique ou radar, et ce à différentes échelles, permet une compréhension des phénomènes hydrologiques naturels et de leur interaction avec les activités anthropiques (Panda and Wahr, 2016; Vissa et al., 2019). En particulier, les données d'humidité du sol et de gravimétrie sont des données uniques en hydrologie pour le suivi des ressources en eau dans les zones fortement irriguées. Les estimations d'humidité du sol en surface telles celles des missions SMOS et SMAP sont précieuses pour la gestion de l'eau pour l'irrigation, puisqu'elles constituent les ressources en eau utilisables pour les plantes. Les événements d'irrigation peuvent notamment être détectés par la comparaison de ces données avec les précipitations (Brocca et al., 2018). Les données GRACE, quant à elles, fournissent des anomalies de lame d'eau intégrée sur toute la colonne verticale, qui incluent donc l'eau stockée par la canopée, la neige, l'eau de surface, l'humidité du sol, et enfin l'eau souterraine. Ces données possèdent donc un grand potentiel pour l'estimation de l'eau souterraine, ce qui les rendent uniques par rapport à la plupart des autres données satellitaires qui n'estiment que des variables de surface.

Cependant la faible résolution des données GRACE (300-400 km) ne permet pas l'obtention de cartes à une échelle pertinente pour l'usage des sols. Cette dernière décennie, de nombreuses études ont cherché à obtenir des cartes mensuelles d'eau souterraine à une résolution supérieure par désagrégation des données GRACE, par leur assimilation dans des modèles hydrologiques (Giroto et al., 2016; Houborg et al., 2012; Nie et al., 2019; Schumacher et al., 2018; Tian et al., 2017; Zaitchik et al., 2008) ou avec des approches empiriques basées sur le caractère supposément multi-échelle de la relation entre les données GRACE et des variables auxiliaires (Ali et al., 2021; Jyolsna et al., 2021; Karunakalage et al., 2021; Sahour et al., 2020; Seyoum and Milewski, 2017; Vishwakarma et al., 2021; Zhang et al., 2021a,b).

Le suivi quantitatif du volume du RHS peut également être réalisé par imagerie satellitaire. Les méthodes actuelles d'estimation de volume par les données satellitaires reposent sur l'association de mesures altimétriques ou de modèles numériques de terrain (MNT), avec des cartes d'eau produites à partir d'images optiques ou radar (Avisse et al., 2017; Baup et al., 2014; Bhagwat et al., 2019; Bonnema and Hossain, 2017; Crétaux et al., 2011; Duan and Bastiaanssen, 2013; Frappart et al., 2006). Ces estimations se font généralement sur de grands lacs isolés, et sont difficilement applicables sur les tanks du RHS du fait de la faible résolution des MNT globaux (généralement 30 m) et l'absence de contrôle sur la date d'acquisition qui ne garantit pas l'accès à leur bathymétrie. Ces études portent plus rarement sur un ensemble de petits réservoirs (Arendt et al., 2021; Vanthof and Kelly, 2019) et nécessitent des MNT de plus haute résolution. Avec l'arrivée de nouvelles missions pour la cartographie de la topographie de la Terre à très haute résolution comme CO3D (Lebègue et al., 2020), l'application des méthodes d'estimation de volume avec des cartes d'eau produites à partir d'images satellites à fort taux de revisite présente un grand potentiel pour le suivi volumétrique du RHS en quasi-temps réel et le développement de systèmes d'alerte de sécheresse, autrement inenvisageable pour ces tanks sans système de suivi *in situ*.



### 1.4.3 Objectifs et organisation du manuscrit

L'objectif général de la thèse est de cartographier les ressources en eau et de comprendre leur dynamique dans l'État du Telangana, dans l'optique d'une gestion optimisée des ressources en eau dans les régions semi-arides fortement irriguées. Ceci est réalisé à l'aide d'une synergie entre plusieurs données de télédétection multi-capteurs et multi-échelles permettant une approche spatialisée et un suivi temporel régulier. L'utilisation de données satellitaires globales vise à la généralisation et la reproductibilité sur d'autres zones du monde des approches développées. Ces travaux se concentrent en particulier sur le suivi du volume d'eau souterraine dans les aquifères, principal réservoir exploité pour l'irrigation, et du système de récupération des eaux de pluies, réservoir promu pour sa facilitation de la recharge mais peu étudié et dont le fonctionnement est méconnu. Les données *in situ* disponibles sur la zone d'étude, du fait de leur finesse et du caractère direct de leurs mesures, sont utilisées pour la validation des approches développées dans ces travaux. Ces données comportent les piézomètres du *Groundwater Department* du Telangana, la cartographie de six tanks du RHS acquises lors d'une campagne terrain de l'IFCGR (Perrin et al., 2012), et une campagne saisonnière d'échantillonnage d'occupation du sol sur 4 ans réalisée dans le cadre d'une collaboration entre le CESBIO et l'IFCGR.

Pour répondre à ces objectifs généraux, les données satellitaires sont mises à contribution du suivi des ressources en eau et de leur dynamique à fine résolution dans tous les compartiments : l'eau de surface (eau stockées dans les retenues d'eau), l'eau du sol (l'humidité du sol et l'eau utilisée pour l'irrigation) et l'eau souterraine. Ce manuscrit est organisé en six chapitres et comporte trois axes principaux de recherche selon les objectifs définis :

- Ce chapitre 1 contextualise les travaux de la thèse en exposant les problématiques liées aux ressources en eau dans le Telangana pour montrer comment les données de télédétection peuvent contribuer au suivi de ces ressources,
- Le chapitre 2 présente les données *in situ*, satellitaires et les modèles utilisés dans le cadre de ces travaux et leurs caractéristiques conditionnant leur utilisation,
- Le chapitre 3 constitue le premier axe de recherche. Il consiste en l'inter-comparaison et la mise en cohérence des dynamiques temporelles et spatiales des données *in situ*, satellitaires, et de modèles disponibles sur le Telangana pour l'amélioration de la compréhension des phénomènes hydrologiques, de leur évolution et de leur interaction avec les activités humaines. Une attention particulière est apportée au potentiel des données SMOS et GRACE pour la levée de certains verrous de connaissances hydrologiques sur l'irrigation et l'eau souterraine,
- Le chapitre 4 constitue le deuxième axe de recherche, et porte sur le suivi de l'eau souterraine à fine échelle par la désagrégation des données GRACE. Les problématiques liées aux techniques de désagrégation sont exposées dans ce chapitre, ainsi que les résultats de plusieurs modèles de désagrégation appliqués sur le Telangana. Un travail sur les méthodes de validation des produits GRACE désagrégés est également présenté sous la forme d'un article soumis à *Hydrology and Earth System Sciences* (HESS), qui synthétise les méthodes de validation des produits GRACE désagrégés utilisées dans la littérature, et propose un nouveau schéma de validation comblant les lacunes existantes dans le domaine,
- Le chapitre 5 constitue le troisième et dernier axe de recherche, et porte sur le potentiel des MNT très haute résolution pour le suivi du volume du RHS et les perspectives applicatives pour la compréhension de l'impact du RHS sur l'irrigation et la recharge des aquifères. Une

quantification de l'apport des MNT très haute résolution par rapport aux MNT globaux pour le suivi volumétrique du RHS est présentée dans un article soumis à *Remote Sensing* (MDPI),

- Le chapitre 6 présente les conclusions des trois axes de recherche et propose des perspectives pour l'amélioration des méthodes de suivis des ressources en eau par télédétection.



# Chapitre 2

## Données

### Sommaire

---

<b>2.1</b>	<b>Introduction</b>	<b>22</b>
<b>2.2</b>	<b>Données <i>in situ</i></b>	<b>23</b>
2.2.1	Mesures piézométriques	23
2.2.2	Suivi des petits réservoirs	23
2.2.3	Occupation du sol	25
<b>2.3</b>	<b>Données satellitaires</b>	<b>27</b>
2.3.1	Variation du stock d'eau total (GRACE)	27
2.3.2	Humidité du sol	33
2.3.3	Indice de végétation et évapotranspiration MODIS	34
2.3.4	Eau de surface	34
2.3.5	Précipitations TRMM	36
<b>2.4</b>	<b>Modèles</b>	<b>37</b>
2.4.1	GLDAS	37
2.4.2	GLEAM	38
2.4.3	Évapotranspiration potentielle	38
<b>2.5</b>	<b>Ce qu'il faut retenir</b>	<b>39</b>

---

## 2.1 Introduction

La gestion de l'eau douce est essentielle dans les régions où elle est fortement sollicitée pour les activités humaines. C'est en particulier le cas dans l'État du Telangana dans le Sud de l'Inde, où l'irrigation intense et les usages domestiques d'une population dense exercent une grande pression sur les ressources en eau, principalement souterraines mais aussi celles stockées en surface, qui sont plus ou moins reconstituées chaque année avec les pluies de mousson. La gestion parcimonieuse des ressources en eau est d'autant plus essentielle dans le Telangana du fait de la faible capacité de stockage de ses aquifères granitiques, où, en l'absence de ressources anciennes comme dans les aquifères profonds du nord de l'Inde, les pluies de mousson constituent la seule source d'eau pour l'année.

Le suivi des ressources en eau et la compréhension de l'impact anthropique sur les dynamiques hydrologiques naturelles nécessitent l'accès à un catalogue exhaustif de variables hydro-agronomiques : les principaux flux hydriques entrants et sortants (précipitations, évapotranspiration, ruissellement), l'état des stocks d'eau (eau de surface, humidité du sol, aquifère), l'usage des sols (végétation naturelle, terres cultivées, espace urbanisé), l'état de la végétation naturelle et des cultures et de leurs différents stades.

Certaines de ces variables peuvent être directement quantifiées ou estimées avec des mesures *in situ* (niveau de l'aquifère, niveau des réservoirs d'eau de surface, précipitations) ou obtenues avec des études locales (occupation du sol, quantité d'eau utilisée pour l'irrigation). Ces moyens permettent l'acquisition d'une donnée d'une grande précision, mais peu informative à large échelle du fait de la nature ponctuelle de la donnée et du coût élevé de la mise en place de réseaux d'observation denses.

La télédétection s'affranchit de ce problème en estimant ces variables hydrologiques à plus large échelle, généralement globales, spatialisées et à des résolutions spatiales et temporelles variées. Les mesures physiques d'une grande diversité de capteurs et de méthodes d'acquisition (réflectance, température de brillance, gravimétrie, stéréoscopie) permettent d'estimer des variables hydrologiques à l'aide de modèles de transfert radiatif ou physiques (stock d'eau total avec des mesures gravimétriques, humidité du sol de surface et précipitations avec des mesures micro-ondes), et de développer des indicateurs (de végétation, de sécheresse) ou des méthodes d'estimation de certaines variables (cartographie et quantification de l'eau de surface, cartographie de l'occupation du sol).

Certaines variables ne sont pas directement accessibles ou facilement dérivables à partir de mesures satellitaires comme l'évapotranspiration, le ruissellement ou l'humidité dans les couches profondes du sol. Aussi, l'utilisation de modèles hydrologiques décrivant les cycles de l'eau et d'énergie à la surface de la Terre permet d'estimer ces variables à partir de données de forçage météorologique et/ou de l'assimilation de données satellitaires.

Les travaux réalisés ici s'intéressent à la cartographie des ressources en eau sur l'État du Telangana à partir des données satellitaires, et en particulier à l'impact de l'irrigation sur ces ressources. Une multitude de variables issues de données satellitaires et de modèles sont utilisées pour représenter les processus environnementaux. Les données *in situ* disponibles sur la zone d'étude servent à la validation des approches utilisées. Ce chapitre décrit l'ensemble des données utilisées dans ces travaux en terme d'hétérogénéité de capteurs, de résolutions temporelles et spatiales, d'incertitudes et de prétraitements effectués pour extraire les variables d'intérêt, caractéristiques déterminantes dans le développement des approches décrites dans les chapitres suivants. Une première partie décrit les données *in situ* disponibles sur l'État du Telangana pour le suivi de l'eau et de ses usages. La deuxième partie présente les données satellitaires utilisées et les variables qui en sont dérivées, présentées dans l'ordre des couches les plus profondes aux plus superficielles. La dernière partie

présente les modèles utilisés pour l'estimation de certaines variables difficilement accessibles par satellite mais nécessaires à la représentation exhaustive des processus hydrologiques, en particulier pour la fermeture du bilan hydrologique. La table 2.1 résume les spécificités des données *in situ*, satellitaires et sorties de modèle présentées dans ce chapitre.

## 2.2 Données *in situ*

### 2.2.1 Mesures piézométriques

Le département de l'eau souterraine du Telangana (*Groundwater Department*, <https://gwd.telangana.gov.in/>, dernier accès le 23 juin 2022) possède un réseau de piézomètres d'observation sur le Telangana où la piézométrie est relevée ponctuellement une fois par mois de manière automatique ou manuelle selon les stations. Les données sont fournies par l'*India Water Resources Information System* (WRIS, <https://indiawris.gov.in/wris>, dernier accès le 19 janvier 2022). Un filtre est utilisée pour supprimer les valeurs entre 0 et 1 m pendant la saison des pluies, car jugées inexactes. Les séries temporelles peuvent être interrompues (absence de relevé, niveau d'eau en dessous du capteur). Pour ces travaux, seules les stations avec une série temporelle d'au minimum sept points sur la période entre mai 2007 et février 2020 sont conservées, pour un total de 1006 puits répartis sur le Telangana (figure 1.2). Des cartes mensuelles de niveau d'eau à 5 km de résolution sont produites par interpolation des données ponctuelles par pondération inverse à la distance des puits (IDW - *Inverse Distance Weighting*), méthode choisie pour la souplesse de sa mise en œuvre pour la production d'un grand nombre de cartes à plusieurs dates différentes.

### 2.2.2 Suivi des petits réservoirs

Les petits réservoirs (aussi appelés "tanks") du système de récupération des eaux de pluie (RHS - *Rainwater Harvesting System*) sont peu étudiés du fait de leur petite taille, tant au niveau de leur recensement que de leur suivi. Dans la région, quelques tanks ont fait l'objet d'un suivi individuel de leur volume dans le temps, notamment pour comprendre plus précisément leur fonctionnement (infiltration, évaporation) et leur impact sur la recharge (Boisson et al., 2014, 2015; Massuel et al., 2014, 2008; Nicolas et al., 2019; Perrin et al., 2012; Wintgens et al., 2016). Ces études ont majoritairement été réalisées par le CEFIRES (Centre Franco-Indien de Recherche sur les Eaux Souterraines), une collaboration entre le BRGM (service géologique national français) et le NGRI (*National Geophysical Research Institute*, Inde).

En particulier, Les données de Perrin et al. (2012) acquises en 2010 seront utilisées pour ces travaux. Les volumes et surfaces de cinq tanks sont mesurés par deux méthodes différentes dans la région de Gajwell, à environ 50 km au nord d'Hyderabad. La première méthode consiste à cartographier les surfaces en eau à intervalles de temps réguliers par suivi GPS pendant la saison sèche (tank 6 sur la figure 2.1a selon la numérotation de Perrin et al. (2012)). Avec la deuxième méthode, le fond des tanks vides est cartographié avec un GPS différentiel (DGPS). Ceci concerne quatre tanks (4, 5, 15 et 27 sur la figure 2.1a) pour 1146 points de mesure. Pour les deux méthodes, les altitudes des points de mesures sont interpolées avec la méthode TIN (interpolation par triangulation irrégulière) pour produire des modèles numériques de terrain (MNT), puis les surfaces et volumes de chaque tank sont calculés pour chaque valeur d'altitude avec un incrément de 10 cm (figure 2.1b).

TABLE 2.1 – Récapitulatif des données utilisées. GWL : niveau d'eau souterraine (*Ground Water Level*), TWSA : anomalie de stock d'eau total (*Total Water Storage Anomaly*), SSM : humidité de surface du sol (*Surface Soil Moisture*), RZSM : humidité du sol en zone racinaire (*RootZone Soil Moisture*), MNT : Modèle Numérique de Terrain, ET : ÉvapoTranspiration, PET : évapotranspiration potentielle (*Potential EvapoTranspiration*). Projections : UTM zone 44N (EPSG :32644), WGS84 (EPSG :4326), EASE2 grid (EPSG :6933), UTM zone 43N (EPSG :32643)

Grandeur	Unité	Source	Type	Capteur	Résolution temporelle	Résolution spatiale native	Résolution de distribution	Projection (EPSG)
GWL	m	WRIS	<i>in situ</i>	piézomètre	1 mois	ponctuel	ponctuel	UTM zone 44N
Altitude	m	-	<i>in situ</i>	GPS différentiel	-	ponctuel	ponctuel	UTM zone 44N
Occupation du sol	-	Sentinel-2 L2A	<i>in situ</i> + classification	optique	1 saison	10 m	10 m	UTM zone 44N
TWSA	cm	GRACE JPL RL06M	satellitaire	gravimétrie	1 mois	≈ 300 km	0.5°	WGS84
SSM	m <sup>3</sup> /m <sup>3</sup>	SMOS L3SM RE07	satellitaire	radiomètre	3 jours	40 km	25 km	EASE2 grid
RZSM	m <sup>3</sup> /m <sup>3</sup>	SMOS	satellitaire	radiomètre + modèle	1 jour	40 km	25 km	EASE2 grid
SSM	m <sup>3</sup> /m <sup>3</sup>	SMAP SPL3SMP.006	satellitaire	radiomètre	2-3 jours	40 km	36 km	EASE2 grid
RZSM	m <sup>3</sup> /m <sup>3</sup>	SMAP SPL4SMGP.004	satellitaire + modèle	radiomètre	3 heures	-	9 km	EASE2 grid
SSM	m <sup>3</sup> /m <sup>3</sup>	ESA CCI v2	satellitaire	capteurs passifs	1 jour	40 km	25 km	WGS84
MNT	m	Pléiades	satellitaire	stéréoscopie optique	ponctuel	2 m	2 m	UTM zone 44N et 43N
NDVI	-	MODIS	satellitaire	optique	1 mois	250 m	1 km	ellipsoïdale
ET	kg/m <sup>2</sup>	MODIS MOD13A3v006	modèle	-	8 jours	-	500 m	ellipsoïdale
Surfaces en eau	-	LandSat	classification	optique	1 mois	30 m	30 m	WGS84
Surfaces en eau	-	Sentinel-2	classification	optique	5 jours	10 m	10 m	UTM zone 44N et 43N
Précipitations	mm/heure	TRMM 3B43 v7	satellitaire	-	1 jour	multiples	0.25°	WGS84
RZSM, SSM, ruissellement	kg/m <sup>2</sup>	GLDAS-NOAH v2.1	modèle	-	3 heures	-	0.25°	WGS84
RZSM, SSM	m <sup>3</sup> /m <sup>3</sup>	GLEAM v3.5b	modèle	-	1 jour	-	0.25°	WGS84
PET	mm	RM-OBS/PU v2	modèle	-	3 heures	-	0.25°	WGS84

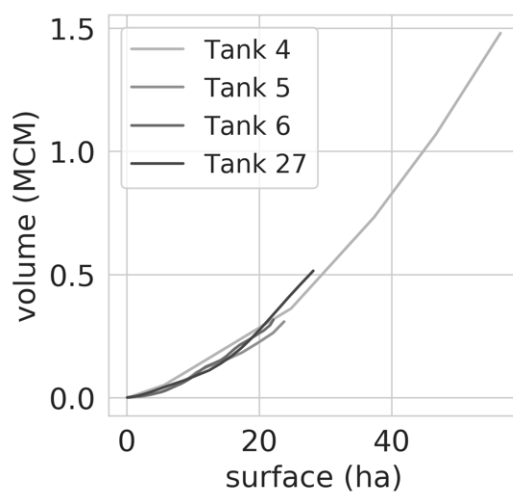
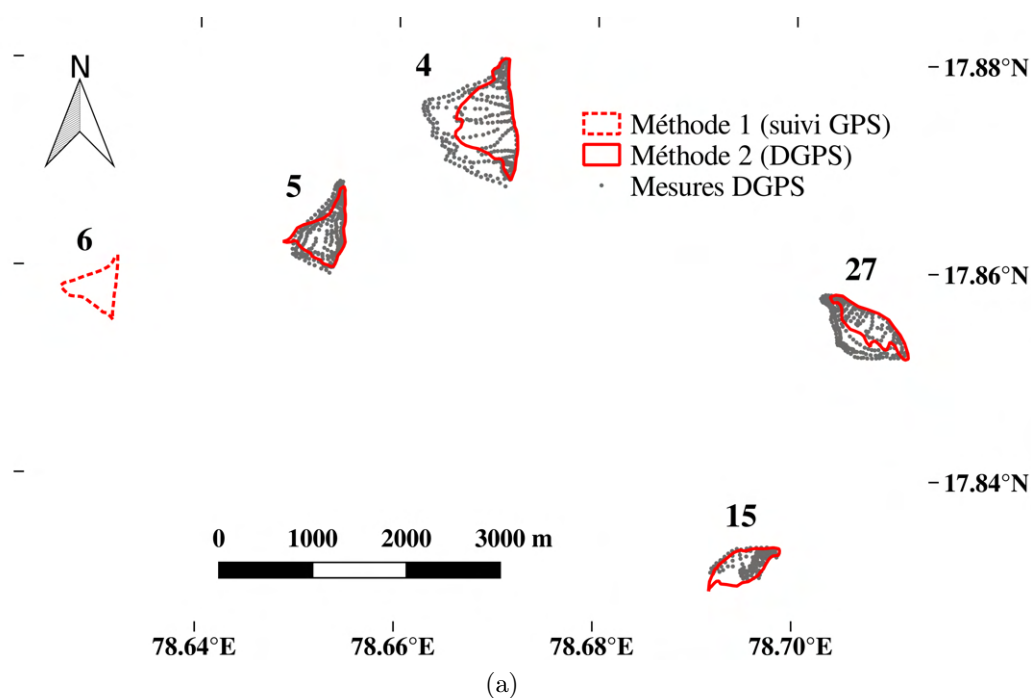


FIGURE 2.1 – (a) Localisation des tanks cartographiés, numérotés comme dans Perrin et al. (2012). Le tank 6, délimité par la ligne rouge pointillée, est cartographié par suivi GPS des surfaces à intervalles de temps réguliers (méthode 1). Les bathymétries des tanks 4, 5, 15 et 27 (ligne pleine) sont cartographiées par DGPS (méthode 2) pour un total de 1146 mesures (points gris) sur les quatre tanks. (b) Volumes (en millions de mètres cubes - MCM) et surfaces (en hectares - ha) des tanks calculés avec un incrément de hauteur de 10 cm.

### 2.2.3 Occupation du sol

Des cartes d'occupation du sol pour les saisons culturales de Kharif et Rabi de 2016 à 2019 sont produites à partir de données terrain sur huit tuiles (tuilage des images Sentinel-2) de  $100 \times 100 \text{ km}^2$  (tuiles indiquées en figure 2.2). Les données terrain sont acquises sur plusieurs jours consécutifs durant chaque saison par identification de parcelles sur un itinéraire couvrant le Telangana, dans



le cadre d'une collaboration entre le CESBIO et l'IFCGR. La figure 2.2 montre l'emplacement des 1669 parcelles labellisées pour la saison humide 2019. Les parcelles ont été classées selon neuf classes : riz, légumes divers, maïs, vergers, forêt, sol nu, urbain, eau et coton.

Les cartes sont produites avec l'algorithme IOTA2 (Infrastructure pour l'Occupation des sols par Traitement Automatique Incorporant les Orfeo Toolbox Applications, <https://framagit.org/iota2-project/iota2>, dernier accès le 11 mai 2022), développé par le CESBIO (Centre d'Étude Spatiale de la BIOSphère), à partir d'une série temporelle tous les cinq jours d'images optiques Sentinel-2 L2A. Les données Sentinel-2 L2A sont corrigées par l'algorithme MAJA (Hagolle et al., 2017) et téléchargées sur la plateforme <https://www.theia-land.fr/> (dernier accès le 16 mai 2022). IOTA2 réalise une classification multitemporelle à l'échelle du pixel avec un algorithme de classification *random forest* (Breiman, 2001) en interpolant les réflectances des images Sentinel-2 à un intervalle de temps régulier (ici 10 jours), remplissant ainsi les dates laissées vides par les images ennuagées et harmonisant les dates d'acquisitions qui varient avec les orbites des satellites. Les classifications en Kharif (respectivement Rabi) pour chaque année sont réalisées avec les images acquises entre juillet et novembre (respectivement entre janvier et avril) pour correspondre au calendrier cultural.

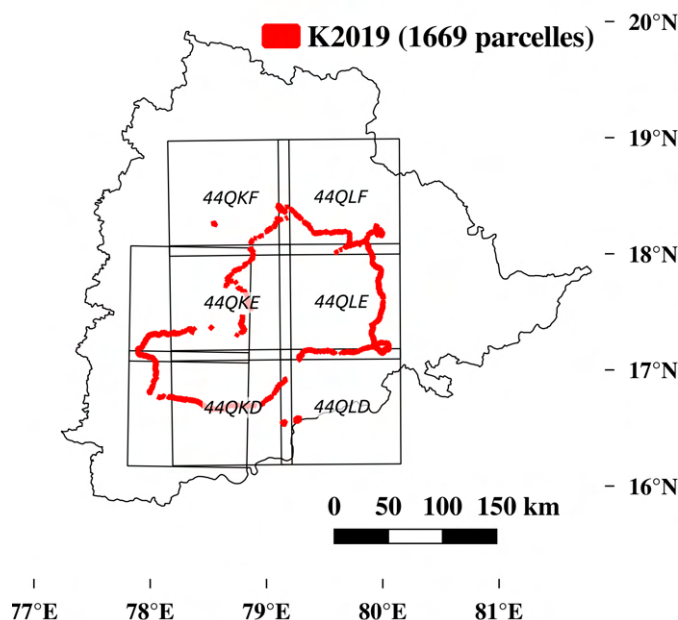


FIGURE 2.2 – Emplacement des parcelles labellisées pour la saison humide (Kharif) 2019 et des tuiles Sentinel-2.

La figure 2.3 montre les distributions des précisions et rappels des classifications pour les huit périodes. Le rappel d'une classe  $C$  (équation 2.1) mesure la capacité de l'algorithme à reconnaître un pixel de la classe (proportion de pixels  $C$  qui sont classés  $C$ ), et la précision (équation 2.2) mesure la pureté de la classe (proportion de pixels classés  $C$  qui sont effectivement  $C$ ). Les classes les mieux identifiées sont le riz, le couvert forestier, l'urbain et l'eau. Les rappels sont moins bons que les précisions, indiquant que ces classes ne sont pas toujours bien reconnues. Les précisions et rappels des autres classes sont plus faibles (rarement au-delà de 80 %), aussi ces classes doivent être interprétées avec prudence. En particulier, les confusions sur le riz sont plus fréquentes en Kharif en raison du faible nombre d'images non ennuagées. L'utilisation d'images radar Sentinel-1 permet de s'affranchir du problème des nuages, mais induit d'autres confusions (Ferrant et al.,

2019, 2017).

$$rappel = \frac{VP}{VP + FN} \quad (2.1)$$

$$precision = \frac{VP}{VP + FP} \quad (2.2)$$

VP : vrai positif, FP : faux positif, FN : faux négatif.

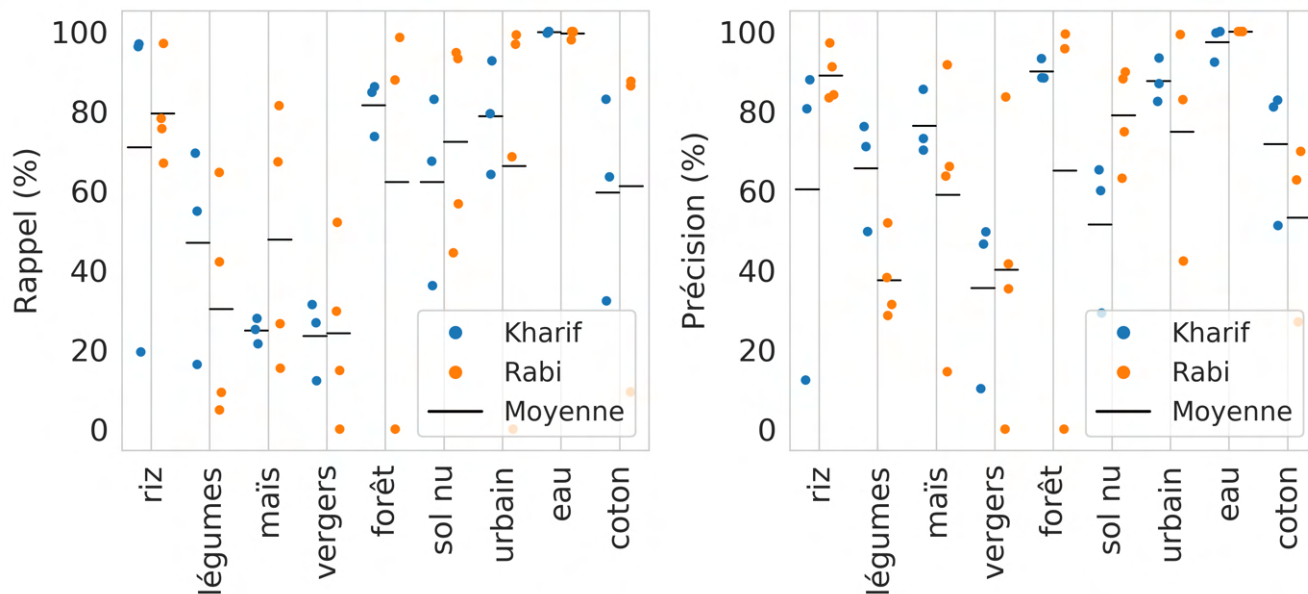


FIGURE 2.3 – Rappel et précision des cartes d’occupation du sol pour chaque classe et chaque saison. La moyenne des métriques est indiquée par une barre noire horizontale.

## 2.3 Données satellitaires

### 2.3.1 Variation du stock d’eau total (GRACE)

#### 2.3.1.1 La mission GRACE

La mission GRACE (*Gravity Recovery and Climate Experiment*), lancée le 17 mars 2002, est une constellation de deux satellites co-orbitants sur une orbite quasi polaire ( $89.5^\circ$ ), à une altitude relativement basse d’environ 500 km et séparés d’une distance d’environ 220 km, et réalisant 16 révolutions quotidiennes.

Chaque satellite dispose de cinq instruments (<https://podaac.jpl.nasa.gov/grace>, dernier accès le 11 mai 2022) :

- KBR (K-Band Ranging System) mesure les changements de distance entre les deux satellites avec une précision micrométrique, (Frappart and Ramillien, 2018; Longuevergne et al., 2010) en utilisant les bandes K et Ka,

- Récepteur GPS Black Jack (Black Jack Global Positioning System Receiver) détermine la position des satellites au dessus de la Terre avec une précision centimétrique (Cooley and Landerer, 2019) et la vitesse de chaque satellite,
- ACC (Super STAR Accelerometer) mesure les forces non-gravitationnelles qui s’appliquent aux satellites (pression radiative du Soleil et de la Terre, traînée atmosphérique),
- SCA (Star Camera Assembly) est un ensemble de deux caméras qui déterminent la position des satellites par rapport aux étoiles,
- LRR (Laser Retro Reflector) détermine avec précision l’orbite des satellites.

La mission GRACE s’est terminée le 12 octobre 2017, et est succédée par GRACE Follow-On (GRACE-FO), lancée le 22 mai 2018. Elle est très similaire à GRACE, à deux exceptions principales près (<https://podaac.jpl.nasa.gov/GRACE-FO>, dernier accès le 11 mai 2022) :

- Le récepteur GPS Black Jack est remplacé par le récepteur TriG-RO (Tri-GNSS (GPS+Galileo+GLONASS) Radio Occultation) capable de capter non seulement les signaux de la constellation GPS comme Black Jack, mais aussi d’autres signaux GNSS (Galileo, CDMA GLONASS, Compass), et ainsi de déterminer la position des satellites avec plus de précision,
- Le système SCA comportent trois caméras au lieu de deux,
- GRACE-FO dispose d’un interféromètre LRI (*Laser Ranging Interferometer*), permettant d’améliorer la précision de la mesure de la distance entre les deux satellites d’un facteur 10 (Cooley and Landerer, 2019).

### 2.3.1.2 Les données GRACE

GRACE, par la configuration particulière de sa constellation, permet d’estimer les variations spatiales et temporelles du champ gravitationnel terrestre (ou potentiel gravitationnel), qui reflètent les mouvements de masse, intégrés sur l’axe vertical sans limite de profondeur. Les variations du champ gravitationnel terrestre à une petite échelle de temps (quelques décennies) sont principalement dues aux variations de masse des stocks d’eau dans les enveloppes fluides de la Terre (Tapley et al., 2004a; Wahr et al., 2004). Elles peuvent néanmoins être liées à d’autres phénomènes comme le déplacement de masses dû à des tremblements de terre (Chen et al., 2007; De Viron et al., 2008; Panet et al., 2007), l’ajustement isostatique (Chambers et al., 2010; Panet et al., 2007; Riva et al., 2009) ou l’affaissement du sol suite à l’exploitation pétrolière (Gido et al., 2020).

Les vitesses, accélérations et positions des satellites mesurées par les instruments des satellites sont reportés dans les produits de niveau 1B (produits L1B). Des solutions mensuelles de niveau 3 (L3) d’anomalies de stock d’eau continental total (cm) (ou TWS : *Total Water Storage*) par rapport à une période donnée sont calculées à partir des produits L1B et mises à disposition par le *Science Data Center*. Il est constitué des trois centres officiels de production : le GFZ (*GeoForschungs Zentrum*, Potsdam, Allemagne), le CSR (*Center for Space Research*, Austin, Texas, Etats-Unis) et le JPL (*Jet Propulsion Laboratory*, Pasadena, California, Etats-Unis) (Frappart and Ramillien, 2018). Deux grands types de solutions L3 sont mises à dispositions : les produits dérivés des harmoniques sphériques (L3SH) et les mascons (L3M).

## Solutions dérivées des harmoniques sphériques (L3SH)

**Produits L2 : harmoniques sphériques.** Un premier moyen d’obtenir des anomalies mensuelles de TWS (TWSA) est de modéliser le potentiel gravitationnel terrestre sous la forme d’une somme d’harmoniques sphériques (analogue à une représentation en série de Fourier pour une géométrie sphérique) à partir des trajectoires des satellites (données L1B). Un potentiel gravitationnel variable dans le temps peut s’écrire comme en équation 2.3 (Schmidt et al., 2008) :

$$V(r, \theta, \lambda, t) = \frac{GM}{R} \sum_{l=2}^{l_{max}} \left(\frac{R}{r}\right)^{l+1} \sum_{m=0}^l (\bar{C}_{lm}(t) \cos(m\lambda) + \bar{S}_{lm}(t) \sin(m\lambda)) \bar{P}_{lm}(\cos \theta) \quad (2.3)$$

avec  $R$  le rayon moyen de la Terre,  $G$  la constante gravitationnelle,  $M$  la masse de la Terre.  $r, \theta, \lambda$  sont les coordonnées sphériques dans un référentiel terrestre du point où le potentiel est évalué ( $r =$  rayon en mètres avec  $r \geq R$ ,  $\theta =$  colatitute comprise entre  $0 - \pi$  à partir du pôle nord,  $\lambda =$  longitude comprise entre  $0 - 2\pi$ ).  $\bar{P}_{lm}$  désigne les fonctions de Legendre.  $l$  est le degré et  $m$  l’ordre azimutal.  $m$  représente le nombre de méridiens (iso-longitudes) et  $l - m$  le nombre de parallèles (iso-latitudes). Plus le degré  $l_{max}$  du modèle est grand, meilleure est la résolution du modèle (un ordre et un degré de 50 donnent une résolution de 400 km (Schmidt et al., 2008)).  $\{C_{lm}(t), S_{lm}(t)\}$  sont les coefficients harmoniques (ou coefficients de Stokes) et représentent la distribution de la masse responsable du champ de pesanteur à l’instant  $t$ .

Dans les faits, les produits L1B ne sont pas utilisés pour caler un modèle de potentiel gravitationnel complet en raison notamment du nombre insuffisant de trajectoires des satellites, mais pour calculer une correction par rapport à un modèle de fond a priori (Bettadpur, 2018; Schmidt et al., 2008). Les coefficients harmoniques corrigés sont reportés dans les produits de niveau 2 (L2).

L’estimation des coefficients harmoniques corrigés se fait à un degré et un ordre suffisamment élevés pour une bonne résolution spatiale, mais limités car l’erreur des coefficients augmente avec le degré (Tapley et al., 2004b; Wahr et al., 2004). En effet, la résolution spatiale dépend de la couverture des traces au sol de GRACE, mais une couverture importante se fait au détriment de la résolution temporelle (voir figure 2.4). En guise de compromis entre la résolution spatiale et temporelle, les coefficients harmoniques sont déterminés sur un mois jusqu’à un ordre et degré de 60, pour une résolution d’environ 330 km (Landerer and Swenson, 2012).

**Traitement des produits L2 : vers un produit L3.** Les coefficients de Stokes déterminés jusqu’à l’ordre et degré 60 souffrent de plusieurs défauts. D’abord, ils sont systématiquement affectés d’un phénomène de stries (*striping*) du nord au sud du fait de l’orbite polaire des satellites et de la faible observabilité sur le gradient est-ouest qui induit des corrélations entre certains coefficients (Watkins et al., 2015). Ces stries sont éliminées par un filtre pour éliminer les erreurs de corrélation (*destriping*). Ensuite, la représentation en harmoniques sphériques implique un phénomène appelé *spatial leakage* : de par sa faible résolution, elle ne permet pas de rendre compte des véritables variations de masse d’eau à des échelles subcontinentales, et les variations amples de potentiel gravitationnel dans les grands bassins polluent les zones voisines, si bien que le signal ne peut pas être interprété à l’échelle locale. La réduction de ces effets se fait généralement avec un filtre gaussien de plusieurs centaines de kilomètres (300 km pour les données officielles, 250 km pour Tiwari et al. (2009), 500 km pour Seo et al. (2006), 750 km ou 1000 km pour Wahr et al. (2004)) pour lisser les données et réduire le bruit en réduisant l’effet des coefficients de plus haut degré (Landerer and Swenson, 2012). Les variations des coefficients de Stokes sont converties en variations de masse à la surface de la Terre (Wahr et al., 1998). Ces variations de masse résiduelles sur les continents

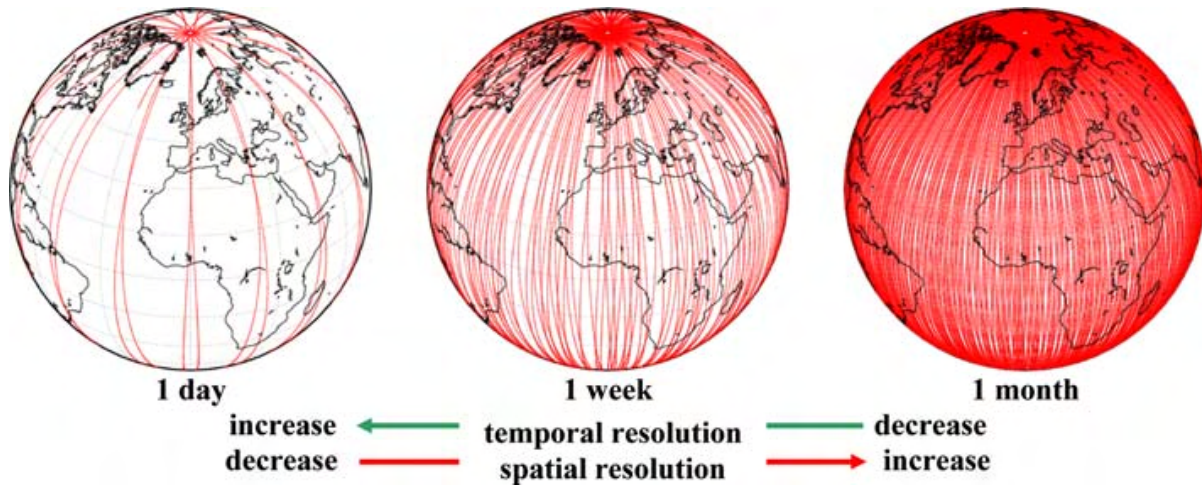


FIGURE 2.4 – Trace au sol de GRACE et compromis entre résolution spatiale et temporelle (Schmidt et al., 2008).

(produits L3) sont principalement dues au stock d’eau terrestre (TWS) (Landerer and Swenson, 2012).

Si les filtres permettent de réduire l’erreur, ils réduisent également la résolution du signal. La puissance du signal en est également atténuée, pouvant empêcher la fermeture du bilan d’eau régional (Landerer and Swenson, 2012). Plusieurs corrections additives ou multiplicatives peuvent être appliquées aux données filtrées de GRACE pour restaurer l’ampleur du signal (Long et al., 2015). Les données L3SH officielles sont fournies avec un facteur multiplicatif  $k$  (*scaling parameter*) appelé paramètre d’échelle, facteur d’échelle ou facteur de gain. Le facteur d’échelle  $\hat{k}$  est calculé par ajustement linéaire entre la série temporelle de TWSA modélisé par le modèle GLDAS-NOAH (Rodell et al., 2004) ( $S_I$ ) et le même signal modifié par les prétraitements appliqués à GRACE ( $S_F$ ), à une échelle de  $1^\circ$  (Landerer and Swenson, 2012) :

$$\hat{k} = \text{Argmin}_k \sum (S_I - k \cdot S_F)^2 \quad (2.4)$$

Le *Science Data Center* fournit les paramètres d’échelle sur une grille de  $1^\circ$  qu’il faut multiplier par les données filtrées de TWSA. Cependant, la résolution des données L3SH reste celle de  $\simeq 300$  km, et les pixels de  $1^\circ \times 1^\circ$  doivent être moyennés sur une zone suffisamment grande (Landerer and Swenson, 2012).

**Avantages et inconvénients.** La grille de facteurs d’échelle suréchantillonne les données L3SH à  $1^\circ \times 1^\circ$  et permet à l’utilisateur de définir arbitrairement une zone d’étude avec plus de flexibilité qu’une grille à 300 km de résolution, et restaure l’ampleur du signal GRACE. Cependant, Landerer and Swenson (2012) soulèvent plusieurs défauts de cette méthode :

- Les facteurs d’échelle sont calculés indépendamment des données GRACE,
- les facteurs d’échelle sont optimisés pour prendre en compte les variations saisonnières et intra-saisonnières, mais ne sont pas adaptés pour des signaux interannuels ou pour observer des tendances sur le long terme,
- les facteurs d’échelle sont calculés sur un modèle ne prenant pas en compte certains flux (fonte des glaciers, abstraction anthropique de l’eau souterraine) et risquent d’être faussés,

notamment sur les zones semi-arides fortement irriguées.

### Solutions mascons (L3M)

Des solutions mascons (pour *mass concentrations*) ont été développées après les solutions harmoniques pour éviter les stries dans les solutions. Les solutions mascons consistent à appliquer des contraintes spatiales et éventuellement temporelles a priori sur les solutions. Des solutions mascons sont produites par deux membres du *Science Data Center*, JPL (Watkins et al., 2015) et CSR (Save et al., 2016), mais aussi par le GFSC (*Goddard Space Flight Center*) (Loomis et al., 2019).

Les solutions mascons sont produites en considérant que le globe peut être divisé en éléments de masse discrets de taille et de forme voulue (mascons sphériques pour JPL, hexagonaux et pentagonaux pour CSR). La contrainte de l'espace par une division en éléments de masse permet d'éliminer les stries induites par l'acquisition exclusivement sur l'axe nord-sud des données GRACE. Cette correction a priori permet d'éviter le *striping* en s'appuyant sur des informations géophysiques plutôt que sur des algorithmes empiriques.

Watkins et al. (2015) décrit deux grands types de solutions L3M contraintes par des mascons (Save et al., 2016; Watkins et al., 2015) :

1. Les mascons sont explicités par une formulation analytique, et les données GRACE (vitesse et/ou accélération des satellites) sont directement utilisées pour estimer les variations de masse sur chaque mascon (cas des mascons JPL (Watkins et al., 2015)),
2. les données GRACE (vitesse et/ou accélération des satellites) sont directement utilisées pour estimer les variations de masse sur chaque mascon, mais chaque mascon est représenté par une expansion en harmoniques sphériques tronquée et a donc une répercussion sur le reste du globe (cas des mascons CSR (Save et al., 2016)).

**Mascons JPL.** Le JPL fournit des mascons mensuels de type 1 (Watkins et al., 2015). Les mascons JPL résolvent le champ gravitationnel global en supposant une répartition de la masse en calottes sphériques de  $3^\circ \times 3^\circ$  et utilisent des données de modèles géophysiques globaux et un cadre Bayésien pour contraindre les variations de masse et éliminer les erreurs de corrélation. La solution obtenue souffre moins de *leakage* que les solutions L3SH. Les mascons obtenus subissent deux étapes de post-traitement :

1. La réduction du phénomène de *leakage* sur les mascons côtiers (CRI - *Coastal Resolution Improvement*) qui consiste à séparer les portions océaniques et terrestres des mascons côtiers à l'aide de modèles (Scanlon et al., 2016; Wiese et al., 2016),
2. Le calcul d'un facteur d'échelle, similaire à celui des données L3SH, à partir du modèle GLDAS-CLM (Community Land Model) (Watkins et al., 2015; Wiese et al., 2016). Ce facteur d'échelle cherche à restaurer le signal perdu par agrégation en mascons de  $3^\circ \times 3^\circ$ . Il est calculé entre les données CLM moyennées spatialement sur l'emplacement des mascons et les données CLM natives. La résolution des facteurs d'échelle est à  $0.5^\circ$  en raison du décalage entre les mascons, localisés sur des parallèles de  $0.5^\circ$ , et le modèle CLM à une résolution de  $1^\circ$ , obligeant la désagrégation préalable des données CLM à  $0.5^\circ$  pour l'ajustement du paramètre d'échelle. Comme pour les données L3SH, la résolution des facteurs d'échelle ne change pas la résolution effective des mascons à  $3^\circ \times 3^\circ$ . Les pixels de  $0.5^\circ$  doivent donc être moyennés sur la zone définie par l'utilisateur, permettant une plus grande flexibilité que l'utilisation de mascons à  $3^\circ$  sur une grille pré-définie.

**Mascons CSR.** Le CSR fournit des solutions mensuelles en mascons hexagonaux (et 12 pentagonaux) à une résolution de  $1^\circ$  (120 km de long) non contraints par des modèles géophysiques (Save et al., 2016). Comme pour les solutions précédentes, ces mascons doivent être moyennés sur des zones au moins aussi grandes que la résolution effective de GRACE (300 km).

**Approches régionales.** Les solutions L3SH ou L3M de CSR ne peuvent pas être simplement interprétées pixel à pixel comme des images satellitaires classiques du fait de la dépendance des signaux induite par le modèle des harmoniques sphériques (Longuevergne et al., 2010). Certaines approches régionales sont ainsi développées pour éviter le *striping* et la contamination du signal par d'autres parties du monde (Frappart et al., 2019; Ramillien et al., 2012).

### 2.3.1.3 Données utilisées, interprétation et utilisation

Les travaux présentés dans cette thèse utilisent les solutions mascons JPL RL06M, comme il est recommandé pour la plupart des usages de GRACE en hydrologie (<https://grace.jpl.nasa.gov/data/choosing-a-solution/>; dernier accès le 10 mai 2022). Après multiplication des mascons par la grille de facteurs d'échelle, les pixels de  $0.5^\circ$  sont moyennés dans une zone d'au moins  $300 \times 300 \text{ km}^2$ . Les données sont téléchargées sur <https://podaac.jpl.nasa.gov> (dernier accès le 11 mai 2022). La série temporelle des données GRACE souffre de l'absence de données sur plusieurs mois en raison des batteries vieillissantes des satellites GRACE en fin de vie (de 2011 à 2017) et des quelques mois de délai entre la fin de la mission GRACE et la mission GRACE-FO.

Les données RL06M sont des séries mensuelles d'anomalies de TWS (TWSA) par rapport à la période 2004-2009 (période par défaut du produit). Elles peuvent être interprétées comme en équation 2.5 :

$$\text{TWSA}_t = \text{TWS}_t - \overline{\text{TWS}}_{2004-2009} \quad (2.5)$$

avec  $\text{TWS}_t$  le stock d'eau total au mois  $t$ ,  $\overline{\text{TWS}}_{2004-2009}$  la moyenne de stock d'eau total sur tous les mois entre janvier 2004 et décembre 2009, et  $\text{TWSA}_t$  l'anomalie sur cette période.

La période de référence de l'anomalie peut être facilement changée en soustrayant la moyenne de TWSA de la période choisie (Cooley and Landerer, 2019). Par exemple, l'anomalie de TWS par rapport à la période 2010-2015  $\text{TWSA}'_t$  est :

$$\text{TWSA}'_t = \text{TWSA}_t - \overline{\text{TWSA}}_{2010-2015} \quad (2.6)$$

L'anomalie de TWS peut être interprétée comme la somme des anomalies des différents stocks d'eau, des couches profondes aux plus superficielles (équation 2.7) :

$$\text{TWSA} = \text{GWSA} + \text{SMA} + \text{SWA} + \text{SNA} + \text{CWSA} \quad (2.7)$$

avec :

- TWSA : anomalie du stock d'eau total mensuel (TWS - *Total Water Storage*),
- GWSA : anomalie du stock d'eau souterraine (GWS - *Ground Water Storage*),
- SMA : anomalie d'humidité du sol dans la zone non saturée (ou zone vadose) (SM - *Soil Moisture*),
- SWA : anomalie du stock d'eau de surface (SW - *Surface Water*),
- SNA : anomalie du stock de neige (SN - *SNow*),
- CWSA : anomalie du stock d'eau dans la biomasse (CWS - *Canopy Water Storage*).

## 2.3.2 Humidité du sol

### 2.3.2.1 SMOS

La mission SMOS (*Soil Moisture and Ocean Salinity*) est une mission de l'ESA, du CNES et du CDTI (agence spatiale d'Espagne) visant à cartographier l'humidité du sol en surface (SSM, *Surface Soil Moisture*) et la salinité des océans. Le satellite a été lancé le 2 novembre 2009 et possède un radiomètre qui mesure les micro-ondes émises depuis la Terre avec une fréquence sensible au contenu en eau ([https://www.esa.int/Applications/Observing\\_the\\_Earth/FutureEO/SMOS/Introducing\\_SMOS](https://www.esa.int/Applications/Observing_the_Earth/FutureEO/SMOS/Introducing_SMOS); <https://smos.cnes.fr/en/SMOS/index.htm>, dernier accès 12 avril 2022). Pour ces travaux, les données L3SM RE07 du Centre Aval de Traitement des Données SMOS (CATDS, <https://www.catds.fr/>, dernier accès le 11 mai 2022) (Al Bitar et al., 2017) sont utilisées. Elles fournissent entre autres la proportion humidité (en  $\text{m}^3/\text{m}^3$ ) dans les cinq premiers centimètres de sol avec une incertitude d'environ  $0.04 \text{ m}^3/\text{m}^3$  à une résolution de 50 km (distribuées à 25 km sur la grille EASE2 (Brodzik et al., 2012)) pour les orbites ascendantes et descendantes (matin et soir) tous les trois jours.

Les données L4 d'humidité en zone racinaire (RZSM) (en  $\text{m}^3/\text{m}^3$ ) (Al Bitar et al., 2013), produites à partir des données SMOS L3SM RE04, sont également utilisées. Elles sont obtenues en (i) modélisant l'infiltration de l'eau dans les couches 5-40 cm et 40-200 cm par deux modèles de bilan hydriques différents (filtre exponentiel et équation de Richards respectivement) et (ii) en intégrant le taux de transpiration avec la méthode de la FAO (Sheffield and Wood, 2011), le NDVI à 16 jours de MODIS et les données climatiques de NCEP (National Centers for Environmental Prediction).

### 2.3.2.2 SMAP

La mission SMAP (*Soil Moisture Active Passive*) vise à cartographier l'humidité du sol à une résolution supérieure à SMOS. Le satellite, lancé le 31 janvier 2015, possède un radiomètre d'une résolution de 40 km et un radar à bande L d'une résolution de 3 km. La synergie de ces instruments devait permettre la désagrégation des données d'humidité du sol à une résolution d'environ 10 km (Das et al., 2014). Depuis l'arrêt du fonctionnement du radar en juillet 2015, les données SMAP sont issues du radiomètre seulement et sont produites et fournies à une résolution similaire à celle de SMOS (<https://smap.jpl.nasa.gov/>, dernier accès le 11 mai 2022). Dans ces travaux, les données L3 V006 d'humidité dans les 5 premiers centimètres du sol (en  $\text{m}^3/\text{m}^3$ ) (SMAP SPL3SMP.006) ainsi que les données L4 d'humidité en zone racinaire (SPL4SMGP.004) sont utilisées. Les données L3 sont distribuées à une résolution de 36 km sur la grille EASE2 (Brodzik et al., 2012) pour les orbites ascendantes et descendantes tous les deux à trois jours. Les données L4 sont fournies toutes les trois heures à une résolution de 9 km. Elles sont obtenues par assimilation de la température de brillance de SMAP dans le modèle GEOS-5 CLSM (*Catchment Land Surface Model*) avec un filtre de Kalman d'ensemble (Reichle et al., 2017).

Les données L3 et L4 sont téléchargeables sur les plateformes <https://appears.earthdatacloud.nasa.gov> et <https://nsidc.org/data/smap/smap-data.html> (dernier accès le 11 mai 2022).

### 2.3.2.3 ESA CCI

Les données CCI (*Climate Change Initiative*) de l'ESA (*European Space Agency*) fournissent des séries temporelles globales d'humidité de surface (en  $\text{m}^3/\text{m}^3$ ) issues de la fusion des mesures de différents satellites. Trois jeux de données sont disponibles : le produit "ACTIVE" (actif) et le



produit “PASSIVE” (passif) qui sont créés en fusionnant les produits d’humidité des diffusomètres et des radiomètres respectivement, et le produit “COMBINED” (combiné) qui fusionne ces deux jeux de données. Ces jeux de données ont l’avantage de couvrir une période temporelle de 40 ans, de Novembre 1978 au 31 décembre 2020, alors que les satellites de SMOS et SMAP ne sont en activité que depuis 2010 et 2015 respectivement. Les données “PASSIVE” V2 sont utilisées pour ces travaux car elles combinent les données de capteurs similaires à ceux de SMOS et SMAP. Ce sont des données journalières fournies à une résolution de  $0.25^\circ$ . Elles sont téléchargeables sur <https://www.esa-soilmoisture-cci.org/> (dernier accès le 11 mai 2022).

### 2.3.3 Indice de végétation et évapotranspiration MODIS

Les deux satellites Terra et Aqua ont été lancés respectivement en 1999 et 2002. Ils sont équipés du capteur MODIS (*Moderate-Resolution Imaging Spectroradiometer*) qui acquiert les données dans 36 bandes spectrales de 0.4 à 14.4  $\mu\text{m}$  avec une résolution spatiale de 250 m à 1 km. Chacun des deux satellites cartographie la Terre tous les un ou deux jours avec un temps de passage décalé de 3 heures environ (10 : 30 pour Terra et 13 : 30 pour Aqua).

Ces travaux utilisent deux produits de MODIS : le NDVI (*Normalized Difference Vegetation Index*) mensuel à 1 km (MOD13A3v006) (Didan et al., 2015) et l’évapotranspiration (ET) tous les huit jours à 500 m (MOD16A2GFv006) (Running et al., 2019), téléchargés sur <https://appears.earthdatacloud.nasa.gov> (dernier accès le 11 mai 2022).

Le NDVI est un indice calculé à partir des réflectances des bandes spectrales rouge (R) et proche infrarouge (NIR) (équation 2.8) dont la valeur se rapproche de 1 avec la densité de végétation verte.

$$\frac{NIR - R}{NIR + R} \quad (2.8)$$

L’ET est calculée comme la somme de l’évaporation de la canopée humide, la transpiration de la canopée sèche et l’évaporation du sol tous les huit jours par l’algorithme MOD16 avec une approche de bilan énergétique basé sur l’équation de Penman-Monteith (Monteith, 1965). L’estimation de l’ET se fait à partir de données météorologiques journalières (radiation, température et pression de l’air, humidité) et d’autres données produites à partir des réflectances MODIS à 8 ou 16 jours (occupation du sol, indice foliaire, albédo) (Mu et al., 2007, 2011; Running et al., 2019).

### 2.3.4 Eau de surface

La détection de l’eau de surface sur des images satellitaires optiques ou radar permet de cartographier les étendues d’eau. Les travaux présentés dans cette thèse utilisent (i) les produits de surface en eau saisonniers dérivés des missions Landsat de la base de données *Global Surface Water* et (ii) des cartes d’eau produites à partir des images optiques récentes Sentinel-2. Les cartes d’eau dérivées de Sentinel-2 sont utilisées en combinaison avec des modèles numériques de terrain (MNT) Pléiades très haute résolution pour quantifier les volumes des petits réservoirs composant le système de récupération des eaux de pluie (chapitre 5).

#### 2.3.4.1 *Global Surface Water*

*Global Surface Water* est une base de données produisant des cartes de surface en eau à partir des images optiques de Landsat à une résolution de 30 m depuis 1984 à l’échelle globale (Peckel et al., 2016). Plusieurs produits sont proposés par la base de données et téléchargeables sur

<https://global-surface-water.appspot.com/> (dernier accès le 8 mai 2022), notamment l'historique mensuel des surfaces en eau, l'occurrence (% d'observations d'eau par rapport au nombre d'observations valides) et l'étendue maximale d'eau observée sur la série temporelle. Les travaux présentés dans le chapitre 5 utilisent l'étendue maximale, qui indique les pixels ayant été classés au moins une fois en eau sur toute la série temporelle.

Les cartes d'eau mensuelles sont produites à l'échelle globale à partir des réflectances de plus de trois millions d'images d'archives USGS L1T produites à partir des images Landsat 5, 7 et 8 (Dwyer, 2019). D'autres variables sont ajoutées aux réflectances pour la détection de l'eau, en particulier, le NDVI et la transformation en modèle de couleur HSV (Teinte-Saturation-Valeur ou *Hue-Saturation-Value*), plus adaptée à l'analyse des images, à partir des réflectances infrarouge à courte longueur d'onde (SWIR2), proche infrarouge (NIR) rouge, verte et bleue.

La construction du modèle de détection d'eau est développée sur une base de données construite sur 64 254 échantillons obtenus par interprétation visuelle, représentative de la variabilité spectrale des classes eau et non-eau à l'échelle globale. Les enveloppes des groupes de pixels de même classe (clusters) sont tracées manuellement sur l'espace multidimensionnel des attributs des pixels et les équations caractérisant les clusters sont obtenues par triangulation de Delaunay (Delaunay, 1934). Des confusions peuvent survenir en raison de la superposition des clusters, et la classification est raffinée par un système expert sous la forme d'un arbre de décision séquentiel combinant les caractéristiques multispectrales et multitemporelles des images Landsat, des données auxiliaires (modèles numériques de terrain, ombres, coulées de lave, etc.), et d'autres éléments de connaissance experte pour les indéterminations (par exemple un pixel fréquemment dans le cluster en eau sera plus vraisemblablement en eau). Ce système expert est raffiné par itérations jusqu'à que les améliorations observées ne soient plus notables, puis appliqué à l'échelle globale sur toute la série temporelle.

La validation des cartes obtenues est faite sur 40 124 points de contrôle distribués géographiquement sur l'ensemble de la série temporelle. L'erreur de commission (pixels non eau classés en eau) est estimée à moins de 1% sur 12 856 échantillons, et l'erreur d'omission (pixels eau classés non eau) à moins de 5% sur 27 268 échantillons.

#### 2.3.4.2 Modèle numérique de terrain très haute résolution (Pléiades)

Quatre MNT à 2 m de résolution produits à partir d'images stéréoscopiques Pléiades à 50 cm de résolution sont utilisés. Trois d'entre eux sont produits à partir d'images acquises en juin 2016 (disponibles sur le catalogue d'Airbus Defense and Space, emprises (a), (b) et (c) sur la figure 2.5) et le dernier a fait l'objet d'une acquisition en juin 2019 (emprise (d)), et couvrent respectivement 444, 322, 463 et 584 km<sup>2</sup>.

Les MNT de chaque zone ont été générés avec des paires d'images stéréoscopiques panchromatiques de Pleiades et le logiciel ASP (Ames Stereo Pipeline) de la NASA (Beyer et al., 2019). Les images ont toutes été acquises à la fin de la saison sèche où la quasi-absence d'eau au fond des réservoirs permet l'accès à leur bathymétrie. Les zones du MNT présentant une absence de donnée sont comblées par interpolation, notamment les zones en eau (lit des rivières ou eau résiduelle dans les tanks par exemple), les nuages ou les sols très clairs dont la réflectance est saturée.

#### 2.3.4.3 Sentinel-2

Des cartes d'eau sont fabriquées à partir d'une série temporelle d'images optiques Sentinel-2 L2A. Les données Sentinel-2 L2A sont corrigées par l'algorithme MAJA (Hagolle et al., 2017) et

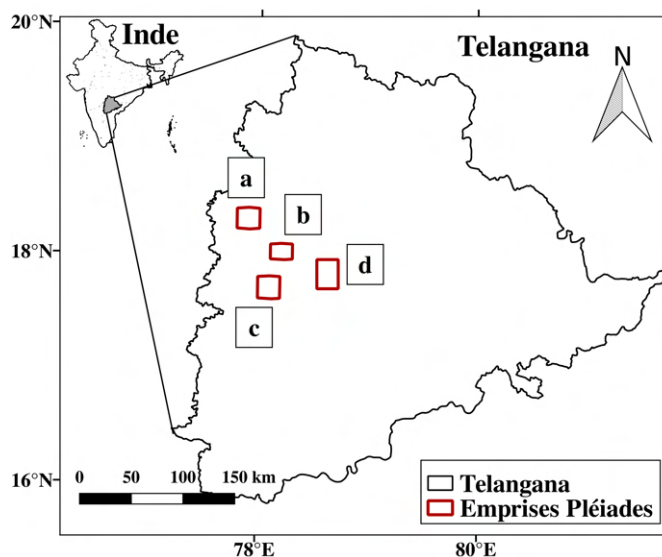


FIGURE 2.5 – Localisation des MNT Pléiades.

téléchargées sur la plateforme <https://www.theia-land.fr/> (dernier accès le 16 mai 2022). Ces cartes sont réalisées sur quatre tuiles Sentinel-2 qui couvrent les quatre emprises Pléiades (section 2.3.4.2) : T44QGA (40 images), T44QHA (44 images), T44QHV (44 images) et T44QKE (71 images).

Pour chaque date d’acquisition, un modèle *random forest* (Breiman, 2001) est entraîné et validé avec un échantillon aléatoire de 10 000 pixels eau et non-eau, et utilisé pour faire une classification binaire des images. La sélection des échantillons peut être automatisée avec des bases de données comme l’occurrence de *Global Surface Water* (sélection des pixels d’eau permanente), cependant l’eau permanente ou quasi permanente est rare dans cette partie de l’Inde et ne permet pas d’échantillonner différents types d’eau. Les échantillons en eau sont donc identifiés avec un seuillage de NDWI (Normalized Difference Water Index) à 0 (Chowdary et al., 2008; McFEETERS, 1996; Vanthof and Kelly, 2019). Les incertitudes du seuillage de NDWI est compensé par la robustesse de l’algorithme *random forest* au bruit dans les labels jusqu’à 20% (Pelletier et al., 2017). Les échantillons non-eau sont sélectionnés aléatoirement à une distance tampon des surfaces ayant un NDWI négatif (surfaces supposément en eau).

La précision de ces cartes est évaluée contre des images optiques très haute résolution de Google Earth Pro (acquises par Maxar Technologies ou CNES/Airbus). Les surfaces en eau de 16 tanks ont été détournées à la main à différentes dates qui correspondent à celle d’une image de Sentinel-2 à cinq jours près, pour un total de 29 points. Les surfaces dérivées des images Sentinel-2 sont généralement sous-estimées par rapport aux images haute résolution, avec une erreur relative moyenne de +17.5 %.

### 2.3.5 Précipitations TRMM

La mission TRMM (*Tropical Rainfall Measurement Mission*), lancée en 1997, quantifie les précipitations à l’aide de plusieurs instruments : un radar météorologique (PR - *Precipitation Radar*) d’une résolution de 5 km, un capteur passif micro-ondes (TMI - *TRMM Microwave Imager*) mesurant l’intensité des radiations aux fréquences 10.7, 19.4, 21.3, 37, et 85.5 GHz à une résolution de 50 km, et un capteur infrarouge (VIRS - *Visible Infrared Scanner*) mesurant les réflectances à

0.63, 1.6, 10.8 et 12 micromètres à une résolution de 2.4 km. La mission TRMM fournit plusieurs produits de précipitations estimées avec différents algorithmes. Les travaux présentés dans cette thèse utilisent les données 3B43 Version 7 qui fournissent le taux de précipitation mensuel moyen horaire (en mm/heure) sur une grille de  $25^\circ \times 25^\circ$  entre les latitudes  $50^\circ\text{S}$  et  $50^\circ\text{N}$ . Elles sont produites à partir des estimations de précipitations mensuelles avec une approche basée sur la loi du rayonnement de Planck et les données micro-ondes TMI, combinées à des données de pluviomètres (MacRitchie (2017), <https://gpm.nasa.gov/missions/TRMM/satellite/TMI>, dernier accès le 7 juin 2022). Les taux de précipitations moyens mensuels sont fournis avec une erreur quadratique moyenne (RMSE - *Root Mean Squared Error*) spatialisée calculée par rapport aux mesures pluviométriques. Le RMSE varie de 0.02 à 0.33 mm/heure (0.11 mm/heure en moyenne) sur le Telangana durant les mois de fortes précipitations (de juin à octobre) de 2002 à 2020, et peut représenter de 19 à 151% du taux estimé (40% en moyenne). Les données 3B42 Version 7 sont également utilisées, et fournissent des taux de précipitations tous les 3 heures à une résolution de  $0.25^\circ$  issus de la fusion des estimations des différents capteurs avec les mesures pluviométriques.

## 2.4 Modèles

Les variables hydrologiques ne sont généralement pas directement accessibles par télédétection, et leur estimation à partir de données satellitaires nécessite des modèles de transfert radiatif ou physiques et/ou des données de forçage, comme c'est le cas pour les précipitations de TRMM, l'humidité de surface de SMOS et SMAP, ou l'ET de MODIS. Les modèles hydrologiques comme les modèles de surface terrestre (LSM - *Land Surface Models*) permettent de décrire les cycles de l'eau et d'énergie à la surface de la Terre, et d'estimer en particulier l'évapotranspiration, le ruissellement ou l'humidité des premiers mètres de sol à partir de données de forçage et/ou de l'assimilation de données satellitaires.

Dans ces travaux, la sensibilité des estimations de SSM et RZSM est évaluée par comparaison des estimations dérivées de la température de brillance de SMOS et SMAP mais également avec les sorties de deux modèles : GLDAS-NOAH et GLEAM. En l'absence de données satellitaires ou de réseau dense de mesures *in situ*, l'évapotranspiration potentielle et l'estimation du ruissellement nécessaires à la fermeture du bilan des échanges hydriques sont également modélisés.

### 2.4.1 GLDAS

GLDAS (*Global Land Data Assimilation System*) est un ensemble de modèles de surface qui génère une grande diversité de variables d'état et de flux de surface en combinant des données satellitaires et des données terrain (Rodell et al., 2004). Quatre modèles GLDAS sont actuellement disponibles : NOAH, CLSM (*Catchment Land Surface Model*), CLM (*Community Land Model*) et VIC (*Variable Infiltration Capacity*). Ces travaux utilisent les sorties du modèle GLDAS-NOAH, qui est le plus fréquemment utilisé dans les études sur les données GRACE.

Plusieurs versions de GLDAS sont disponibles, avec (version 2.2) ou sans assimilation de données (versions 2.0 et 2.1). Pour ces travaux, les données version 2.1, indépendantes des données GRACE, seront préférées à la version 2.2, qui les assimile. En particulier, les données mensuelles de SSM, RZSM et de ruissellement à  $0.25^\circ$  de 2002 à 2021 sont utilisées. Les données sont accessibles sur le GES DISC (*National Aeronautics and Space Administration Goddard Earth Science Data Information and Services Center*, <https://disc.gsfc.nasa.gov/>, dernier accès le 11 mai 2022).

## 2.4.2 GLEAM

Le modèle GLEAM (*Global Land Evaporation Amsterdam Model*) est un ensemble d’algorithmes qui estime l’humidité du sol ainsi que les différentes composantes de l’évapotranspiration à partir de données satellitaires (dont CCI et MODIS) (Martens et al., 2017; Miralles et al., 2011). Ces travaux utilisent les données de la version 3.5b de SSM (SMsurf) et de RZSM (SMroot), distribuées à  $0.25^\circ$  sur la période 2003-2020. Les données sont téléchargeables sur <https://www.gleam.eu/> (dernier accès le 11 mai 2022).

## 2.4.3 Évapotranspiration potentielle

Le jeu de données “RM-OBS/PU Potential Evapotranspiration and Supporting Forcing L4 3-Hourly  $0.25 \times 0.25$  degrees” propose des séries temporelles d’évapotranspiration potentielles calculées à partir de données météorologiques et de rayonnement issues d’une série de réanalyses, de données satellitaires et de données de terrain spatialisées (Peng, 2019; Peng et al., 2019). Trois jeux de données sont fournis : l’évapotranspiration potentielle de l’eau libre calculée par l’équation de Penman (Penman and Keen, 1948), celle calculée par l’équation de Priestley-Taylor (Priestley and Taylor, 1972), et l’évapotranspiration potentielle de référence calculée par l’équation de Penman-Monteith selon l’approche de la FAO56 (Walter et al., 2012). Ces trois séries sont fournies toutes les trois heures à une résolution de  $0.25 \times 0.25^\circ$  de 1984 à 2016 par le jeu de données PET\_PU\_3H025\_002, téléchargeable sur <https://disc.gsfc.nasa.gov/> (dernier accès le 11 mai 2022).

## 2.5 Ce qu'il faut retenir

- Une diversité de données satellitaires et de modèles multi-sources permet d'estimer une multitude de variables environnementales à large échelle.
  - En particulier, les données gravimétriques GRACE, les données d'humidité du sol SMOS et SMAP et les modèles numériques de terrain très haute résolution combinés à des séries temporelles de surfaces en eau sont des données uniques pour respectivement le suivi de l'eau souterraine, de l'irrigation et du volume des petites retenues d'eau du RHS.
  - Les différences d'unité, de résolution spatiale et temporelle, d'étendue spatiale et de période d'acquisition requièrent une homogénéisation des données préalable à leur comparaison.
  - Les données *in situ*, disponibles à moins large échelle que les données satellitaires, sont utilisées pour la validation des approches développées dans ces travaux.
-



# Chapitre 3

## Potentiel des données GRACE et SMOS pour le suivi des stocks d'eau souterraine

### Sommaire

---

<b>3.1</b>	<b>Introduction</b>	<b>41</b>
<b>3.2</b>	<b>Homogénéisation des données</b>	<b>43</b>
3.2.1	Résolution spatiale et temporelle, projection	43
3.2.2	Unité	43
3.2.3	Spatialisation des données piézométriques	43
<b>3.3</b>	<b>Évolution des ressources en eau à l'échelle du Telangana</b>	<b>44</b>
3.3.1	Interprétation qualitative des séries temporelles	44
3.3.2	Tendances à long terme	45
<b>3.4</b>	<b>Déconvolution de GRACE à l'échelle du Telangana</b>	<b>47</b>
3.4.1	Équation générale de bilan d'eau	47
3.4.2	Sensibilité du GWS estimé à différents jeux de données	49
3.4.3	Discussion	51
<b>3.5</b>	<b>Potentiel de SMOS</b>	<b>52</b>
3.5.1	Un cycle annuel d'humidité bimodal	52
3.5.2	Interprétation à l'échelle de SMOS	54
3.5.3	Quantification de l'eau utilisée pour l'irrigation avec SMOS	57
3.5.4	Indicator of flood-irrigated crops from SMOS and SMAP soil moisture products in southern India	59
<b>3.6</b>	<b>Ce qu'il faut retenir</b>	<b>65</b>

---

### 3.1 Introduction

L'étude et la compréhension des phénomènes hydrologiques et anthropiques est essentielle pour une meilleure gestion des ressources en eau, en particulier dans les zones semi-arides fortement anthropisées où les processus hydrologiques naturels sont perturbés par les activités humaines.



L'Inde est un exemple flagrant de la surexploitation des ressources pour l'irrigation. 80% de l'eau consommée est destinée à l'irrigation, dont 65% provenant de l'eau souterraine. En particulier, dans l'État du Telangana (Inde du Sud), l'irrigation du riz représente 80% de la consommation en eau souterraine (Ferrant et al., 2014; Perrin et al., 2012). Les pénuries d'eau sont fréquentes du fait de la fréquence des épisodes de recharges et de l'intensité des pompages.

Le suivi local de l'eau souterraine sur le Telangana est assuré par un réseau de plus de 1000 piézomètres d'observation, mais reste difficile du fait de la forte hétérogénéité spatiale de la mesure, elle même représentative d'un état très localisé de la nappe. Cette nature ponctuelle de la donnée nécessite un grand nombre d'observations pour rendre plus robuste l'estimation de l'état de la nappe au niveau régional. Contrairement aux études de terrain coûteuses et localisées, la télédétection permet de restituer directement ou indirectement une diversité de phénomènes hydrologiques (ruissellement, recharge, stockage transitoire de surface ou souterrain, évapotranspiration liée à la végétation naturelle et irriguée) dans le temps et dans l'espace à grande échelle. Une variété de données spatialisées issues de satellites ou de modèles permet le suivi des stocks et flux d'eau (anomalie du stock d'eau total de GRACE, l'humidité de surface de SMOS, précipitations TRMM) et de la végétation (NDVI de MODIS, cartes d'occupation du sol produites à partir des données optiques Sentinel-2) avec une grande diversité de capteurs, de résolutions temporelles, et d'incertitudes de restitution des variables environnementales ou hydro-agronomiques.

Les données GRACE et SMOS sont des données uniques en hydrologie. Les données gravimétriques de GRACE fournissent depuis 2002 des anomalies mensuelles de stock d'eau sans limite de profondeur, incluant notamment les ressources en eau souterraine. Les données SMOS suivent quant à elles tous les trois jours l'humidité des cinq premiers centimètres du sol, signature des phénomènes naturels (précipitations) et des activités anthropiques (irrigation). Ces données d'exception apportent un suivi spatialisé, régulier et global depuis 20 ans pour GRACE et 12 ans pour SMOS de ces deux stocks essentiels pour la compréhension des phénomènes hydrologiques et des usages de l'eau, auparavant réalisable uniquement par des réseaux d'observation généralement épars ou inexistant car coûteux à mettre en place.

Malgré les apports des données satellitaires (séries temporelles spatialisées longues à grande échelle), la précision des données *in situ* demeure essentielle pour la compréhension des systèmes à échelle locale, et ces dernières sont utilisées pour la validation des résultats issus d'approches utilisant des mesures satellitaires et des sorties de modèles.

Ce chapitre s'intéresse donc à l'intercomparaison et la mise en cohérence des dynamiques temporelles et spatiales des données disponibles (*in situ*, satellitaires, modèles) sur la zone granitique du Telangana afin de déterminer l'impact de l'irrigation sur les ressources en eau souterraine. Une attention particulière est apportée au potentiel des données GRACE et SMOS pour le suivi des ressources en eau souterraine et de ses usages, et à comment ces séries temporelles spatialisées et régulières pourraient permettre de lever certains verrous de connaissances hydrologiques.

Une première partie décrit le travail préliminaire d'homogénéisation nécessaire pour la comparaison d'une multiplicité de données hétérogènes. Une seconde partie s'intéresse à l'évolution temporelle du stock d'eau total de GRACE et d'eau souterraine des données *in situ* à l'échelle du Telangana, et la troisième partie à la quantification du stock d'eau souterraine par GRACE et d'autres données satellitaires ou sorties de modèles spatialisées. La dernière partie est consacrée au potentiel des données SMOS pour la détection et la quantification des activités humaines d'irrigation, et comporte un article soumis à *IEEE Geoscience and Remote Sensing Letters*.

## 3.2 Homogénéisation des données

### 3.2.1 Résolution spatiale et temporelle, projection

La comparaison des données hétérogènes traitées dans ce chapitre nécessite une homogénéisation préalable de leurs résolutions spatiale et temporelle. Elles sont adaptées pour correspondre à celles des données GRACE, qui sont les plus limitantes en terme de résolution spatiale (300 km) et temporelle (1 mois). Comme expliqué en section 2.3.1.3, les données GRACE RL06M sont fournies avec une grille de facteurs d'échelle de  $0.5^\circ$  ce qui permet à l'utilisateur de moyennner les pixels ré-échantillonnés dans une zone d'étude arbitrairement choisie de surface suffisamment grande (au minimum  $300 \times 300 \text{ km}^2$ ). La zone d'étude délimitée ici concerne la zone granitique du Telangana, qui constitue un aquifère hétérogène dont les propriétés à l'échelle régionale, notamment la porosité efficace, a fait l'objet d'études du BRGM et du NGRI (Dewandel et al., 2017a; Maréchal et al., 2006). Cette zone, détournée à la résolution des données GRACE suréchantillonnées ( $0.5^\circ$ ), couvre une surface de  $113\,000 \text{ km}^2$  (zone rose sur la figure 1.2).

Les données déjà spatialisées comme les données satellitaires ou sorties de modèles sont reprojctées dans un même système de projection (ici la projection de GRACE, WGS84) et rééchantillonnées à la même résolution spatiale (ici moyennées dans la zone d'étude de  $113\,000 \text{ km}^2$ ). Les données de résolution temporelle supérieure à GRACE sont agrégées temporellement à un mois.

### 3.2.2 Unité

Les données satellitaires ou sorties de modèles quantifient des stocks et flux d'eau à différentes unités. Par exemple, la pluie de TRMM est en mm, l'humidité de surface de SMOS en fraction d'humidité dans le sol, et l'évapotranspiration de MODIS en  $\text{kg}/\text{m}^2$ . Les données GRACE sont encore une fois les plus limitantes en terme d'unité, puisqu'elles restituent des anomalies (et non une quantité absolue) par rapport à la moyenne sur la période 2004-2009 de stock d'eau total (TWS - *Total Water Storage*) en lame d'eau équivalente en cm. Les hauteurs d'eau quantifiées par les différents jeux de données (pluie, humidité du sol SMOS et SMAP, humidité du sol des modèles, évapotranspiration de MODIS) sont donc converties en anomalies de hauteur d'eau (cm) par rapport à la moyenne d'une même période de référence, adaptée en fonction des données comparées. Par exemple, la période 2004-2009 ne peut pas être utilisée pour comparer les données GRACE aux données SMOS et SMAP, acquises à partir de 2010 et 2015. La période doit être suffisamment longue (plusieurs années) pour approcher la moyenne de la série. L'anomalie  $SA_t$  d'une grandeur  $S$  à l'instant  $t$  par rapport à une période  $P$  est calculée par l'équation 3.1 :

$$SA_t = S_t - \overline{S}_P \quad (3.1)$$

avec  $S(t)$  la quantité  $S$  à l'instant  $t$ ,  $\overline{S}_P$  la moyenne de la quantité  $S$  sur la période  $P$ .

### 3.2.3 Spatialisation des données piézométriques

Les données piézométriques enregistrent une variation mensuelle locale de la profondeur de la nappe sur 1006 puits suivis de 2007 à 2020 sur la zone d'étude. L'interpolation spatiale de ces profondeurs par maille de 5 km permet de caractériser une variation régionale de la nappe. Les hauteurs d'eau sont interpolées par pondération inverse à la distance des puits (méthode IDW - *Inverse Distance Weighting*). L'anomalie de stock d'eau souterraine (GWSA - *Ground Water*

*Storage*) est le produit de l’anomalie de hauteur d’eau (GWLA) et de la porosité efficace de l’aquifère (Sy, *specific yield*) :

$$\text{GWSA}_t = -\text{Sy} \cdot \text{GWLA}_t = \text{Sy} \cdot (\overline{\text{GWL}_P} - \text{GWL}_t) \quad (3.2)$$

Le Sy des différentes couches de l’aquifère granitique du Telangana a fait l’objet d’études du BRGM (Bureau de Recherches Géologiques et Minières) et du NGRI (*Indian National Geophysical Research Institute*) sur plusieurs bassins de tailles variées (58 à 983 km<sup>2</sup>) (Dewandel et al., 2010, 2012, 2017a) mais il n’est pas connu à large échelle. Dans cette étude, le Sy est calibré par ajustement linéaire entre les cartes d’anomalies de hauteur d’eau *in situ* (GWLA) et une estimation du GWSA par des données satellitaires spatialisées  $\text{GWSA}_{sat}$  (différence entre l’anomalie TWS de GRACE et l’anomalie d’humidité en zone racinaire (RZSM) de GLEAM) sur la période 2007-2020. Le Sy est estimé en minimisant le critère des moindres carrés (somme des carrés des écarts entre le GWSA *in situ* et  $\text{GWSA}_{sat}$ ) :

$$\begin{aligned} \text{Sy} &= \text{Argmin}_{\text{Sy}} \sum_t (\text{GWSA}_{sat_t} - \text{GWSA}_{in\ situ_t})^2 \\ &= \text{Argmin}_{\text{Sy}} \sum_t (\text{TWSA}_{\text{GRACE}_t} - \text{RZSMA}_{\text{GLEAM}_t} - \text{Sy} \cdot \text{GWLA}_t)^2 \end{aligned}$$

avec  $\text{GWSA}_{sat_t}$  et  $\text{GWSA}_{in\ situ_t}$  les estimations d’anomalie de stock d’eau souterraine (par rapport à la moyenne sur la période 2007-2020) au mois  $t$  issues respectivement des données satellitaires et de modèle, et des mesures piézométriques,  $\text{TWSA}_{\text{GRACE}_t}$ ,  $\text{RZSMA}_{\text{GLEAM}_t}$ ,  $\text{GWLA}_t$  les anomalies de TWS, RZSM et GWL issues respectivement de GRACE, GLEAM et des mesures piézométriques, et Argmin la fonction désignant l’ensemble des valeurs de Sy pour laquelle la quantité  $(\text{GWSA}_{sat_t} - \text{GWSA}_{in\ situ_t})^2$  atteint sa valeur minimale.

Le Sy est estimé à 4.7%, une valeur légèrement haute par rapport à celle attendue, mais du bon ordre de grandeur et cohérente car intermédiaire entre les Sy des deux couches de l’aquifère granitique (10% pour la saprolite, 1% pour le granite fracturé, voir section 1.2.1).

## 3.3 Évolution des ressources en eau à l’échelle du Telangana

### 3.3.1 Interprétation qualitative des séries temporelles

L’État du Telangana est soumis à un climat tropical sec/semi-aride, caractérisé par une saison de mousson (juin à octobre) durant laquelle 90% des précipitations annuelles se produisent. Ce cycle annuel, visible sur les précipitations TRMM, est également restitué par les séries temporelles du TWS de GRACE et du GWS dérivé des données *in situ* (figure 3.1). Ce climat bimodal marque fortement les fluctuations des variables hydrologiques. Les ressources de surface et souterraines sont plus ou moins reconstituées chaque année selon le cumul des précipitations de la mousson, tandis que l’absence de pluie combinée aux prélèvements pour l’irrigation et à la faible capacité de l’aquifère contribuent à les vider rapidement pendant la saison sèche. L’amplitude des fluctuations annuelles de TWS sur le Telangana, pouvant atteindre 60 cm, est parmi les plus importantes enregistrées par la mission GRACE depuis 2002 à l’échelle mondiale. Les trois séries temporelles de pluie, TWS et GWS sont cohérentes malgré l’hétérogénéité de leurs modes d’acquisition, notamment pour l’identification des années sèches de 2009 et 2016 : les faibles précipitations de la mousson (“mousson sèche”) ne suffisent pas à reconstituer les stocks d’eau totaux et souterrains, qui

sont particulièrement bas pendant la saison sèche suivante. Le cycle hydrologique annuel domine les signaux (alternance saison humide/saison sèche), mais le TWS et les précipitations semblent subir des perturbations climatiques à court terme qui obéissent à un cycle d'environ 3 ans, dus aux phénomènes El Niño et La Niña (Asoka et al., 2017; Vissa et al., 2019).

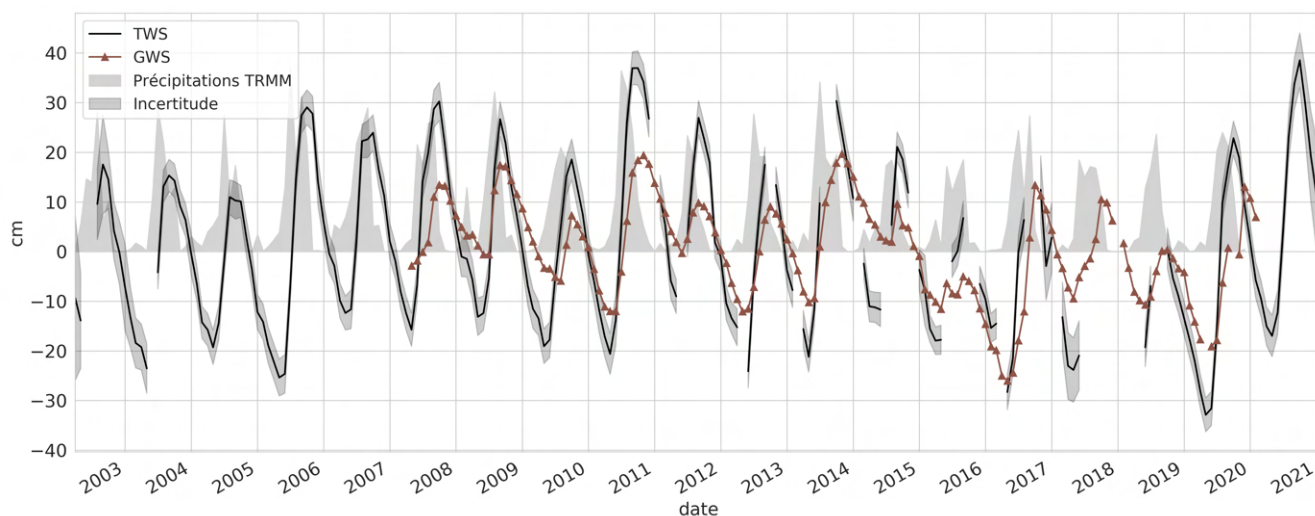


FIGURE 3.1 – Séries temporelles de pluie, TWS et GWS (anomalies 2008-2020). L'enveloppe d'incertitude du TWS correspond aux incertitudes des mascons fournies avec le produit RL06M moyennées dans la zone d'étude.

### 3.3.2 Tendances à long terme

Plusieurs régions du monde, en particulier l'Inde du Nord, sont sujettes à une baisse rapide des ressources en eau souterraines, observation appuyée par des études de tendance sur des puits d'observation (Asoka et al., 2017) ou sur les données GRACE (Chen et al., 2014; Rodell et al., 2018; Tiwari et al., 2009). Les mêmes études dans l'Inde du Sud semblent suggérer le phénomène inverse avec une tendance positive calculée à partir du TWS de GRACE depuis 2002 (Cao and Roy, 2020; Rodell et al., 2018; Tiwari et al., 2009; Vissa et al., 2019), confirmées par une sélection de données piézométriques du CGWB (*Indian Central GroundWater Board*) donnant les mêmes tendances caractérisant une remontée du niveau de nappe dans ces régions. Ces observations sont souvent attribuées à une augmentation des précipitations et au succès de politiques locales (Asoka et al., 2017; Bhanja et al., 2017; Panda and Wahr, 2016), mais vont à l'encontre des études terrain et des enquêtes auprès des agriculteurs qui notent une augmentation du nombre de puits défectueux et un stress hydrique croissant de la végétation (Merriott, 2016; Srinivasan et al., 2015).

Deux principaux phénomènes expliquent ces conclusions contradictoires sur l'Inde du Sud : les caractéristiques des aquifères et les biais d'observations sur les données *in situ*. D'abord, les aquifères granitiques et basaltiques de l'Inde du Sud (Deolankar, 1980; Maréchal et al., 2006) possèdent une faible capacité de stockage par rapport aux aquifères profonds des roches sédimentaires de l'Inde du Nord, où la profondeur d'eau peut atteindre une centaine de mètres (Chidambaram et al., 2009; Singh et al., 2005). Les premiers se remplissent et se vident en grande partie chaque année, alors que les ressources anciennes des seconds disparaissent au fil des années avec leur sur-exploitation (figure 3.2a), et l'interprétation des tendances calculées sur les séries de TWS sur ces deux aquifères diffèrent. Ceci est illustré en figure 3.2a par les tendances des séries de TWS sur

TABLE 3.1 – Tendances annuelles calculées sur les séries de TWS de GRACE dans la littérature.

Référence	Période	Zone	
		Inde du Nord	Inde du Sud
Asoka et al. (2017)	2002-2013	-2 cm/an	+1-2 cm/an
Rodell et al. (2018)	2002-2018	-2 cm/an	+0.7 cm/an
Tiwari et al. (2009)	2002-2008	-2.9 cm/an	-

l'État du Telangana et l'Inde du Nord (États du Punjab, Rajasthan et Haryana). Comme dans Tiwari et al. (2009) et Rodell et al. (2018), le cycle saisonnier est préalablement éliminé du TWS par la méthode STL (*Seasonal Trend decomposition based on Loess*) (Cleveland et al., 1990), laissant un signal résiduel (en gris sur la figure 3.2a). La pente est ensuite calculée avec l'estimateur de Theil-Sen (Gilbert, 1987) et soumise au test non paramétrique modifié de Mann-Kendall (Hamed and Ramachandra Rao, 1998) utilisé pour évaluer le degré de significativité des tendances (Asoka et al., 2017; Jena et al., 2021). La tendance est significative pour une p-valeur du test inférieure à 0.05. Le TWS sur l'Inde du Nord décline très clairement avec une perte de 2.6 cm/an estimée sur la période 2002-2021, cohérente avec les autres estimations trouvées dans la littérature (table 3.1). Le TWS sur le Telangana, en revanche, ne semble pas subir de changement aussi important sur le long terme. En effet, comme indiqué en section 3.3.1, les précipitations annuelles sont soumises aux oscillations El Niño qui induisent une alternance de périodes sèches et humides de quelques années. Aussi, la tendance positive de 1.1 cm visible entre 2002 et 2013 (figure 3.2a), confirmant les résultats de Asoka et al. (2017) et Rodell et al. (2018), rend surtout compte de la succession de la grande sécheresse précédant le début d'acquisition des données GRACE et d'une période plus humide lui succédant. Ces périodes climatiques s'équilibrent sur des périodes plus longues (les faibles précipitations indiquent une nouvelle période sèche entre 2014 à 2020, suivie par des précipitations intenses en 2020), expliquant l'absence de tendance significative sur la période 2002-2021 (figure 3.2a). Aussi, à ce jour, les tendances du signal GRACE sur un aquifère peu profond comme celui Telangana reflètent des périodes climatiques de court terme, et non une évolution durable des ressources en eau (baisse ou augmentation) comme c'est le cas sur un aquifère à forte capacité de stockage comme en Inde du Nord.

Les études de pente sur les puits d'observation sont aussi soumises à plusieurs biais d'observation, que Hora et al. (2019) présentent comme le paradoxe de l'Inde du sud (*Groundwater Recovery Paradox in South India*). Les tendances positives observées sont dues au "biais du survivant", qui consiste à sélectionner pour l'étude seulement les puits répondant à certains critères. En particulier, les puits conservés pour l'analyse sont généralement ceux dont la série temporelle montre une tendance significative (Asoka et al., 2017) ou est la plus complète (Asoka et al., 2017; Bhanja et al., 2016). Or, des séries temporelles lacunaires peuvent être expliquées par des niveaux de remplissage extrêmement faibles, en dessous du capteur et donc indétectables, mais qui sont pourtant une indication essentielle du stress. Quant aux pentes positives observées sur le TWS de GRACE, Hora et al. (2019) suggèrent que la résolution spatiale est insuffisante pour identifier les diminutions locales des stocks d'eau souterraine, ce qui est particulièrement problématique dans le cas d'un aquifère granitique hétérogène comme celui de l'Inde du Sud (Dewandel et al., 2011; Perrin et al., 2011).

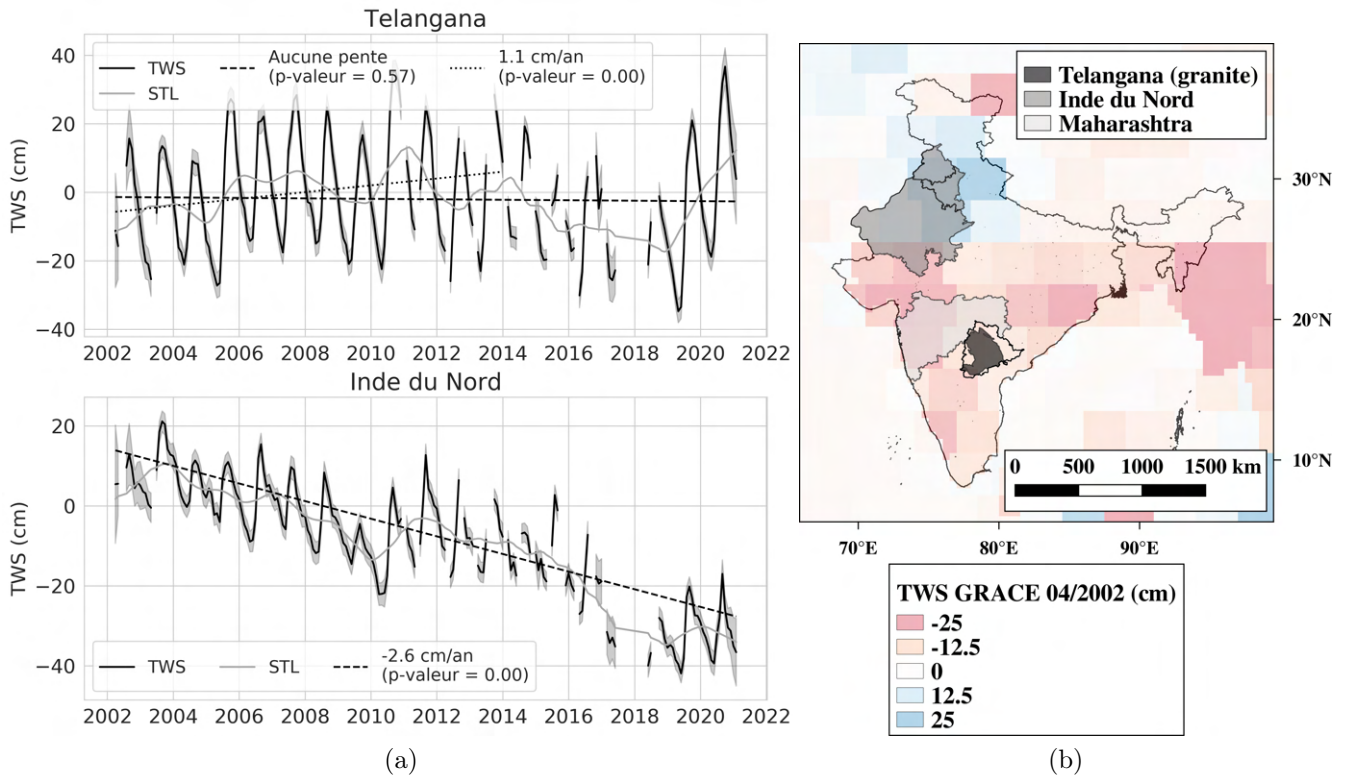


FIGURE 3.2 – (a) Comparaison du TWS de GRACE (anomalies 2004-2009) de l’État du Telangana (Inde du Sud) et des États du Punjab, Rajasthan et Haryana (Inde du Nord). La pente (ligne pointillée) est calculée sur le signal entre 2002 et 2021 dont le cycle annuel est éliminé par STL (*Seasonal Trend decomposition based on Loess*, courbe grise) et soumise à un test de Mann-Kendall. Pour le Telangana, une pente est également calculée entre 2002 et 2013 (ligne pointillée fine). Les tendances significatives sont indiquées par une p-valeur inférieure à 0.05. (b) Localisation de différents États Indiens. Le fond de carte est l’anomalie de TWS de GRACE d’avril 2002 (mascons JPL RL06M).

## 3.4 Déconvolution de GRACE à l’échelle du Telangana

### 3.4.1 Équation générale de bilan d’eau

Les anomalies de TWS de GRACE sont des données informatives précieuses en hydrologie continentale. Plusieurs études ont estimé les variations de GWS à large échelle (Breña-Naranjo et al., 2014; Cao and Roy, 2020; Frappart et al., 2019; Papa et al., 2015; Rodell et al., 2018; Rzepecka and Birylo, 2020; Tiwari et al., 2009; Zhang et al., 2020) à partir du TWS de GRACE, en séparant les contribution des différents réservoirs hydrologiques (déconvolution du TWS) (voir Frappart and Ramillien (2018) pour une synthèse bibliographique). En effet, le produit mensuel d’anomalies de masse de surface fourni par GRACE représente l’anomalie mensuelle du stock d’eau total (équation 3.3) :

$$\text{TWSA} = \text{GWSA} + \text{SMA} + \text{SWA} + \text{SNA} + \text{CWSA} \quad (3.3)$$

avec :

- TWSA : anomalie mensuelle du stock d’eau total (TWS - *Total Water Storage*),

- GWSA : anomalie mensuelle du stock d'eau souterraine (GWS - *Groundwater Storage*),
- SMA : anomalie mensuelle d'humidité du sol dans la zone non saturée (ou zone vadose) (SM - *Soil Moisture*),
- SWA : anomalie mensuelle du stock d'eau de surface (SW - *Surface Water*),
- SNA : anomalie mensuelle du stock de neige (SN - *SNow*),
- CWSA : anomalie mensuelle du stock d'eau dans la biomasse (CWS - *Canopy Water Storage*).

Contrairement à GWSA, les autres composantes de cette équation sont quantifiables par les données satellitaires (détection de la neige, de l'eau de surface, et de l'humidité du sol jusqu'à une certaine profondeur grâce à des modèles ou aux récentes missions satellitaires SMOS et SMAP). Un bilan résiduel permet d'isoler GWSA en soustrayant du TWSA de GRACE les variations de stock des autres réservoirs (équation 3.4) :

$$\text{GWSA} = \text{TWSA} - \text{SMA} - \text{SWA} - \text{SNA} - \text{CWSA} \quad (3.4)$$

### 3.4.1.1 État de l'art

Ce bilan est à adapter en fonction de la zone d'étude. L'influence de l'eau stockée par la biomasse étant en dessous du seuil de détectabilité de GRACE (Rodell et al., 2005), ce terme est rarement considéré dans le bilan (Frappart et al., 2019; Rodell et al., 2009). Quand il est effectivement intégré par souci de fermeture du bilan, il est simulé par des modèles de surface terrestre (LSM, *Land Surface Models*) comme GLDAS (Sun et al., 2019; Vissa et al., 2019). SNA peut être simulée par des modèles comme GLDAS (Rodell et al., 2007) et a fait l'objet d'estimations avec des modèles numériques de terrain (MNT) très haute résolution (Deschamps-Berger et al., 2020), mais est négligée sur les zones rarement ou non enneigées (Frappart et al., 2019; Papa et al., 2015). Le volume d'eau de surface SWA peut être quantifié par imagerie satellitaire, en combinant l'étendue des surfaces en eau détectées sur des images radar ou optiques à l'altimétrie radar (Frappart et al., 2019; Papa et al., 2015) ou des MNT (voir introduction de l'article en section 5.3 pour une revue de ces méthodes). Dans les zones semi-arides et arides, l'eau stockée en surface est souvent négligée et la variation de TWS est attribuée à l'eau souterraine et à l'humidité du sol SMA et GWSA (Frappart et al., 2019; Papa et al., 2015; Sun et al., 2010; Yeh et al., 2006).

L'humidité du sol dans la couche superficielle (0-5 cm) et dans les couches plus profondes est difficile à évaluer. Certaines régions bénéficient d'un réseau d'observation dense comme l'ISWS *Illinois State Water Survey* (Yeh et al., 2006), mais de telles données sont rarement disponibles. L'humidité du sol dans la zone non saturée (SM) est donc généralement estimée avec RZSM (humidité jusqu'à 1 ou 2 m de sol) évaluée par des modèles : NLDAS (North American Land Data Assimilation System) (Breña-Naranjo et al., 2014; Sahour et al., 2020; Seyoum and Milewski, 2017; Sun et al., 2010, 2012) ou son équivalent global GLDAS (Global Land Data Assimilation System) (Ali et al., 2021; Chen et al., 2019; Jyolsna et al., 2021; Karunakalage et al., 2021; Rodell et al., 2007, 2009; Yin et al., 2018, 2020; Zhang et al., 2021b; Zuo et al., 2021), WGHM (*WaterGAP global hydrology model*) (Frappart et al., 2019), EALCO (*Ecological Assimilation of Land and Climate Observations*) (Zhong et al., 2021). Depuis 2009 et 2015 les missions SMOS et SMAP fournissent un produit global d'humidité du sol à 0-5 cm, et des estimations de l'humidité en zone racinaire (jusqu'à un mètre de profondeur).

Le GWSA est généralement validé par des données piézométriques converties en anomalies de GWS avec une valeur moyenne de  $S_y$  (Breña-Naranjo et al., 2014; Ferrant et al., 2014; Frappart et al., 2019; Miro and Famiglietti, 2018; Zaitchik et al., 2008). Une valeur moyenne de  $S_y$  permet d'évaluer un ordre de grandeur, car sa variabilité est forte.

### 3.4.1.2 Cas du Telangana

La contribution de l'eau de la canopée (CWSA) et la neige (SNA) est négligeable dans le TWS du Telangana. L'eau de surface est représentée par les rivières, les grands lacs de barrage et le système de récupération des eaux de pluies (RHS, *Rainwater Harvesting System*) composé de centaines de milliers de petits réservoirs. Les rivières ne sont pas pérennes et ne sont actives que quelques mois pendant et après la mousson, à l'exception des fleuves majeurs de la Godavari et la Krishna (Figure 1.2). Le NRLD (*National Registry of Large Dam*) estime la capacité maximale des principaux grands lacs de barrages référencés autour de 113 mm. La capacité maximale du RHS sur le Telangana a été estimée à 24 mm (voir chapitre 5). Ce réservoir potentiel de 137 mm peut représenter jusqu'à 22% de la fluctuation annuelle du TWS de GRACE. Cependant les retenues d'eau sont rarement simultanément pleines et leur remplissage est limité, et l'eau stockée en surface peut être négligée la plupart du temps. Aussi, seules les contributions de l'humidité du sol et de l'eau souterraine au signal GRACE sont considérées comme significatives, et l'équation 3.4 peut être simplifiée en :

$$\text{TWSA} = \text{GWSA} + \text{SMA} \quad (3.5)$$

### 3.4.2 Sensibilité du GWS estimé à différents jeux de données

La déconvolution du TWS est réalisée à l'échelle de la zone granitique du Telangana d'étendue similaire à la résolution de GRACE (113 000 km<sup>2</sup>) (zone rose sur la figure 1.2). À l'instar de la majorité des approches de déconvolution, SMA est modélisé avec l'anomalie de RZSM de quatre jeux de données : les sorties des modèles GLDAS-NOAH et GLEAM, et les produits L4 de SMOS et SMAP. La profondeur de la zone racinaire varie selon les modèles (2 m pour GLDAS-NOAH, 1 m pour SMOS et SMAP, non précisé pour GLEAM), mais pour cette application, le RZSM des quatre modèles est converti en hauteur d'eau équivalente pour une zone racinaire d'une profondeur d'1 m. La figure 3.3 représente la variation temporelle de l'humidité du sol des différents produits de RZSM. Tous les produits sont logiquement corrélés aux précipitations, à l'exception du produit L4 de SMOS qui montre une augmentation presque systématique de l'humidité du sol entre décembre et mars, alors même que cette période sèche est marquée par de très faibles précipitations. Cette augmentation de l'humidité, coïncidant avec le début de la saison culturale de la saison sèche, semble être causée par l'irrigation des cultures et particulièrement l'inondation des rizières (voir section 3.5.1).

Les estimations de GWSA sont comparées avec le GWSA dérivé des données *in situ* mensuelles (produites comme expliqué en section 3.2.3) avec le coefficient de corrélation de Pearson (R) et le coefficient de détermination (R<sup>2</sup>). Ces résultats sont comparés au cas "sans déconvolution" où le TWSA de GRACE est directement comparé au GWSA *in situ*. La figure 3.4 représente les séries temporelles de GWSA estimé par déconvolution, le TWSA de GRACE et le GWSA dérivé des données *in situ* ainsi que les nuages de points associés. La longueur variable des séries temporelles des RZSM rend difficile la comparaison des résultats, aussi les R et R<sup>2</sup> sont calculés sur des séries de même taille sur deux périodes : la période de 2010-2018 (intersection des périodes temporelles



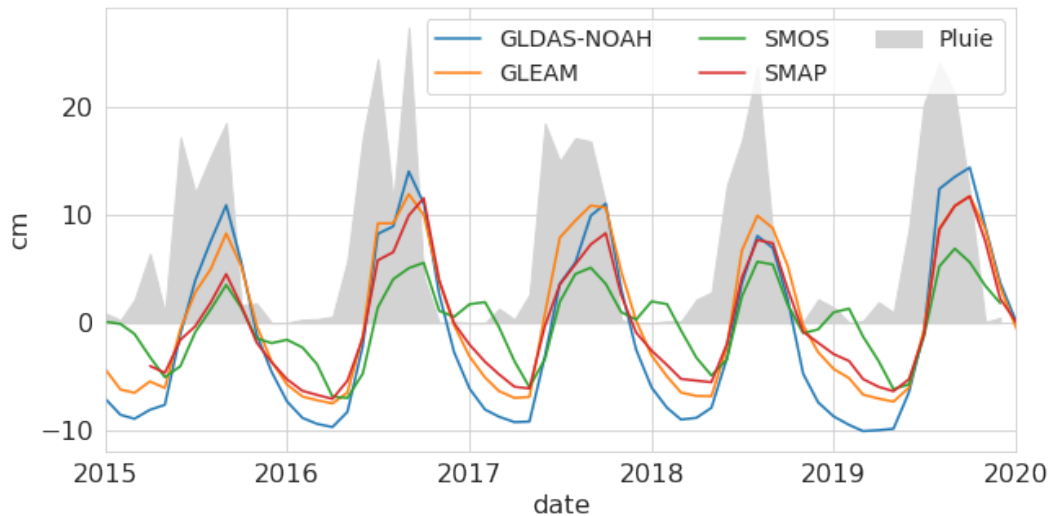


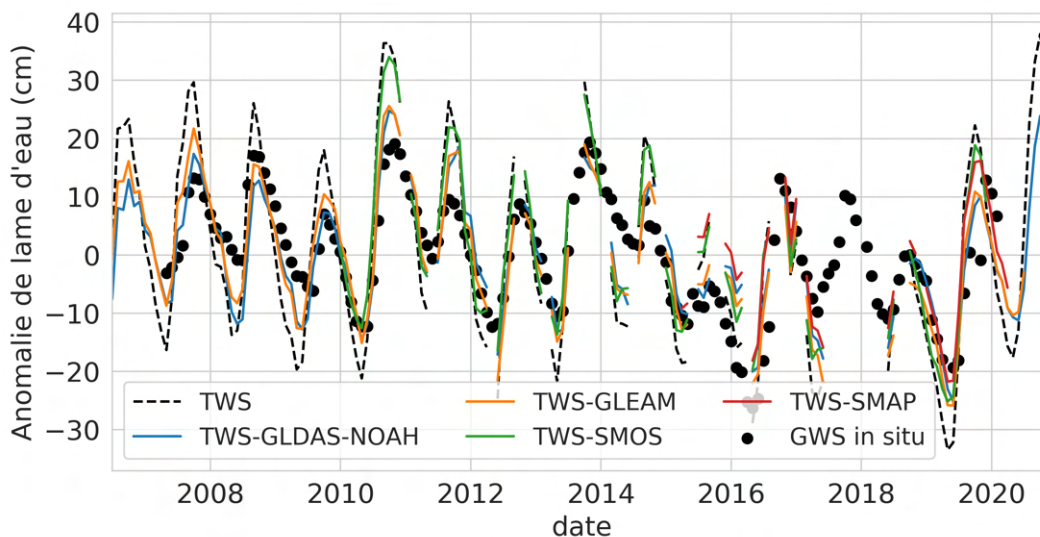
FIGURE 3.3 – Anomalies de RZSM (période 2015-2020) et précipitations.

TABLE 3.2 –  $R$  et  $R^2$  entre les données *in situ* et le GWSA estimé par déconvolution pour différents jeux de données. La dernière colonne compare directement le GWS *in situ* et TWSA de GRACE. Les métriques sont calculées pour différentes périodes : la série temporelle complète, la période 2010-2018 et la période 2015-2018.

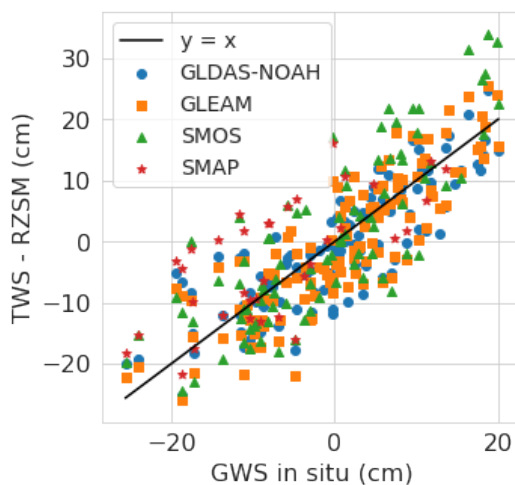
RZSM (nombre de mois)	GLDAS NOAH (122)		GLEAM (122)		SMOS (88)		SMAP (36)		- (122)	
	R	$R^2$	R	$R^2$	R	$R^2$	R	$R^2$	R	$R^2$
Période totale	0.84	0.67	0.85	0.64	0.82	0.44	0.69	0.20	0.79	0.00
2010-2018 (90 points)	0.87	0.70	0.86	0.66	0.82	0.44	-	-	0.81	0.14
2015-2018 (34 points)	0.75	0.37	0.71	0.27	0.65	0.05	0.69	0.20	0.66	-0.32

de GLDAS, GLEAM et SMOS) et 2015-2018 (couverte par l'ensemble des jeux de données). Les métriques sont reportés dans la table 3.2.

Sur ces deux périodes, retrancher le RZSM améliore grandement le  $R^2$ . Sans déconvolution, il est de 0.14 (respectivement  $-0.32$ ) pour la période 2010-2018 (respectivement 2015-2018), et varie de 0.44 à 0.70 (respectivement 0.05 à 0.37) avec déconvolution. Le  $R$  est également amélioré, excepté pour la déconvolution avec SMOS sur la période 2015-2018. Le RZSM de GLDAS (respectivement GLEAM) produit de meilleurs résultats que SMOS sur la période 2010-2018, avec un  $R$  de 0.87 (respectivement 0.86) contre 0.82, et  $R^2$  de 0.70 (respectivement 0.66) contre 0.44. La déconvolution avec SMOS est la moins bonne sur la période 2015-2018 avec un  $R$  de 0.65 et un  $R^2$  de 0.05. Le  $R^2$  plus bas peut être expliqué en partie par la calibration de GWS avec GLEAM. En effet, les trois jeux de données GLDAS, GLEAM et SMAP sont issus de modèles qui dépendent principalement des précipitations et présentent des amplitudes annuelles plus fortes que SMOS (voir figure 3.3), et la calibration de GWS avec GLEAM favorise les modèles similaires. Le Sy calibré avec SMOS est de 4.9% au lieu de 4.7%, ce qui augmente le  $R^2$  de SMOS à 0.50 sur la période 2010-2018 et à 0.11 sur la période 2015-2018, tout en conservant le même coefficient de corrélation.



(a)



(b)

FIGURE 3.4 – (a) GWSA évalué avec les quatre jeux de données RZSM (GLDAS-NOAH, GLEAM, SMOS SMAP) (lignes colorées), TWSA de GRACE (ligne noire pointillée) et GWSA dérivé des données *in situ* (points noirs). (b) Nuage de points du GWSA évalué avec les quatre jeux de données RZSM et de GWSA *in situ*.

### 3.4.3 Discussion

Dans la littérature, la comparaison des données *in situ* avec le GWSA évalué par des méthodes similaires donne des résultats variés selon la zone d'étude (voir table 3.3). Les corrélations sont généralement bonnes entre les données *in situ* et le GWSA évalué par déconvolution sur le Telangana, compte tenu des disparités entre les modes et fréquences d'acquisitions des différents jeux de données utilisés dans ce type d'approches (GRACE, modèle, *in situ*), et parfois meilleures que celles des études similaires. Dans le meilleur des cas, le GWS évalué par déconvolution explique 70% de la variance du GWS *in situ* sur la période 2010-2018.

Certaines sources d'incertitudes peuvent néanmoins dégrader la qualité des résultats. Notamment, les stocks d'eau de surface et de la canopée sont négligés dans l'équation de déconvolution,

TABLE 3.3 – Plages de valeurs de  $R$  et  $R^2$  entre le GWS *in situ* et issu de la déconvolution de GRACE dans la littérature.

Référence	Zone d'étude	$R$	$R^2$
(Singh and Saravanan, 2020)	Inde du Sud		0.52 – 0.65 (pré-mousson) 0.53 – 0.90 (post-mousson)
(Breña-Naranjo et al., 2014)	Aquifère des Hautes Plaines (États-Unis)		0 – 0.18
(Yeh et al., 2006)	Illinois (États-Unis)	0.66	
(Yin et al., 2020)	Tasmanie	0.64–0.88	

alors qu'ils pourraient avoir une contribution significative pendant la saison humide. Aussi, l'humidité du sol dans la zone non saturée (SMA) est modélisée ici par l'humidité dans le premier mètre de sol (RZSMA), négligeant ainsi les variations d'eau situées plus profondément jusqu'à la nappe souterraine. Cette approximation est partiellement compensée par la méthode de calibration de  $S_y$  (section 3.2.3) qui permet d'ajuster l'amplitude de GWSA en cohérence avec celle du RZSMA. La sous-estimation ou l'élimination de ces termes du bilan d'eau pourrait néanmoins être responsable de la valeur légèrement haute de  $S_y$  (4.7%). Enfin, la surface de la zone d'étude (113 000 km<sup>2</sup>) pourrait aussi être trop faible par rapport à la résolution de GRACE. En effet, Rodell et al. (2007) jugent insatisfaisant l'accord entre le GWS évalué par déconvolution du TWS de GRACE et les données *in situ* à l'échelle de 500 000 km<sup>2</sup> sur l'État du Mississippi (États-Unis), et suggèrent que la méthode est plus appropriée pour des échelles plus grandes (900 000 km<sup>2</sup>).

## 3.5 Potentiel de SMOS

### 3.5.1 Un cycle annuel d'humidité bimodal

Le cycle annuel bimodal du RZSM de SMOS mis en évidence en figure 3.3 est également visible sur l'humidité de surface (SSM, *Surface Soil Moisture*) de SMOS, SMAP et CCI dans une moindre mesure (figure 3.5), et contraste avec le cycle unimodal des SSM de GLDAS-NOAH et GLEAM. Le premier pic en saison humide, fortement corrélé aux précipitations est présent sur tous les jeux de données. Un autre pic d'humidité, de moindre amplitude que le pic en saison humide, survient pendant la saison sèche entre décembre et mars, ce qui correspond à une saison culturale avec irrigation. Son amplitude varie selon les années et l'intensité des précipitations. En particulier, le pic de la saison sèche au début de l'année 2016, aux précipitations particulièrement faibles, est moins marqué que les autres années. Ce second pic est absent du SSM modélisé par GLDAS-NOAH et GLEAM, deux modèles ne prenant pas en compte l'irrigation et dépendant principalement des précipitations.

Le cycle annuel bimodal mis en évidence sur le SSM se retrouve sur le RZSM de SMOS, mais est absent du RZSM de SMAP (figure 3.3), ce qui s'explique par leur mode de production. Le RZSM de SMOS est produit avec un modèle simulant l'infiltration de l'eau de la couche superficielle (SSM SMOS) dans les couches plus profondes, tandis que le RSZM de SMAP est une sortie du modèle GLDAS-CLSM avec assimilation de la température de brillance de SMAP. GLDAS-CLSM ne modélise pas l'irrigation, et de fait, lisse le pic en saison sèche détecté par le radiomètre du

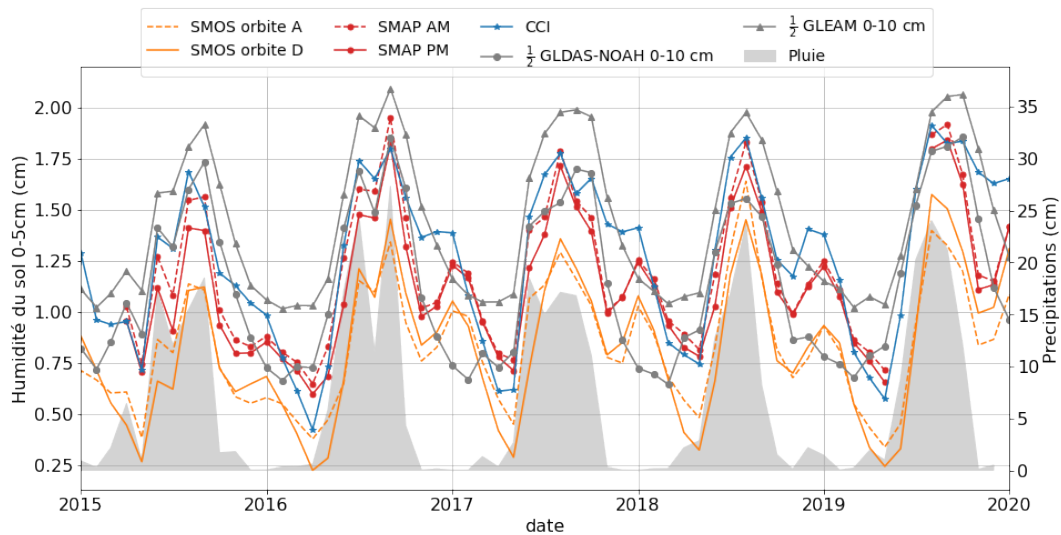


FIGURE 3.5 – Humidité de surface (0-5 cm) de SMOS, SMAP et CCI et des modèles GLDAS-NOAH et GLEAM (humidité 0-10 cm divisée par deux).

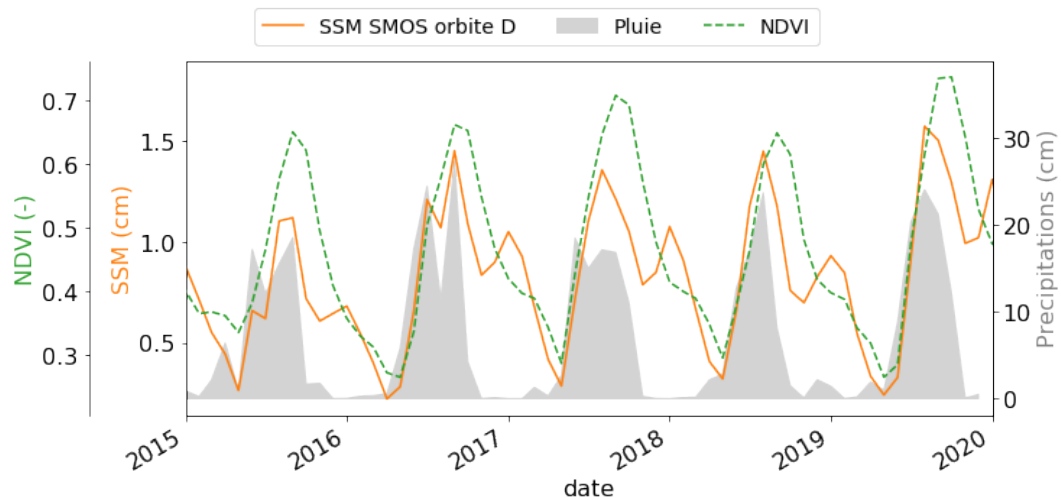
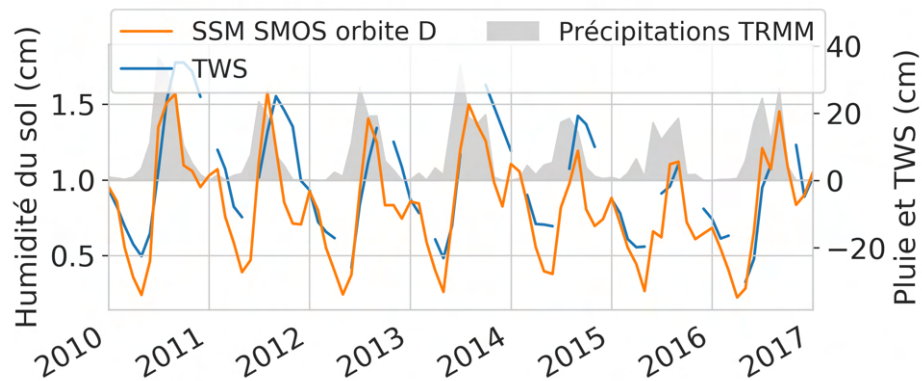


FIGURE 3.6 – Séries temporelles de NDVI, d'humidité de surface (SSM) de SMOS et de pluie sur la zone granitique du Telangana.

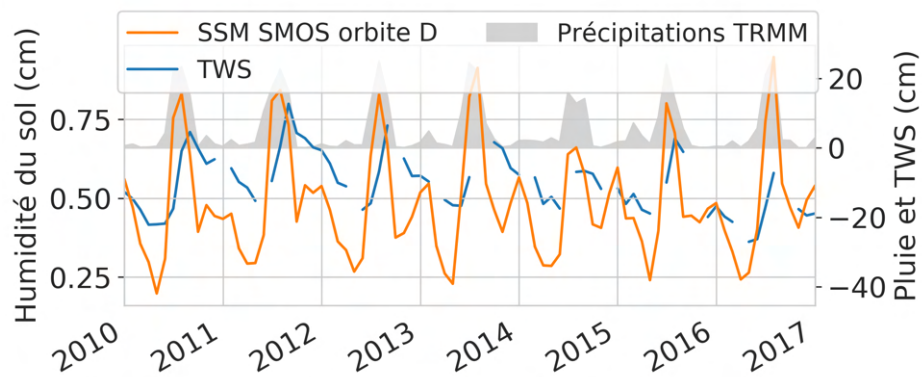
satellite qu'il ne peut pas interpréter comme un artefact anthropique.

La bimodalité du cycle annuel de SMOS a été recherchée dans différentes régions de l'Inde (dont les localisations sont indiquées en figure 3.2b) : l'État du Telangana, l'Inde du Nord et l'État du Maharashtra. Le pic de SMOS en saison sèche est présent en Inde du Nord, mais il est moins marqué dans le Maharashtra où on observe plutôt une rupture de pente (figure 3.7). Le Maharashtra est pourtant un État où plus de 50% des surfaces sont irriguées (Verma and Katpatal, 2020), mais les ressources souterraines utilisables pour l'irrigation sont moins importantes que dans le Telangana et en Inde du Nord (Bhanja and Mukherjee, 2019), induisant probablement la mise en place de cultures moins gourmandes en eau et moins détectables par imagerie satellitaire. La figure 3.2b met également en évidence le décalage temporel entre précipitations, SSM et TWS en saison humide. Le pic d'humidité, induit par les précipitations et l'irrigation des cultures, est atteint au milieu de la mousson généralement un mois après le pic de cumul de précipitations. Le pic de TWS est

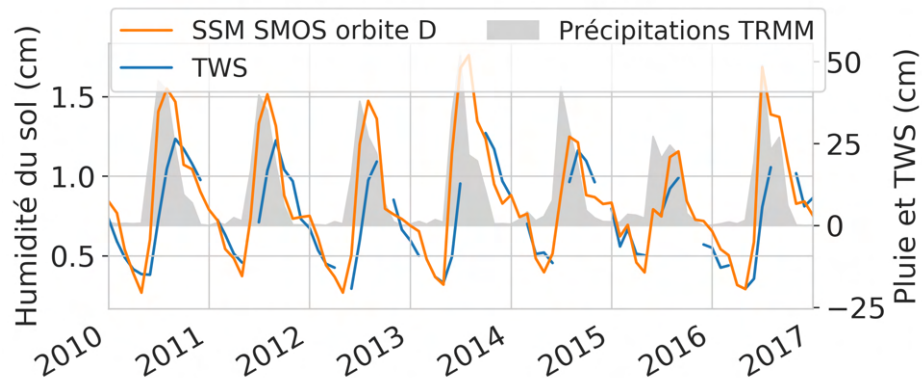
atteint le mois suivant, témoignant de la durée d'infiltration de l'eau de pluie dans l'aquifère.



(a) Telangana (granite)



(b) Inde du nord



(c) Maharashtra

FIGURE 3.7 – Humidité de surface de SMOS sur différentes régions d'Inde. La localisation des trois régions est indiquée en figure 3.2b.

### 3.5.2 Interprétation à l'échelle de SMOS

La présence d'un pic d'humidité pendant la saison sèche sur les données satellitaires et son absence sur les sorties de modèle semble indiquer une origine anthropique : l'irrigation des sols pendant la deuxième saison culturale. Cette partie vise à évaluer cette hypothèse en reliant le

SSM de SMOS à d'autres variables permettant d'identifier l'activité agricole : le NDVI de MODIS (indice de végétation), les cartes d'occupation du sol produites avec IOTA2, et les données piézométriques *in situ* (indiquant le degré d'exploitation de l'aquifère). La profondeur de la nappe GWL (*GroundWater Level*) est directement utilisé dans cette étude, contrairement aux approches de déconvolution qui nécessitent sa conversion en GWS pour la comparaison quantitative de différents stocks. Les sens de variations de GWS et GWL sont opposés : le GWL diminue avec le remplissage de la nappe. Pour plus de précision et de finesse dans l'interprétation, cette étude est réalisée à l'échelle de la grille SMOS (25 km) et non à l'échelle de la zone granitique du Telangana (113 000 km<sup>2</sup>). La zone d'étude est délimitée par l'emprise des cartes d'occupation du sol IOTA2 (voir figure 3.8). Les séries temporelles des différentes variables étudiées sont moyennées dans chaque pixel de 25×25 km<sup>2</sup>, dont deux exemples sont présentés en figure 3.9.

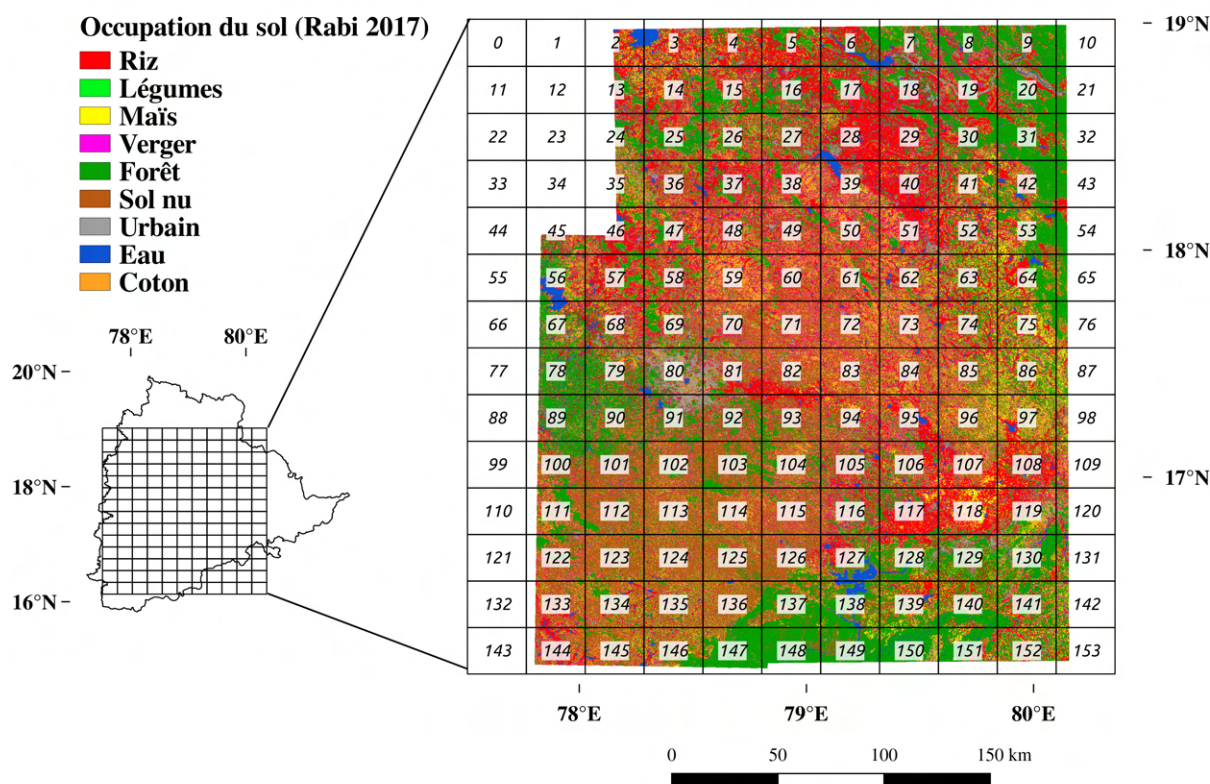
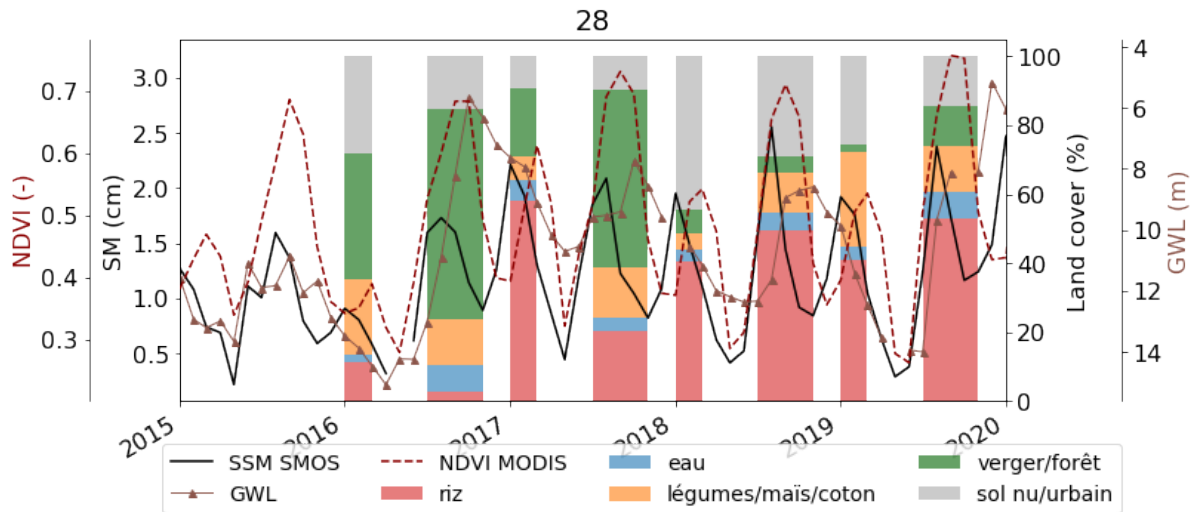
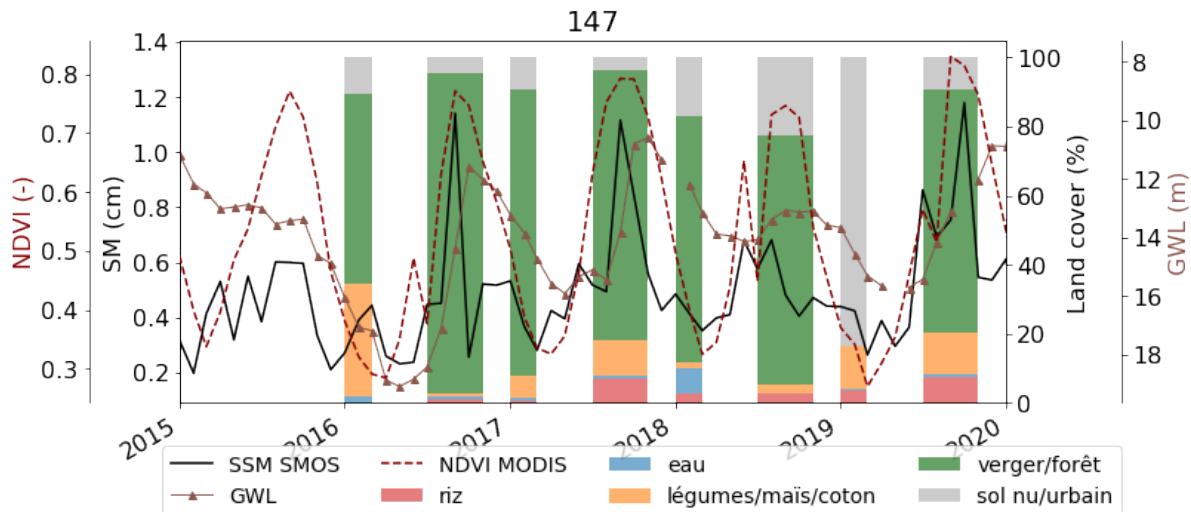


FIGURE 3.8 – Grille SMOS (25 km de résolution) et occupation du sol (IOTA2) pour la saison sèche (Rabi) 2017.

**Étude qualitative.** Dans les zones agricoles fortement irriguées en période sèche comme en figure 3.9a (plus de 40% de riz en période sèche entre 2017 et 2020), les séries temporelles de SSM SMOS et NDVI présentent un cycle annuel bimodal avec un pic en saison sèche. Le décalage des cycles est cohérent : le pic en période sèche du NDVI (attribué à la croissance des cultures) survient deux mois après celui de SSM SMOS (attribué à l'irrigation). Sur des zones moins agricoles comme en figure 3.9b (plus de 50% de forêts), ni le SSM de SMOS ni le NDVI ne présentent de pic marqué en saison sèche, et le pic de NDVI pendant la saison humide est plus étalé dans le temps par rapport à une zone agricole, témoignant de la pérennité des forêts par rapport aux cultures. Sur les deux pixels, l'amplitude des pics de SSM SMOS semble être corrélée à la quantité d'eau souterraine de



(a) Zone cultivée (riz)



(b) Forêt

FIGURE 3.9 – Séries temporelles de NDVI, SSM SMOS, GWL et occupation du sol sur deux pixels de  $25 \times 25 \text{ km}^2$ . Les numéros des pixels correspondent à ceux indiqués sur la figure 3.8. Axe 1 : NDVI (rouge pointillé). Axe 2 : SSM de SMOS (noir) en cm (moyenne des deux orbites). Axe 3 : occupation du sol (%). Axe 4 : Niveaux piézométriques mesurés (marron) en m. Les saisons sèches sont indiquées par des bandes d'occupation du sol moins larges qu'en saison humide.

la même année hydrologique, qui correspond à la succession d'une saison humide et d'une saison sèche. À l'échelle du Telangana, la série temporelle de SMOS présente également un cycle annuel bimodal très marqué contrairement au NDVI (figure 3.6). En effet, la diversité des occupations du sol observable à fine échelle ou des différentes pratiques culturales comme les dates de semis n'est pas correctement restituée par une moyenne spatiale à large échelle.

**Étude quantitative.** Cette partie vise à identifier des corrélations entre l'amplitude du pic de SSM SMOS en saison sèche et d'autres variables reflétant les activités agricoles. Ici, ce pic est caractérisé chaque année par son maximum entre décembre et avril. Les coefficients de corrélation de Pearson ( $R$ ) sont calculés entre ce maximum et l'occupation du sol pendant la saison sèche, le

pic de NDVI pendant la saison sèche et le GWL. Le GWL est caractérisé par trois indicateurs : le minimum en saison humide (GWLK), le maximum en saison sèche (GWLR) et GWLR-GWLK, qui correspondent respectivement au stock maximum d'eau souterraine à la fin de la mousson, au stock minimum après la saison sèche, et à la quantité d'eau utilisée pendant l'année hydrologique. Les corrélations et p-valeurs sur toutes les années et tous les pixels de  $25 \times 25$  km<sup>2</sup> sont reportées en table 3.4. L'intensité du pic de SSM SMOS est fortement corrélée avec la présence de riz ( $R = 0.81$ , nuage de points en figure 3.10a) pour toutes les cases de  $25 \times 25$  km<sup>2</sup> confondues, et cette tendance est également visible à l'échelle des cases individuelles pour les quatre années où les cartes sont disponibles (résultats non présentés). Les corrélations avec les autres types occupations du sol sont moins marquées ( $R$  faibles, p-valeurs non significatives ou  $> 0.01$ ) et sont difficilement interprétables car ces couverts sont mal identifiés par l'algorithme de classification (figure 2.3). La corrélation entre NDVI et SMOS est moins claire ( $R = 0.65$  et nuage de points diffus en figure 3.10b) car calculée sur tous les pixels quel que soit leur degré d'anthropisation. GWLR et GWLK sont corrélés négativement avec l'amplitude du pic de SSM en saison sèche ( $R = -0.54$ , figure 3.10c), indiquant que pendant les années très humides, les ressources élevées en début d'année (faible GWLK) permettent une irrigation intense des cultures en saison sèche (fort pic de SSM SMOS) sans trop impacter l'aquifère (faible GWLR).

TABLE 3.4 – Corrélations entre l'amplitude du pic d'humidité de SMOS en saison sèche (moyenne des deux orbites), l'occupation du sol, le pic de NDVI en saison sèche et les données piézométriques. (coefficient de corrélation / p-value). Les corrélations en italique sont non significatives (avec une p-valeur supérieure à 0.05).

		R	p-valeur
NDVI	Amplitude du pic en saison sèche	0.65	1.2e-187
Occupation du sol	Riz	0.81	1.6e-101
	Eau	0.12	0.014
	Légumes / Maïs / Coton	-0.11	0.024
	Vergers / Forêt	<i>-0.055</i>	<i>0.25</i>
	Sol nu / Urbain	-0.36	1.8e-14
	GWL	GWLR	-0.54
	GWLK	-0.54	5.3e-123
	GWLR-GWLK	-0.19	3.3e-15

**Conclusion.** Le pic d'humidité en saison sèche observé sur SSM SMOS à l'échelle de  $25 \times 25$  km<sup>2</sup> et du Telangana ( $113\,000$  km<sup>2</sup>) semble être causé par l'irrigation des cultures pendant la saison sèche, en particulier l'inondation des rizières. Cette hypothèse est confirmée par (i) la présence du pic d'humidité en saison sèche uniquement sur les régions agricoles fortement irriguées, (ii) son absence sur les séries de SSM modélisées sans prise en compte de l'irrigation, (iii) la cohérence temporelle des pics en saison sèche de SSM SMOS et de NDVI, et (iv) la corrélation entre l'amplitude du pic de SSM SMOS et la présence de riz.

### 3.5.3 Quantification de l'eau utilisée pour l'irrigation avec SMOS

Dans l'État du Telangana, le pic de SSM en saison sèche, attribué à l'irrigation, est décorrélié des précipitations qui sont presque inexistantes durant cette période. L'influence quasi-totale de



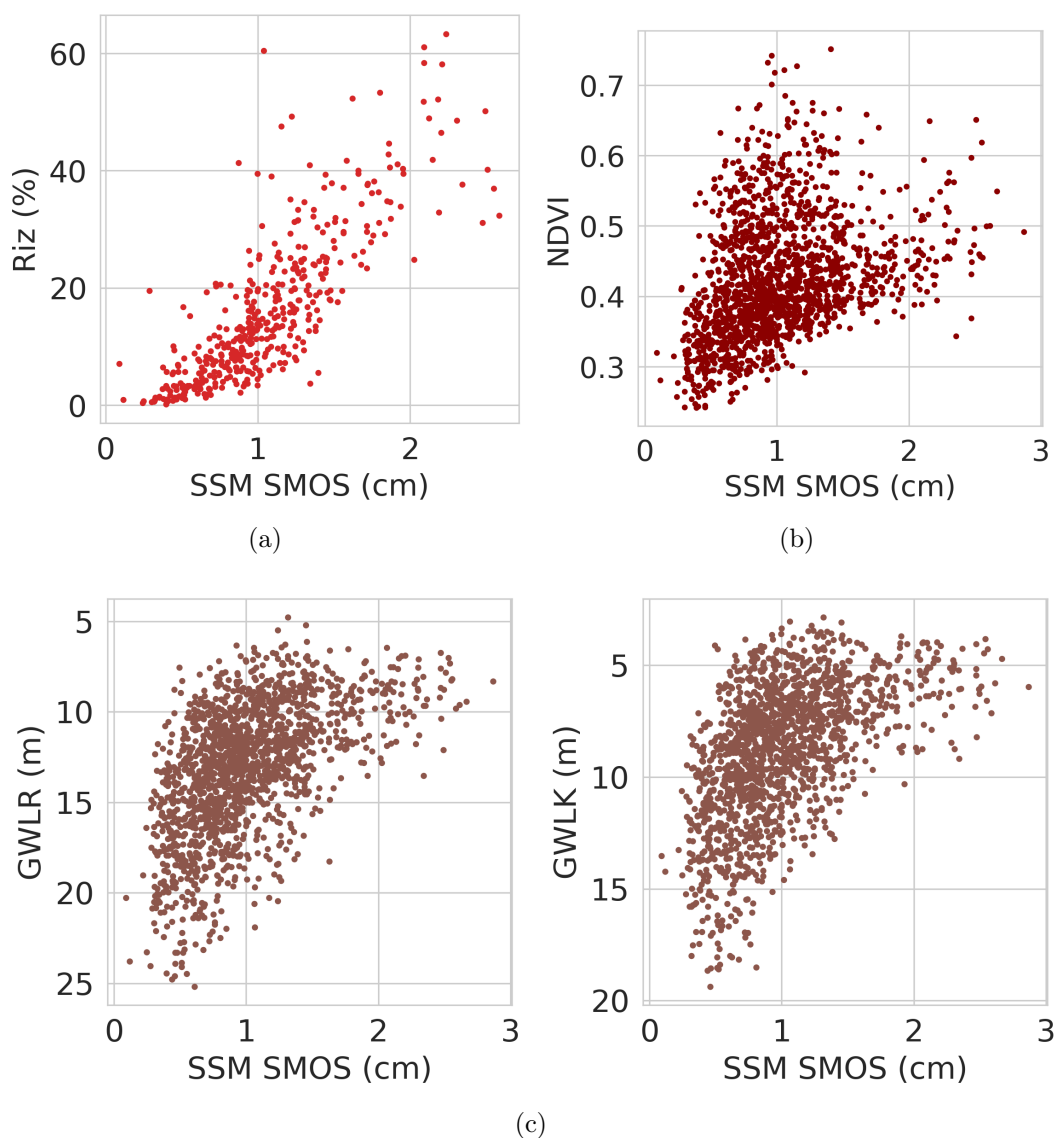


FIGURE 3.10 – (a) Proportion de riz en Rabi, (b) amplitude du pic de NDVI en saison sèche et (c) GWLR et GWLK en fonction du pic de SSM SMOS en saison sèche.

l'irrigation sur le pic de SSM SMOS en saison sèche a motivé l'étude de son potentiel pour estimer la quantité d'eau utilisée pour l'irrigation, dont les résultats préliminaires sont présentés dans cette section.

Plusieurs algorithmes utilisent les données d'humidité pour quantifier ou rectifier les précipitations : SM2RAIN (Brocca et al., 2014), PrISM (Precipitation Inferred from Soil Moisture) (Pellarin et al., 2020), Soil Moisture Analysis Rainfall Tool (SMART) (Crow et al., 2011, 2009) ou API-mod (Pellarin et al., 2013). En particulier, la méthode SM2RAIN a été appliquée pour quantifier l'irrigation à partir des données d'humidité Brocca et al. (2018). L'algorithme SM2RAIN est basé sur l'inversion du bilan hydrique du sol :

$$Z^* \frac{dS(t)}{dt} = r(t) + i(t) - g(t) - sr(t) - e(t) \quad (3.6)$$

avec  $Z^*$  (mm) la capacité de l'eau utile du sol,  $S(t)$  (-) la saturation relative du sol,  $t$  (jour) le

temps,  $r(t)$  (mm/jours) les précipitations,  $i(t)$  (mm/jour) l'irrigation,  $g(t)$  (mm/jour) le drainage,  $sr(t)$  (mm/jour) le ruissellement et  $e(t)$  (mm/jour) l'évapotranspiration réelle. En négligeant le ruissellement  $sr(t)$ , le taux d'irrigation peut être calculé comme suit :

$$i(t) = Z^* \frac{dS(t)}{dt} + g(t) + e(t) - r(t) \quad (3.7)$$

$$= Z^* \frac{dS(t)}{dt} + K_S S(t)^a + ET_{pot} S(t) - r(t) \quad (3.8)$$

$$(3.9)$$

avec  $K_S$  (mm/jour) la conductivité hydraulique du sol à saturation,  $a$  (-) un paramètre traduisant la non-linéarité entre le taux de drainage et la saturation du sol, et  $ET_{pot}$  l'évapotranspiration potentielle.

La calibration des paramètres ( $Z^*$ ,  $K_S$ ,  $a$ ) se fait lors des jours de pluie ( $r(t) > 0$ ) en supposant que l'irrigation est nulle ( $i(t) = 0$ ). Une fois calibrée, l'équation 3.7 est appliquée pour estimer l'irrigation à partir des données de pluie, d'humidité et d'évapotranspiration. Cependant, la méthode peine à séparer les contributions de la pluie et de l'irrigation sur les régions où l'on irrigue pendant les mois où les précipitations sont fortes, comme c'est le cas dans le Telangana. De fait, l'application de cette méthode sur cette région avec les données de pluie TRMM, l'évapotranspiration potentielle PET (Shen et al., 2021), et l'humidité de surface de SMOS à l'échelle de tuiles de 0.5 degrés donne des valeurs de ( $Z^*$ ,  $K_S$ ,  $a$ ) parfois incohérentes lors de la calibration (valeurs négatives par exemple). De plus, l'irrigation estimée à partir de ce type de méthode est difficile à valider du fait de l'indisponibilité de données d'irrigation sur la zone d'étude.

### 3.5.4 Indicator of flood-irrigated crops from SMOS and SMAP soil moisture products in southern India

Claire Pascal<sup>1</sup>, Sylvain Ferrant<sup>1</sup>, Nemesio Rodriguez-Fernandez<sup>1</sup>, Yann Kerr<sup>1</sup>, Adrien Selles<sup>2</sup>, Olivier Merlin<sup>1</sup>.

Soumis à *IEEE Geoscience and Remote Sensing Letters* (GRSL).

---

1. Centre d'Étude Spatiale de la BIOSphère (CESBIO-UPS-CNRS-IRD-CNES-INRAE), 18 av. Ed. Belin, Toulouse CEDEX 9, 31401, France

2. Bureau de Recherches Géologiques et Minières (BRGM), Université de Montpellier, 1039 rue de Pinville, Montpellier, 34000, France

# Indicator of flood-irrigated crops from SMOS and SMAP soil moisture products in southern India

Claire Pascal<sup>1</sup>, Sylvain Ferrant<sup>1</sup>, Nemesio Rodriguez-Fernandez<sup>1</sup>, Yann Kerr<sup>1</sup>, Adrien Selles<sup>2</sup>, Olivier Merlin<sup>1</sup>  
<sup>1</sup>Centre d'Étude Spatiale de la BIOSphère (CESBIO-UPS-CNRS-IRD-CNES-INRAE), 18 av. Ed. Belin, Toulouse  
 CEDEX 9, 31401, France

<sup>2</sup>Bureau de Recherches Géologiques et Minières (BRGM), Université de Montpellier, 1039 rue de Pinville,  
 Montpellier, 34000, France

**Abstract**—Spaceborne L-band data have the potential to monitor flooded and irrigated areas. However, so far little has been done to quantitatively assess the impact of flood-irrigated crops on SMOS and SMAP soil moisture data. This paper demonstrates the ability of SMOS and SMAP surface soil moisture (SSM) retrievals to quantify the fraction of flood-irrigated area at the seasonal scale and at a 25 km resolution in the Telangana State in southern India. Over irrigated areas, both SMOS level 3 (L3) SSM and SMAP L3 enhanced SSM products present a bimodal annual cycle, with a peak of SSM during the monsoon (wet) season corresponding to rainfall and irrigation, and a peak during the dry season due to irrigation activities solely. The second peak is absent or has a very small amplitude in areas where rice represents a small fraction (typically below 5-10%). More importantly, the amplitude of the second SSM peak is significantly correlated to the rice cover fraction within 25×25 km<sup>2</sup> pixels (R=0.81 for SMOS and 0.77 for SMAP), showing its potential to assess crop fraction and hence the water used for irrigation. The SMOS/SMAP L3 SSM peak during the dry period occurs several months before the harvest, constituting an indicator for rice stocks at the end of the season. However the irrigation signature is absent from the SMAP level 4 SSM product derived from the assimilation of SMAP brightness temperatures in a land surface model, which indicates that the data assimilation scheme is inefficient to reconstitute irrigation information.

**Index Terms**—Soil moisture, SMOS, SMAP, Irrigation, Water Management, Rice Cropping, Regional Scale.

## I. INTRODUCTION

With the increase of agricultural production during the last century, irrigation has become the first consumer of freshwater, representing 70% of global freshwater withdrawals [1]. The increase of soil moisture by irrigation practices also alters the water cycle [2]. Knowledge of irrigated land is hence crucial to manage the water resource and to ensure food security. It is also important to understand the impact of irrigation on the water cycle and climatology, and to participate to the development of hydrological models that account for irrigation in a realistic manner.

The dynamic mapping of inundated areas can be achieved with microwave radiometers thanks to the high sensitivity of the microwave brightness temperature (T<sub>b</sub>) to open water. Such data are insensitive to clouds and hence allow continuous monitoring, in spite of having a lower spatial resolution than optical data. The T<sub>b</sub> at various frequencies has been used to detect flooded pixels [3]–[6] or to estimate an open water fraction within the microwave pixel by separating the

contributions of land and water. Especially, T<sub>b</sub> data from the multi-band Advanced Microwave Scanning Radiometer for EOS (AMSR-E) sensor [7], [8], the L-band Soil Moisture and Ocean Salinity (SMOS) mission [9], and the L-band Soil Moisture Active Passive Data (SMAP) mission [5] have been used at the 5 km and ≈40 km resolution for AMSR-E (89 GHz) and SMOS/SMAP, respectively. Note that higher resolutions can be achieved by downscaling the microwave data using disaggregation algorithms and ancillary data [5], [10].

The particular sensitivity of the L-band (1.4 GHz) to water has been used to monitor the soil water content. Since 2010 and 2015 respectively, SMOS and SMAP satellite missions have provided surface soil moisture (SSM) estimates every 3 days at a ≈40 km resolution. L-band-derived SSM data have been proven to be sensitive to irrigation, in particular in heavily and flood irrigated areas [11], [12], demonstrating their potential for irrigated area mapping and irrigation timing detection. Notably, several indicators have been identified to detect irrigation from the amplitude of SSM values in time series [11] and using the spatial relative difference within an area revealing the spatial distribution of irrigated sites [13]. Also, irrigated areas can be detected by comparing the SSM retrievals with the SSM simulated by a land surface model that does not take irrigation into account: larger observed SSM values are attributed to irrigation [12]–[14]. Automatic discrimination of irrigated areas from other land types has also been achieved from SSM data with unsupervised classification algorithms [13].

As a step further, L-band-derived SSM data have also been used to estimate the quantity of water used for irrigation [15]–[18]. Especially, the SM2RAIN algorithm [19] is based on the inversion of the soil water balance equation. This requires prior calibration during rainy periods, by assuming that there is no irrigation during those periods. Such an approach has been effective in quantifying irrigation water use over large irrigated areas and where irrigation and precipitation periods are distinct. However, SM2RAIN is not able to separate the contributions of rainfall and irrigation to the observed SSM dynamics [15]. Moreover, its application to smaller and sparse irrigated areas would require disaggregated SSM data, which may not be available at the required spatio-temporal scales with a sufficient accuracy [20].

To this day, therefore, studies have mainly focused on the

binary discrimination of irrigated and non-irrigated pixels from the SSM signals. However the microwave pixels (several tens of km) are generally much larger than the size of agricultural fields or even often the size of irrigation districts, and hence the above methods do not provide sub-pixel information about the portion of irrigated areas. To the knowledge of the authors, no demonstration has yet proved that L-band-derived SSM can be used to quantify a fraction of irrigated areas within SMOS and SMAP pixels, in the same way that open fraction water can be estimated from the Tb data. In this context, this paper aims to quantify the sensitivity of SMOS and SMAP SSM data to the percentage of irrigated rice within 25 km resolution (SMOS sampling grid) pixels. The study is undertaken over the Telangana State in southern India and for four years when field data are available (2016 to 2019 included). The reference data of rice fraction at the 25 km resolution are derived from an existing supervised classification algorithm [21], [22] based on Sentinel-2 data and previously evaluated using ground measurements in the area.

## II. STUDY AREA AND DATA

### A. Study area

The study focuses on the Telangana State in southern India (see Figure 1). Highly rural and irrigated (15% of the State area [23]), Telangana is dominated by a monsoon climate alternating with a wet and dry season. Monsoon rainfall occurs between July and October and ranges between 540 and 1300 mm a year with an average of 879 mm (Indian Meteorological Department, <https://mausam.imd.gov.in/>). Annual precipitation sustains generally two growing seasons of irrigated crops in a year, one during the wet season (from July to November) and one during the dry season (from January to April). Orchards, vegetables and maize are irrigated, but rice is the most water-consuming crop. Rice is flooded from early July until early September and from early January until early March in the wet and dry season respectively.

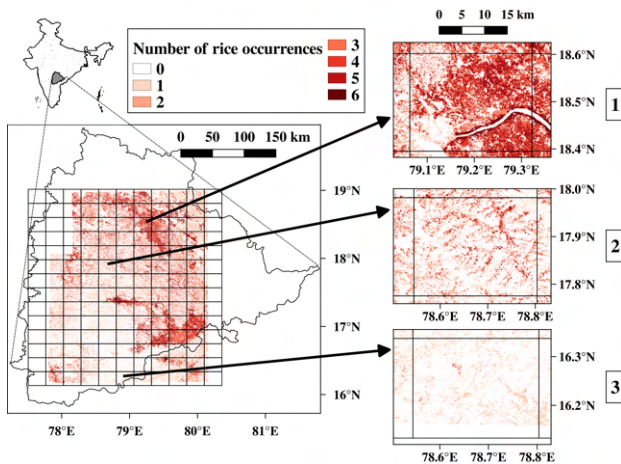


Fig. 1. Overview of the study area within the Telangana state in India overlaid with the 25 km resolution SMOS sampling grid and the number of rice occurrences during the 6 seasons when land cover maps are available.

### B. Data

This section presents the reference rice cover maps, SMOS/SMAP SSM data and other ancillary (vegetation index and rainfall) data used to interpret the results. All data were resampled (by simple averaging) on the same 25 km resolution SMOS grid.

1) *Land cover maps*: Reference land cover maps were produced from Sentinel-2 reflectances for wet and dry seasons from 2016 to 2019 separately on 8 Sentinel-2 tiles of  $100 \times 100$  km<sup>2</sup> each at 10 m resolution. Sentinel-2 data were processed by the MAJA processing chain [24] which detects clouds, cloud and topographic shadows, and corrects reflectance levels for atmospheric phenomena (absorption by atmospheric gases and scattering by air molecules and aerosols). These images were downloaded from Theia data platform (<https://www.theia-land.fr/>). Rice cover maps are produced with the IOTA2 algorithm (Infrastructure pour l'Occupation des sols par Traitement Automatique Incorporant les Orfeo Toolbox Applications, <https://framagit.org/iota2-project/iota2>), that achieves a supervised multitemporal classification with a random forest algorithm on the time series of Sentinel-2 images interpolated at 10 days [22]. The land cover classifications in wet and dry seasons are made on images acquired between July and November and between January and April, respectively.

The samples used for learning and evaluation of the classification algorithm are field data acquired on several consecutive days during each season of each year. Field data consist in visual identification of homogenous land cover and their delimitation as polygons using a GPS to locate them [21]. The performance of the rice cropping classification is evaluated with the F-score. The F-score is the harmonic mean of (i) the recall, i.e. the ability of the algorithm to identify a rice pixel, and (ii) the precision that measures the purity of the rice class. The rice class is generally deemed reliable with a F-score ranging from 75% to 92%. The interannual variability of the performances is linked to (i) the number of samples that varies between the years (262 to 1669) and (ii) the number of available Sentinel-2 images, which can significantly change during the wet season due to cloud coverage. In particular, the maps produced for the wet seasons of 2016 and 2017 are not used in this study because of their poor performance (F-score below 20%). Figure 1 shows the number of rice occurrences at the 10 m resolution during the 6 seasons when land cover maps are available.

The seasonal rice cover fraction is then estimated within each 25 km resolution SMOS grid pixel by averaging all 10 m pixels.

2) *SMOS*: The SMOS mission was launched in November 2009 and has a revisit time of 3 days at the Equator. The SMOS satellite carries an interferometric radiometer working at L-band. The data used for this study are the monthly level 3 (L3) SSM (L3SM RE07) in the first 5 centimeters of soil, provided by the Centre Aval de Traitement des Données SMOS (CATDS, <https://www.catds.fr/>) [25]. The data have a  $\simeq 40$  km resolution on average but are distributed on the 25 km resolution EASE2 grid [26] for both ascending and descending orbits. In this study we use the monthly average SSM of both orbits.

3) *SMAP*: The SMAP satellite, launched in January 2015, carries an L-band passive microwave radiometer. It provides SSM estimates at  $\approx 40$  km resolution with a revisit time of 3 days at the Equator. The data used for this study are the enhanced L3 dataset (L3E) provided on a 9 km resolution grid and obtained by exploiting the oversampling of the antenna overpasses. In this study we use the average SSM of both orbits. We also use the SMAP level 4 (L4) SSM data which are obtained by assimilating the SMAP brightness temperature in the NASA Catchment Land surface Model (CLM) using an ensemble Kalman filter [27]. The model describes the vertical transfer of soil moisture between the surface and root zone reservoirs, and provides fields of SSM and root zone soil moisture globally at 9 km resolution every 3 hours. The model is driven with observation-based surface meteorological forcing data, including precipitation, but does not account for irrigation activities. We use both L3 and L4 SSM data aggregated at the monthly scale and averaged within the 25 km pixels of the SMOS grid.

4) *MODIS NDVI*: The monthly NDVI (Normalized Difference Vegetation Index) allows monitoring the state of vegetation at a temporal scale finer than that of the seasonal land cover maps. We use the monthly NDVI of MODIS (Moderate-Resolution Imaging Spectroradiometer) sensor aboard Terra and Aqua satellites (launched in 1999 and 2002, respectively), at 1 km resolution (MOD13A3v006) and averaged within the 25 km pixels of the SMOS grid.

5) *TRMM rainfall*: The TRMM mission was launched in 1997 to quantify rainfall with several instruments including a precipitation radar, a microwave imager and a visible infrared scanner. We use monthly rainfall rates, provided at a resolution of  $0.25^\circ$  and resampled (bilinear resampling) on the 25 km resolution SMOS grid.

### III. RESULTS AND DISCUSSION

#### A. SSM dynamics over contrasted areas

The sensitivity of L-band sensors to irrigation is illustrated by the comparison of SMOS/SMAP SSM time series on three contrasted  $25 \times 25$  km<sup>2</sup> pixels, numbered from 1 to 3 (see Figure 1): 1) a heavily irrigated one (from 10 to 60% of rice cover) supplied with water by the irrigation systems downstream of large dams, 2) an upstream irrigated area where rice sowing is limited by the lack of water availability, and 3) a pixel with a dominant forest cover. Figure 2 shows the time series of 25 km resolution SMOS L3, SMAP L3E and SMAP L4 SSM monthly anomalies (obtained by subtracting the temporal average of SSM values between 2016 and 2019 pixel-wise) averaged within each  $25 \times 25$  km<sup>2</sup> pixel. The percentage of rice cover is shown for wet (large bands) and dry (thin hatched bands) seasons. The monthly NDVI indicates the crop/vegetation growth at higher temporal resolution.

In heavily irrigated areas (as in pixel 1), the SSM retrieved from SMOS and SMAP sensors has a bimodal annual cycle. During the wet season, the first peak of SSM is correlated with the precipitation, and the contribution of irrigation to SSM cannot be easily disentangled. During the dry season, SMOS and SMAP SSM show a systematic augmentation in the dry

season, between December and March. The maximum SSM value of the second peak is reached generally in January or February, when rice crops are flooded. The absence of precipitation coinciding with this peak suggests an anthropogenic source. The peak of NDVI two months later, associated with crop growth, strengthens this interpretation.

In upstream irrigated areas (as in pixel 2), the rice cover is lower and rarely exceeds 20% because of the lack of water availability, particularly during the dry season. SSM still has a bimodal cycle, but with lower amplitude for the SSM peaks in the dry season (SSM anomaly of  $0.05$  m<sup>3</sup>/m<sup>3</sup> at most), while the NDVI presents a single annual peak.

By contrast, the SSM of non-irrigated areas (as in pixel 3) presents a single peak in the year during the wet season which has its origins in monsoon rainfall. The different behaviors of SSM dynamics over irrigated and not irrigated areas confirm the findings of [11] in the western United States.

Contrary to SMOS and SMAP L3 SSM, SMAP L4 SSM does not show a peak in the dry season in irrigated areas and seems correlated solely to rainfall. This corroborates the limitations of assimilation approaches in representing unmodeled processes [14]. Two reasons can explain this result: first, an incorrect forcing of the model that integrates the contribution of rainfall but not irrigation to soil moisture, and second, the high dependency of assimilation results to the weights and uncertainties associated with observations and models.

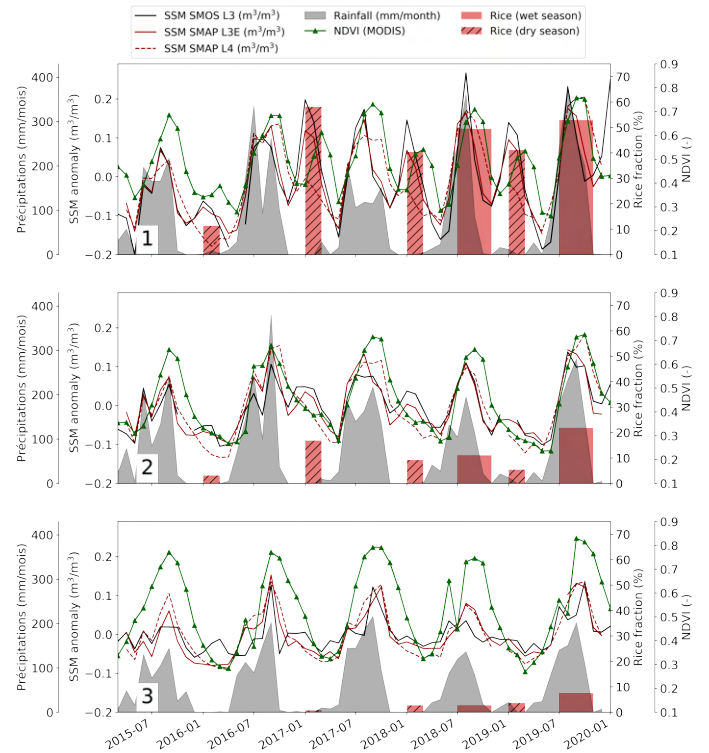


Fig. 2. Time series of monthly anomalies of SMOS and SMAP SSM (obtained by subtracting the temporal average of SSM values between 2016 and 2019 pixel-wise), TRMM rainfall, MODIS NDVI, and seasonal land cover fraction (%) on three pixels of  $25 \times 25$  km<sup>2</sup>. First axis: TRMM rainfall (mm/month). Second axis: SMOS L3 (black line), SMAP L3 (solid dark red line) and SMAP L4 (dotted dark red line) SSM. Third axis: Land cover fraction (%) for wet (large bands) and dry (thin hatched bands) season. Fourth axis: MODIS NDVI (solid green line).

TABLE I  
PEARSON CORRELATION COEFFICIENTS BETWEEN RICE COVER (%) AND SSM FOR EVERY SEASON OF EACH YEAR OR FOR ALL YEARS (“ALL”). ALL SIGNIFICANT CORRELATIONS ARE ASSOCIATED WITH A P-VALUE BELOW 0.05.

Season	Year	SMOS L3		SMAP L3E		SMAP L4	
		R	p-value	R	p-value	R	p-value
Dry	2016	0.67	3.7e-19	0.58	1.7e-13	0.05	0.6
	2017	0.89	8.6e-47	0.77	2.4e-26	0.55	3.9e-12
	2018	0.86	1.5e-41	0.67	6.5e-18	0.36	1.5e-05
	2019	0.62	1.2e-15	0.71	3.8e-22	0.56	1.0e-12
	All	0.81	1.3e-107	0.77	1.6e-96	0.55	8.1e-18
Wet	2018	0.76	4.8e-27	0.63	2.1e-16	0.44	6.9e-08
	2019	0.65	9.5e-18	0.56	1.1e-12	0.24	4.2e-3
	All	0.68	4.8e-38	0.55	7.1e-23	0.34	1.0e-08

### B. Spatio-temporal analysis

At the scale of each 25 km resolution pixel, the amplitude of the SSM peak is generally correlated in time with the corresponding proportion of rice cultivated within the area. This is illustrated in Figure 2 on pixel 1, in which a small proportion of rice is sown during the 2016 dry season (11% of rice) due to a lack of available water after a succession of dry years. This situation explains a low SSM peak amplitude (SSM anomaly of  $-0.05 \text{ m}^3/\text{m}^3$ ). On the other hand, the heavy monsoon rainfall in September 2016 has replenished surface and groundwater that allowed a large rice cultivation in the following 2017 dry season (58% of rice for a SSM anomaly of  $0.1 \text{ m}^3/\text{m}^3$  and  $0.2 \text{ m}^3/\text{m}^3$  for SMOS and SMAP respectively).

We characterize the amplitude of the SSM annual peaks by the maximum value of SSM anomaly between December and March (resp. May and November) in the dry (resp. wet) season. Table I reports the Pearson correlation coefficients between the amplitude of the 3 SSM datasets and the percentage of rice cover (obtained from land cover maps produced as explained in section II-B1) for the dry and wet seasons between 2016 and 2019. The last line presents the correlation coefficients for all years, cumulating both the spatial and temporal effects. Rice cover and SSM SMOS L3 and SMAP L3E show the same spatial pattern (as illustrated for the dry season of 2018 in figure 3a), resulting in spatial correlations close to 1 in Table I (0.81 and 0.77 on average for SMOS and SMAP, respectively). Correlations between rice cover and the SSM peak amplitude are lower during the dry season of 2016 because of the small amplitude of SSM peaks due to the small extent of sown rice in this exceptionally dry year. Correlations tend to be lower during the wet season because the SSM amplitude is due both to rainfall and irrigation. The spatial correlations with SSM SMAP L4 are much lower than with L3 SSM products.

The coefficient of determination for the linear regression fit between rice cover percentage and SSM data (red line in Figure 3b) ranges from 0.32 to 0.65 (resp. from 0.23 to 0.45) for SMOS L3 SSM (resp. SMAP L3 SSM) during the dry season, but is much lower during the wet season ( $<0.23$  and  $<0.15$  for SMOS and SMAP respectively). This demonstrates the potential of SMOS L3 SSM in particular as a predictor for rice percentage cover during the dry season (65% of variance explained in the dry season of 2017).

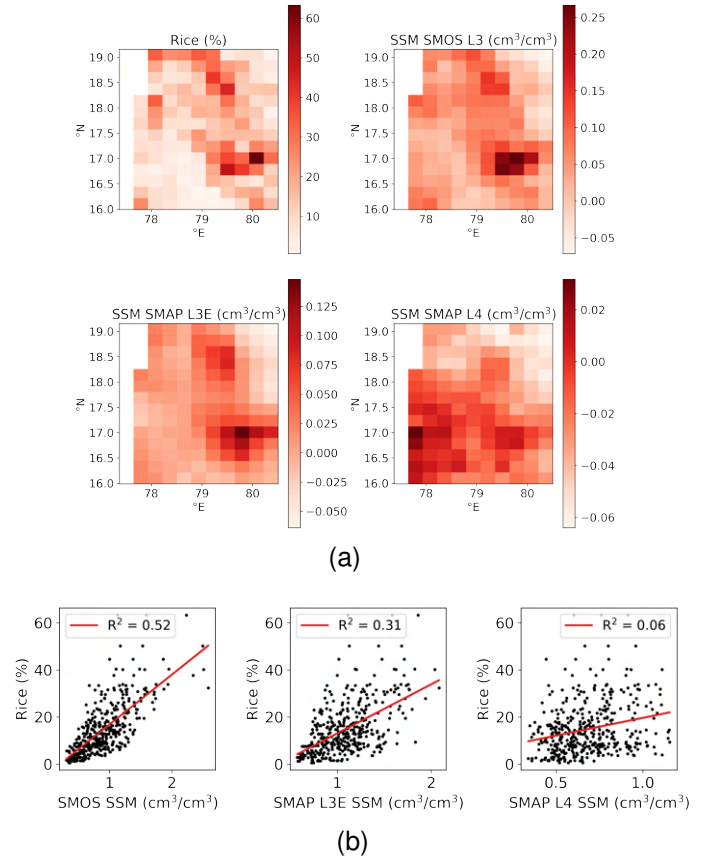


Fig. 3. (a) Maps of rice cover (%) and amplitude of the peak during the 2018 dry season for SMOS L3, SMAP L3E and SMAP L4 SSM anomalies ( $\text{m}^3/\text{m}^3$ ) (obtained by subtracting the temporal average of SSM values between 2016 and 2019 pixel-wise). (b) Scatterplot of rice cover and the SSM products. The linear regression fit to the data points and its coefficient of determination ( $R^2$ ) are indicated by the red line.

## IV. CONCLUSION

SMOS and SMAP L-band data have been used for irrigation mapping and irrigation water quantification because of their sensitivity to soil moisture and open water. This study demonstrates for the first time the potential of SMOS and SMAP L3 SSM data to quantify the fraction of flood-irrigated crops within a 25 km pixel in Telangana, a State in southern India. Areas with a large fraction of cultivated rice present a clear bimodal annual cycle with two SSM peaks in the wet (resp. dry) season associated with the combined effect of precipitation and irrigation (resp. the sole effect of irrigation). We identified a high correlation ( $R=0.81$ ) between rice cover percentage and the amplitude of the SSM peak during the dry season. The correlation with the amplitude of the SSM peak during the wet season is still significant but is lower ( $R=0.68$ ) due to the mixed contribution of irrigation and precipitation to the SSM dynamics. We demonstrated the potential of SMOS L3 SSM for rice cover percentage estimation (up to 65% of variance explained). It should be noted that the particular sensitivity of the L3 SSM product (water content of unsaturated soil) of SMOS and SMAP to flood irrigation is due to the presence of standing water (saturated areas) that should normally be removed by SMOS and SMAP preprocessing algorithms.

This approach could be used in combination with the knowledge of rice water needs and cultural practices to estimate the water used for irrigation. Note that this study was conducted at the scale of the SMOS grid resolution (25 km), a relevant scale for land management but still coarse compared to the size of irrigated perimeters. A higher resolution can be obtained with downscaling algorithms but at the expense of SSM accuracy [20]. Future L-band mission concepts such as SMOS-HR [28] at a spatial resolution of  $\approx 10$  km will allow to make a significant step forward for this application.

#### ACKNOWLEDGMENTS

Field data for the production of land cover maps were surveyed with the help of the IFCGR (Indo-French Center for Groundwater Research), an Indo-French collaboration between the BRGM (Bureau des ressources géologiques et minières) and the NGRI (Indian National Geophysical Research Institute). Support from the Horizon 2020 ACCWA project (grant agreement n°823965) in the context of Marie Skłodowska-Curie research and innovation staff exchange (RISE) program is acknowledged.

#### REFERENCES

- [1] J. A. Foley, N. Ramankutty, K. A. Brauman, E. S. Cassidy, J. S. Gerber, M. Johnston, N. D. Mueller, C. O'Connell, D. K. Ray, P. C. West, C. Balzer, E. M. Bennett, S. R. Carpenter, J. Hill, C. Monfreda, S. Polasky, J. Rockström, J. Sheehan, S. Siebert, D. Tilman, and D. P. M. Zaks, "Solutions for a cultivated planet," *Nature*, vol. 478, no. 7369, pp. 337–342, Oct. 2011.
- [2] Y. Wada, D. Wisser, and M. F. P. Bierkens, "Global modeling of withdrawal, allocation and consumptive use of surface water and groundwater resources," *Earth System Dynamics*, vol. 5, no. 1, pp. 15–40, Jan. 2014.
- [3] M. Rahman, L. Di, E. Yu, L. Lin, C. Zhang, and J. Tang, "Rapid Flood Progress Monitoring in Cropland with NASA SMAP," *Remote Sensing*, vol. 11, no. 2, p. 191, Jan. 2019.
- [4] O. Calla, K. Gadhri, R. Sharma, S. Aghajari, A. Kalla, and G. Rathore, "Microwave remote sensing application for monitoring of floods," *MAUSAM*, vol. 65, no. 2, pp. 141–152, Apr. 2014.
- [5] J. Du, J. S. Kimball, J. Galantowicz, S.-B. Kim, S. K. Chan, R. Reichle, L. A. Jones, and J. D. Watts, "Assessing global surface water inundation dynamics using combined satellite information from SMAP, AMSR2 and Landsat," *Remote Sensing of Environment*, vol. 213, pp. 1–17, Aug. 2018.
- [6] E. Fluet-Chouinard, B. Lehner, L.-M. Rebelo, F. Papa, and S. K. Hamilton, "Development of a global inundation map at high spatial resolution from topographic downscaling of coarse-scale remote sensing data," *Remote Sensing of Environment*, vol. 158, pp. 348–361, Mar. 2015.
- [7] J. Du, J. S. Kimball, L. A. Jones, and J. D. Watts, "Implementation of satellite based fractional water cover indices in the pan-Arctic region using AMSR-E and MODIS," *Remote Sensing of Environment*, vol. 184, pp. 469–481, Oct. 2016.
- [8] B. T. Gouweleuw, A. I. J. M. van Dijk, J. P. Guerschman, P. Dyece, and M. Owe, "Space-based passive microwave soil moisture retrievals and the correction for a dynamic open water fraction," *Hydrology and Earth System Sciences*, vol. 16, no. 6, pp. 1635–1645, Jun. 2012.
- [9] M. Parrens, A. Al Bitar, F. Frappart, F. Papa, S. Calmant, J.-F. Créteaux, J.-P. Wigneron, and Y. Kerr, "Mapping Dynamic Water Fraction under the Tropical Rain Forests of the Amazonian Basin from SMOS Brightness Temperatures," *Water*, vol. 9, no. 5, p. 350, May 2017.
- [10] M. Parrens, A. A. Bitar, F. Frappart, R. Paiva, S. Wongchuig, F. Papa, D. Yamasaki, and Y. Kerr, "High resolution mapping of inundation area in the Amazon basin from a combination of L-band passive microwave, optical and radar datasets," *International Journal of Applied Earth Observation and Geoinformation*, vol. 81, pp. 58–71, Sep. 2019.
- [11] P. M. Lawston, J. A. Santanello Jr, and S. V. Kumar, "Irrigation Signals Detected From SMAP Soil Moisture Retrievals," *Geophysical Research Letters*, vol. 44, no. 23, pp. 11,860–11,867, 2017.
- [12] M. J. Escorihuela and P. Quintana-Seguí, "Comparison of remote sensing and simulated soil moisture datasets in Mediterranean landscapes," *Remote Sensing of Environment*, vol. 180, pp. 99–114, Jul. 2016.
- [13] J. Dari, P. Quintana-Seguí, M. J. Escorihuela, V. Stefan, L. Brocca, and R. Morbidelli, "Detecting and mapping irrigated areas in a Mediterranean environment by using remote sensing soil moisture and a land surface model," *Journal of Hydrology*, vol. 596, p. 126129, May 2021.
- [14] S. V. Kumar, C. D. Peters-Lidard, J. A. Santanello, R. H. Reichle, C. S. Draper, R. D. Koster, G. Nearing, and M. F. Jasinski, "Evaluating the utility of satellite soil moisture retrievals over irrigated areas and the ability of land data assimilation methods to correct for unmodeled processes," *Hydrology and Earth System Sciences*, vol. 19, no. 11, pp. 4463–4478, Nov. 2015.
- [15] L. Brocca, A. Tarpanelli, P. Filippucci, W. Dorigo, F. Zaussinger, A. Gruber, and D. Fernández-Prieto, "How much water is used for irrigation? A new approach exploiting coarse resolution satellite soil moisture products," *International Journal of Applied Earth Observation and Geoinformation*, vol. 73, pp. 752–766, Dec. 2018.
- [16] J. Dari, L. Brocca, P. Quintana-Seguí, M. J. Escorihuela, V. Stefan, and R. Morbidelli, "Exploiting High-Resolution Remote Sensing Soil Moisture to Estimate Irrigation Water Amounts over a Mediterranean Region," *Remote Sensing*, vol. 12, no. 16, p. 2593, Jan. 2020.
- [17] P. Filippucci, A. Tarpanelli, C. Massari, A. Serafini, V. Strati, M. Alberi, K. G. C. Raptis, F. Mantovani, and L. Brocca, "Soil moisture as a potential variable for tracking and quantifying irrigation: A case study with proximal gamma-ray spectroscopy data," *Advances in Water Resources*, vol. 136, p. 103502, Feb. 2020.
- [18] E. Jalilvand, M. Tajrishy, S. A. Ghazi Zadeh Hashemi, and L. Brocca, "Quantification of irrigation water using remote sensing of soil moisture in a semi-arid region," *Remote Sensing of Environment*, vol. 231, p. 111226, Sep. 2019.
- [19] L. Brocca, L. Ciabatta, C. Massari, T. Moramarco, S. Hahn, S. Hase-nauer, R. Kidd, W. Dorigo, W. Wagner, and V. Levizzani, "Soil as a natural rain gauge: Estimating global rainfall from satellite soil moisture data," *Journal of Geophysical Research: Atmospheres*, vol. 119, no. 9, pp. 5128–5141, 2014.
- [20] J. Peng, A. Loew, O. Merlin, and N. E. C. Verhoest, "A review of spatial downscaling of satellite remotely sensed soil moisture," *Reviews of Geophysics*, vol. 55, no. 2, pp. 341–366, 2017.
- [21] S. Ferrant, A. Selles, M. Le Page, P.-A. Herrault, C. Pelletier, A. Al-Bitar, S. Mermoz, S. Gascoin, A. Bouvet, M. Saqalli, B. Dewandel, Y. Caballero, S. Ahmed, J.-C. Maréchal, and Y. Kerr, "Detection of Irrigated Crops from Sentinel-1 and Sentinel-2 Data to Estimate Seasonal Groundwater Use in South India," *Remote Sensing*, vol. 9, no. 11, p. 1119, Nov. 2017.
- [22] J. Inglada, A. Vincent, M. Arias, B. Tardy, D. Morin, and I. Rodes, "Operational High Resolution Land Cover Map Production at the Country Scale Using Satellite Image Time Series," *Remote Sensing*, vol. 9, no. 1, p. 95, Jan. 2017.
- [23] G. Pingle, "Irrigation in Telangana: The Rise and Fall of Tanks," *Economic & Political Weekly*, p. 9, 2011.
- [24] O. Hagolle, M. Huc, C. Desjardins, S. Auer, and R. Richter, "MAJA Algorithm Theoretical Basis Document," Dec. 2017.
- [25] A. Al Bitar, A. Mialon, Y. H. Kerr, F. Cabot, P. Richaume, E. Jacqueline, A. Quesney, A. Mahmoodi, S. Tarot, M. Parrens, A. Al-Yaari, T. Pellarin, N. Rodriguez-Fernandez, and J.-P. Wigneron, "The global SMOS Level 3 daily soil moisture and brightness temperature maps," *Earth System Science Data*, vol. 9, no. 1, pp. 293–315, Jun. 2017.
- [26] M. J. Brodzik, B. Billingsley, T. Haran, B. Raup, and M. H. Savoie, "EASE-Grid 2.0: Incremental but Significant Improvements for Earth-Gridded Data Sets," *ISPRS International Journal of Geo-Information*, vol. 1, no. 1, pp. 32–45, Jun. 2012.
- [27] R. Reichle, R. Koster, G. D. Lannoy, W. Crow, and J. Kimball, "Level 4 Surface and Root Zone Soil Moisture (L4\_SM) Data Product," Dec. 2014.
- [28] N. Rodríguez-Fernández, A. Mialon, O. Merlin, C. Suere, F. Cabot, A. Khazāal, J. Costeraste, B. Palacin, R. Rodríguez-Suquet, T. Tournier, T. Decoopman, É. Anterrieu, M. Colom, J. Morel, Y. Kerr, B. Rougé, J. Boutin, G. Picard, T. Pellarin, M. Escorihuela, A. A. Bitar, and P. Richaume, "SMOS-HR: A High Resolution L-Band Passive Radiometer for Earth Science and Applications," *undefined*, 2019.

---

## 3.6 Ce qu'il faut retenir

- Le TWS de GRACE présente une bonne correspondance avec les données de précipitations TRMM et les données piézométriques.
- Aucune tendance positive à long terme n'a été identifiée sur le signal GRACE entre 2002 et 2020 sur le Telangana. Ceci infirme les résultats de la littérature suggérant une tendance à la hausse du TWS sur l'Inde du Sud, qui semble due aux phénomènes climatiques de court terme et dépendent de la période temporelle d'étude.
- GRACE peut être utilisé pour prédire le stock d'eau souterraine à large échelle, avec une bonne correspondance avec les données *in situ* ( $R^2 = 0.70$ ), compte tenu (i) des disparités entre les modes et fréquences d'acquisitions des différents jeux de données utilisés dans ce type d'approche (GRACE, modèle, *in situ*) et (ii) des faibles propriétés de stockages des aquifères du Telangana et de la recharge modeste.
- Les humidités de surface de SMOS et SMAP présentent un cycle annuel bimodal avec un pic en période sèche (absent sur l'humidité de surface modélisée par GLDAS-NOAH et GLEAM) sur des zones très irriguées, dont l'intensité est fortement corrélée avec la présence de riz ( $R = 0.81$ ).
- Ce pic en saison sèche peut être relié temporellement et quantitativement à d'autres données ayant un lien avec l'irrigation (surface de riz cultivé, NDVI, piézométrie).
- À ce jour, les méthodes actuelles peinent à estimer l'irrigation à partir des données d'humidité dans les régions où les cultures sont irriguées lors des épisodes de pluie.





# Chapitre 4

## Désagrégation des données GRACE

### Sommaire

---

<b>4.1</b>	<b>Introduction</b>	<b>68</b>
<b>4.2</b>	<b>Problématiques et applications sur le Telangana</b>	<b>68</b>
4.2.1	Problématiques	68
4.2.2	Comparaison de différents modèles	70
4.2.3	Limites et perspectives	77
<b>4.3</b>	<b>Evaluating downscaling methods of GRACE data : a case study over a fractured crystalline aquifer in South India</b>	<b>79</b>
<b>4.4</b>	<b>Ce qu'il faut retenir</b>	<b>98</b>

---

## 4.1 Introduction

L'eau souterraine est une ressource essentielle pour l'irrigation, en particulier dans les zones semi-arides comme le sud de l'Inde. Les aquifères s'épuisent dans plusieurs parties du monde depuis des décennies et ces ressources devraient continuer à se raréfier (Wada et al., 2012). La surveillance et la gestion précautionneuse de ces ressources est donc cruciale. Le suivi des stocks d'eau souterraine est généralement assuré par des réseaux de puits d'observation, mais reste difficile du fait de leur couverture éparse, de la nature ponctuelle de la donnée, de leur abandonnement progressif ou des biais d'observation (Hora et al., 2019). En cela, les données gravimétriques fournies par les missions GRACE et GRACE-FO sont des données d'exception en hydrologie. Les mesures des satellites permettent de dériver depuis 2002 des anomalies mensuelles et spatialisées de stock d'eau sans limite de profondeur (TWS - *Total Water Storage*). Les données GRACE sont largement utilisées en hydrologie pour étudier l'évolution à long terme du stock d'eau souterraine (GWS, *GroudWater Storage*) en soustrayant à l'anomalie de TWS les stocks de surface et d'humidité du sol, à échelle régionale ou globale (Breña-Naranjo et al. (2014); Cao and Roy (2020); Frappart et al. (2019); Papa et al. (2015); Rodell et al. (2018); Rzepecka and Birylo (2020); Tiwari et al. (2009); Zhang et al. (2020), en particulier Frappart and Ramillien (2018) pour une synthèse bibliographique). Elles sont une alternative pour étudier les ressources en eau souterraine dans les régions ne possédant pas de réseau dense d'observation.

Cependant, l'utilisation des données GRACE à une échelle pertinente pour les usages agricoles reste limitée en raison de leur faible résolution native (environ 330 km) (Schmidt et al., 2008; Tapley et al., 2004a). Aussi, cette dernière décennie a vu plusieurs études proposer des méthodes pour désagréger les données GRACE pour la production de cartes de GWS à une résolution plus élevée (généralement plusieurs dizaines de km).

Ce chapitre s'intéresse à l'application de ces méthodes sur l'État du Telangana (Inde du Sud), caractérisé par une forte demande en eau souterraine pour l'irrigation et un aquifère granitique de faible capacité de stockage. Une première partie présente les problématiques liées à la désagrégation des données GRACE ainsi que les résultats de plusieurs modèles de désagrégation sur la partie granitique du Telangana. La deuxième partie est un article publié dans *Hydrology and Earth System Sciences* (HESS) qui synthétise les méthodes de validation des produits GRACE désagrégés utilisées dans la littérature, et qui propose un nouveau schéma de validation comblant les lacunes existantes dans le domaine.

## 4.2 Problématiques et applications sur le Telangana

### 4.2.1 Problématiques

#### 4.2.1.1 Méthodes de désagrégation

Deux grands types de méthodes sont utilisées pour la désagrégation des données GRACE dans la littérature : les méthodes basées sur des modèles et des approches basées sur des relations empiriques entre des prédicteurs (respectivement nommées *dynamic downscaling* et *statistical downscaling* dans la littérature).

Les approches type modèle consistent à assimiler les données GRACE dans des modèles hydrologiques qui génèrent GWS à une résolution spatiale et/ou temporelle supérieure à celle de GRACE (Giroto et al., 2016; Houborg et al., 2012; Nie et al., 2019; Schumacher et al., 2018; Tian et al., 2017; Zaitchik et al., 2008). Cette approche souffre cependant (i) de la différence de résolution

entre les données GRACE et les données d'entrée du modèle et (ii) des limitations des modèles : hypothèses, paramétrisation, incertitude des données de forçage météorologique, et particulièrement l'absence de prise en compte dans le modèle des activités anthropiques telles que l'irrigation (Long et al., 2013). Ceci est particulièrement dommageable dans une région telle que le Telangana où les activités d'irrigation sont particulièrement visibles par satellite (voir chapitre 3).

Les approches empiriques peuvent elles-mêmes être divisées en plusieurs catégories. Certaines approches prédisent directement les niveaux d'eau des puits d'observation à partir du TWS de GRACE associé à d'autres prédicteurs à l'aide de réseaux de neurones (Gemitzi and Lakshmi, 2017, 2018; Miro and Famiglietti, 2018; Sun, 2013) ou d'autres modèles empiriques (Milewski et al., 2019). L'importance relative du TWS de GRACE dans la prédiction du modèle peut varier de moins de 10% (Gemitzi and Lakshmi, 2018) à 35% (Miro and Famiglietti, 2018). Ces approches ont cependant des inconvénients : (i) elles ne permettent qu'un suivi très local de l'eau souterraine et (ii) elles utilisent les données *in situ* pour la calibration, et contraignent donc ces méthodes à des zones possédant un réseau dense d'observation. Parallèlement, un autre type de méthode consiste à produire des cartes de GWS désagrégé à partir des données GRACE et n'utilise les données *in situ* que pour la validation. Ce type de méthode consiste à (i) dériver un modèle de TWS à partir de données auxiliaires disponibles à haute résolution (HR), (ii) calibrer ce modèle à basse résolution (LR, *Low Resolution*), (iii) l'appliquer à HR pour prédire TWS et (iv) soustraire la contribution des eaux de surface et d'humidité du sol pour isoler GWS (déconvolution). Cette approche repose sur l'hypothèse que les processus hydrologiques et physiques qui lient ces variables sont les mêmes à toutes les résolutions (Ali et al., 2021; Jyolsna et al., 2021; Karunakalage et al., 2021; Sahour et al., 2020; Seyoum and Milewski, 2017; Vishwakarma et al., 2021; Zhang et al., 2021a,b). Notons que certaines études couplent les deux méthodes (Seyoum et al., 2019; Zhang et al., 2021a).

#### 4.2.1.2 Échelle de désagrégation

Dans la littérature, les approches empiriques sont utilisées pour désagréger les données de GRACE à plusieurs échelles, soit à l'échelle d'un bassin versant comme Seyoum and Milewski (2017) (5 000 km<sup>2</sup> to 20 000 km<sup>2</sup>), soit sur une grille d'une résolution donnée. La résolution de désagrégation est souvent limitée par la résolution la plus grossière des prédicteurs : les précipitations TRMM ou GLDAS à 0.25° (Ali et al., 2021; Jyolsna et al., 2021; Ning et al., 2014; Seyoum et al., 2019; Zhang et al., 2021a), le modèle NLDAS (NASA's North American Land Data Assimilation System) à 0.125° (Sahour et al., 2020), le modèle EALCO (Ecological Assimilation of Land and Climate Observations) à 5 km (Zhong et al., 2021) ou l'évapotranspiration de MODIS à 2 km (Yin et al., 2018). Des méthodes d'interpolation peuvent viser des résolutions encore plus fines, jusqu'au km (Zhang et al., 2021a; Zuo et al., 2021), mais elles ne sont pas basées sur des processus physiques ou hydrologiques.

#### 4.2.1.3 Méthodes de validation

La méthode de validation est essentielle pour évaluer la pertinence des cartes produites pour à terme les utiliser. À ce jour, les stratégies de validation reposent essentiellement sur la correspondance entre les séries temporelles des produits désagrégés et les données *in situ*, par l'appréciation arbitraire de métriques (R, R<sup>2</sup>, RMSE) ou de leurs tendances temporelles. Cependant cette approche de validation manque d'un aspect spatial et d'une comparaison avec les données initiales non désagrégées.

Les méthodes de validation spatiales sont davantage représentées dans la désagrégation des

mesures satellitaires d’humidité du sol d’AMSR-E (Fang et al., 2013; Kim and Hogue, 2012; Zhao and Li, 2013), de SMOS (Merlin et al., 2013, 2015; Sabaghy et al., 2020) et de SMAP (Das et al., 2014; Fang et al., 2018; Sabaghy et al., 2020), ou de la température de surface de MODIS (Merlin et al., 2012), où elles sont généralement utilisées en combinaison avec des méthodes de validation temporelles.

L’article en section 4.3 synthétise les méthodes de validation des données GRACE désagrégées présentes dans la littérature et propose l’utilisation d’un schéma de validation complet, à la fois temporel et spatial, permettant de quantifier l’apport de la désagrégation par rapport aux données LR initiales. Au lieu de calculer seulement une métrique entre le GWS désagrégé et le GWS *in situ*, un gain (similaire à Merlin et al. (2015)) mesure l’amélioration de l’accord avec les données *in situ* entre LR et HR (équation 4.1) :

$$G = \frac{|M_{opt} - M_{LR}| - |M_{opt} - M_{HR}|}{|M_{opt} - M_{LR}| + |M_{opt} - M_{HR}|} \quad (4.1)$$

avec  $M_{LR}$  la métrique calculée entre les données *in situ* HR et le GWS LR,  $M_{HR}$  sa valeur entre les données *in situ* HR et le GWS désagrégé, et  $M_{opt}$  sa valeur optimale (e.g. 1 pour R, 0 pour RMSE). Ce gain est calculé pour plusieurs métriques dans le temps (R,  $R^2$  et RMSE) et dans l’espace (R, RMSE, pente et biais). Un gain positif indique que les données *in situ* sont plus proches des cartes de GWS désagrégé que le GWS à LR (cas sans désagrégation).

#### 4.2.1.4 Échelle d’utilisation de GRACE

Les études de désagrégation utilisent fréquemment les données GRACE à une échelle inférieure à leur résolution théorique (300 à 400 km (Schmidt et al., 2008; Tapley et al., 2004a)). En effet, les grilles de TWS sont fournies avec des facteurs d’échelles à une résolution de  $1^\circ$  pour les solutions dérivées des harmoniques sphériques (SH) (Landerer and Swenson, 2012) et  $0.5^\circ$  pour les mascons du *Jet Propulsion Laboratory* (Wiese et al., 2016). Ces facteurs d’échelle ont pour but de restaurer le signal de GRACE perdu lors de son traitement et permettent aux utilisateurs de moyennner le TWS suréchantillonné sur des zones d’études arbitraires, pourvu qu’elles soient au moins aussi grandes que la résolution de GRACE (Landerer and Swenson, 2012). Les facteurs d’échelles n’ont pas vocation à désagréger les données GRACE du fait de la forte dépendance entre les pixels suréchantillonnés voisins (Landerer and Swenson, 2012). Cependant, beaucoup d’études utilisent directement les solutions SH à  $1^\circ$  (Ali et al., 2021; Jyolsna et al., 2021; Karunakalage et al., 2021; Ning et al., 2014; Seyoum et al., 2019; Yin et al., 2018; Zhang et al., 2021a,b; Zuo et al., 2021) où les mascons à  $0.5^\circ$  (Karunakalage et al., 2021; Nie et al., 2019; Tian et al., 2017) comme donnée à LR, ce qui est bien plus fin que leur vraie résolution. Il n’existe pas d’étude évaluant l’incertitude des produits GRACE désagrégés associée à l’utilisation des données GRACE à la résolution de leur facteur d’échelle au lieu de leur résolution réelle.

## 4.2.2 Comparaison de différents modèles

Cette section s’applique à désagréger les données de TWS de GRACE sur le Telangana tout en abordant les problématiques détaillées en section précédente. Une variété de méthodes de désagrégation ont été testées pour désagréger le TWS mensuel de GRACE en variant les prédicteurs, le modèle, la période (modèle en saison humide/saison sèche), mais seuls quelques résultats sont présentés ici.

**Zone d'étude, LR et HR.** Comme expliqué en section 2.3.1.3, les données GRACE RL06M sont fournies avec une grille de facteur d'échelle de  $0.5^\circ$  permettant à l'utilisateur de moyenniser les pixels ré-échantillonnés dans une zone d'étude arbitrairement choisie de surface suffisamment grande. La zone d'étude est la partie granitique du Telangana, détournée à la résolution de  $0.5^\circ$ , d'une surface de  $113\,000\text{ km}^2$  (zone rose sur la figure 1.2), qui constitue la LR. L'anomalie du TWS de GRACE moyennée sur cette zone ( $TWSA_{GRACE,LR}$ ) est désagrégée à  $0.5^\circ$  (HR).

**Validation.** Le GWS désagrégé est validé avec le GWS calculé à partir des données *in situ* avec une porosité efficace ( $S_y$ , *Specific Yield*) de 4.7% (produites comme en section 3.2.3), appelée GWS-OW (*Observation Well*). Le gain présenté en équation 4.1 mesure l'amélioration du GWS désagrégé par rapport à une référence LR. Cette référence est définie par  $TWSA_{GRACE,LR}$  déconvolué par le RZSM de GLEAM à  $0.5^\circ$ , appelée GWS-LRref. Le gain est calculé pour plusieurs métriques dans le temps (R,  $R^2$  et RMSE) et dans l'espace (R, RMSE, pente et biais). La pente est un indicateur commun pour évaluer les produits désagrégés (Merlin et al., 2015; Sabaghy et al., 2020), puisque qu'il est attendu que le GWS désagrégé soit plus proche des données *in situ* et que sa variabilité soit plus grande qu'à basse résolution.

La résolution visée de  $0.5^\circ$  permet de comparer les résultats des méthodes de désagrégation avec le produit du TWS de GRACE et du facteur d'échelle à  $0.5^\circ$  (appelé SF, *scaling factor*) en le considérant comme un produit désagrégé.

**Types de modèles.** La désagrégation du TWS de GRACE est d'abord réalisée avec des modèles empiriques statiques consistant à prédire les anomalies de TWS avec des données auxiliaires (modèles utilisés dans l'article en section 4.3 et présentés dans la section précédente), qui sont largement représentés dans la littérature. Un nouveau type de modèle est également développé et modélise la variation mensuelle de TWS ( $\Delta TWS$ ). Ce type de modèle, appelé par la suite modèle "dynamique", modélise les flux plutôt que les stocks hydrologiques.

#### 4.2.2.1 Modèles statiques empiriques

**Description.** La méthode de désagrégation utilisée consiste à entraîner un modèle empirique de TWS à LR à partir de variables auxiliaires, puis l'appliquer à HR, de manière similaire à Ali et al. (2021); Jyolsna et al. (2021); Karunakalage et al. (2021); Sahour et al. (2020); Seyoum and Milewski (2017); Vishwakarma et al. (2021); Zhang et al. (2021a,b). Le fonctionnement de la méthode est détaillé en figure 4.1. Les prédicteurs utilisés pour entraîner le modèle sont agrégés à LR ( $113\,000\text{ km}^2$ ) puis un modèle statistique est entraîné à cette résolution pour prédire le TWS en fonction de ces prédicteurs. Ce modèle est ensuite appliqué à HR avec les prédicteurs agrégés à  $0.5^\circ$  avec une méthode bilinéaire pour prédire le TWS à  $0.5^\circ$ . Enfin, le GWS est obtenu par soustraction du RZSM de GLEAM au TWS à  $0.5^\circ$  (déconvolution).

**Données.** Les prédicteurs des modèles statistiques sont choisis de sorte à identifier les principaux phénomènes hydrologiques et anthropiques en essayant de s'affranchir le plus possible des données de modèle : les précipitations, l'indice de végétation NDVI (et sa variation d'un mois sur l'autre  $\Delta NDVI$ ), et l'humidité dans les cinq premiers cm de sol (SSM); ainsi que leurs accumulations sur l'année précédente comme indicateur de l'état de l'aquifère. Certains prédicteurs ont un délai de quelques mois par rapport au TWS du fait de la lenteur des processus hydrologiques (mise en évidence dans le chapitre 3, section 3.5). Ce délai (en mois) est déterminé pour chaque prédicteur en maximisant leur corrélation le TWS de GRACE à LR (Sahour et al., 2020; Seyoum and Milewski,

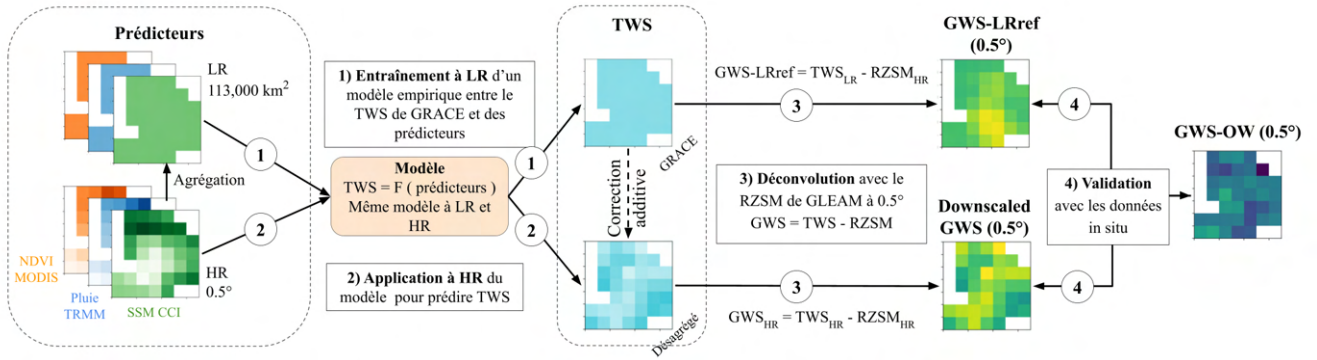


FIGURE 4.1 – Logigramme de la désagrégation empirique statique.

2017). Plus de détails sur les délais optimaux pour chaque prédicteurs sont disponibles dans l'article en section 4.3. À noter que toutes les données (TWS, prédicteurs et données de validation) sont converties en anomalies par rapport à la période 2007-2015 (période couvrant l'ensemble des jeux de données sans contenir une trop grande portion de la série temporelle interrompue de GRACE).

**Modèles.** Deux modèles sont utilisés : un Modèle Linéaire multiple (ML) et un modèle *Random Forest* (RF) (Breiman, 2001), moins interprétable mais plus souple que ML. Plus d'informations sur ces modèles sont dans l'article en section 4.3. Une sélection de variables détermine le modèle le plus parcimonieux. Plus de détails sur la sélection de variables sont disponibles dans l'article en section 4.3.

**Correction additive.** Les prédictions de TWS HR pour chaque pas de temps  $t$  et chaque pixel HR  $i$  sont débiaisées avec une correction additive de sorte que la moyenne spatiale du TWS HR prédit à chaque mois soit égale à  $TWS_{GRACE,LR}$  (équation 4.2) :

$$TWS_{HR,t,i}^{corr} = TWS_{HR,t,i} + TWS_{GRACE,LR,t} - \frac{\sum_i TWS_{HR,t,i}}{n_{pix}} \quad (4.2)$$

avec  $TWS_{HR,t,i}$  le TWS HR prédit par le modèle pour le mois  $t$  et le pixel  $i$ ,  $TWS_{HR,t,i}^{corr}$  le TWS corrigé,  $TWS_{GRACE,LR,t}$  le TWS LR de GRACE à la date  $t$ , et  $\frac{\sum_i TWS_{HR,t,i}}{n_{pix}}$  la moyenne spatiale du TWS HR prédit à la date  $t$ .

**Calibration des modèles à LR.** Les modèles ML et RF sont entraînés sur 80 % de l'ensemble de la série temporelle (174 points au total). Les variables sélectionnées et les performances du modèle sont rapportées dans la table 4.1. Le modèle RF a un meilleur  $R^2$  que le modèle ML (0.97 contre 0.90), mais le RMSE sur l'ensemble de test est bien plus grand que sur l'ensemble d'apprentissage (4,6 cm contre 1.9 cm), révélant que le RF souffre de surapprentissage. Le RMSE sur l'ensemble de test est respectivement de 5.0 et 4.6 cm pour les modèles ML et RF, ce qui représente 7% et 6% de l'amplitude totale de  $TWS_{GRACE,LR}$  sur la région pendant la période d'étude (71 cm). Les deux modèles semblent pouvoir prédire  $TWS_{GRACE,LR}$  avec de bonnes performances.

**Résultats HR.** Après prédiction du GWS désagrégé par application du modèle à HR et déconvolution, les gains par rapport à GWS-LRref sont calculés dans le temps ( $R$ ,  $R^2$ , RMSE) et dans l'espace ( $R$ , RMSE, pente et biais). Le calcul des gains spatiaux et temporels est illustré en

TABLE 4.1 – Sélection de variables et performances des modèles prédictifs de TWS à basse résolution (LR). Le nombre entre parenthèses est le délai optimal (en mois) par rapport au TWS. Les variables avec le suffixe “cum.” sont cumulées sur l’année précédente. La performance du modèle est évaluée contre le TWS de GRACE à LR avec le  $R^2$  et le RMSE sur les échantillons d’apprentissage et de test.

Modèle	Variables sélectionnées							Performances du modèle		
	Pluie (2)	NDVI (0)	SM CCI (1)	$\Delta$ NDVI (2)	Pluie cum.	NDVI cum.	SM CCI cum.	RMSE appren-tissage (cm)	RMSE test (cm)	$R^2$ avec $TWS_{GRACE,LR}$
ML		X	X	X			X	5.2	5.0	0.90
RF	X	X	X	X	X		X	1.9	4.6	0.97

figure 4.2, et la formule de chaque gain est détaillée dans l’article en section 4.3. Les gains calculés sont récapitulés en figure 4.3. SF et ML dégradent la corrélation temporelle avec GWS-OW (gains négatifs sur  $R^2$ , R et RMSE, voir figure 4.3a) sur la plupart des pixels, contrairement au modèle RF qui semble apporter une amélioration significative (les gains moyens sur R,  $R^2$  et RMSE sont de +9.0, +0.79 et +5.4%). Les gains temporels de R sont meilleurs avec la correction du biais (en moyenne 6.5 et 6.7% pour ML CORR et RF CORR), tandis que les distributions des gains de  $R^2$  et RMSE sont plus condensées mais avec des moyennes plus basses de 3.2 et 1.6% pour ML CORR et 4.0 et 1.9% pour RF CORR. Dans le domaine spatial (voir figure 4.3b), la qualité du GWS SF est discutable. Le gain quasi-nul sur le biais était attendu puisque la seule différence entre le  $TWS_{GRACE,LR}$  et TWS SF est un facteur multiplicatif généralement proche de 1. GWS SF montre des gains positifs pour la pente et R sur la plupart des séries temporelles (12.9 et 13.9% respectivement en moyenne), mais au prix d’une incertitude plus élevée (-8.5% de gain sur le RMSE). Les nuages de points (résultats non présentés) indiquent que le gain positif sur la pente (i.e. la pente se rapproche de 1) est le plus souvent dû à une dispersion accrue causée par ce qui semble être un bruit supplémentaire apporté par la grille des facteurs d’échelle à chaque pas de temps. L’amélioration de la représentativité spatiale par ML CORR et RF CORR se traduit par des gains majoritairement positifs sur les pentes et R (23.0 et 28.8% pour ML CORR, 18.2 et 27.2% pour RF CORR) tout en maintenant un gain positif sur le RMSE (2.0 et 2.2% pour ML CORR et RF CORR respectivement). La correction additive explique le gain nul sur le biais pour ML CORR et RF CORR, par opposition à la distribution erratique des gains pour ML et RF, cependant elle baisse le gain moyen sur le RMSE (qui est de 3.9 et 7.9% pour ML et RF respectivement).

#### 4.2.2.2 Modèles dynamiques

Les modèles dynamiques développés dans cette étude s’appuient sur la modélisation des mouvements d’eau mensuels  $\Delta TWS$ . La modélisation des flux permet a priori de représenter les tendances interannuelles qui ne peuvent pas être prédites avec une méthode statique qui modélise les stocks à un instant donné. Cette section évalue le potentiel de ces modèles dynamiques.

La désagrégation du GWS à partir du TWS de GRACE se fait en quatre étapes : (i) modélisation de la variation de TWS d’un mois sur l’autre ( $\Delta TWS_t = TWS_t - TWS_{t-1}$ ) au lieu du TWS, (ii) initialisation du TWS mensuel par une valeur de TWS (*in situ* ou GRACE par exemple), (iii)



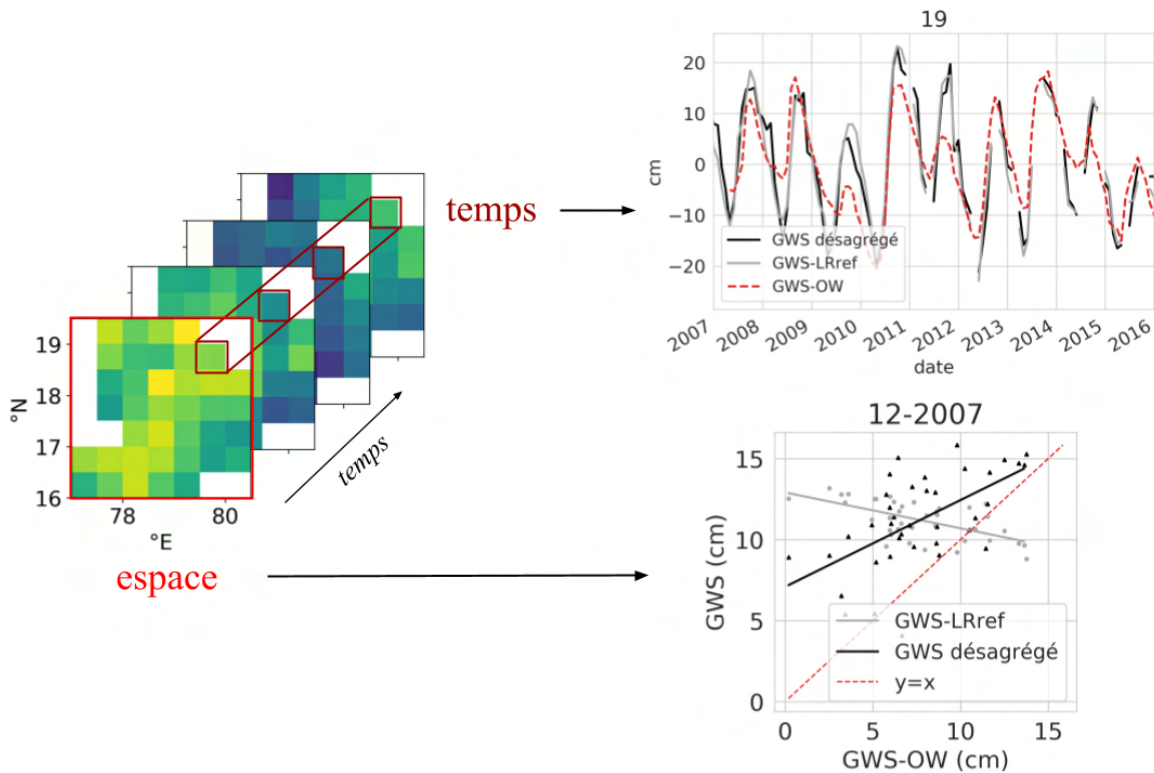


FIGURE 4.2 – Schéma explicatif du calcul des gains spatiaux et temporels. Les gains temporels sont calculés sur les séries temporelles de chaque pixel de  $0.5^\circ \times 0.5^\circ$  (illustration sur le pixel 19). Les gains spatiaux sont calculés sur le nuage de points à chaque pas de temps (illustration pour le mois de décembre 2007). Dans les deux cas, chaque métrique  $M_{HR}$  (voir équation 4.1) est calculée entre le GWS *in situ* (GWS-OW) et le GWS désagrégé, et  $M_{LR}$  entre GWS-OW et le GWS à basse résolution (GWS-LRref).

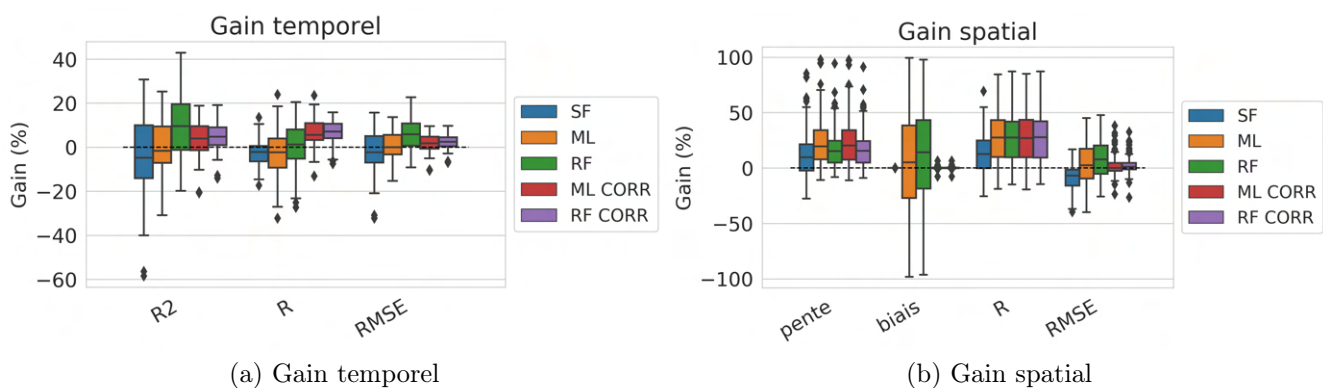


FIGURE 4.3 – Distribution des gains (a) temporels et (b) spatiaux pour le GWS produit à partir du facteur d’échelle (SF) et des modèles linéaire (ML) et random forest (RF). “CORR” désigne la correction additive.

reconstitution de la série temporelle par ajout des  $\Delta TWS$  modélisés à chaque pas de temps (équation 4.3), et enfin (iv) le TWS est déconvolué en GWS et validé contre les données *in situ*. Deux

types de modèles sont comparés : un modèle basé sur le bilan hydrique et un modèle empirique.

$$\text{TWS}_t = \Delta\text{TWS}_t + \text{TWS}_{t-1} \quad (4.3)$$

**Bilan hydrique.**  $\Delta\text{TWS}_t$  à HR est modélisé comme la différence des flux entrants (précipitations  $P_t$ ) et des flux sortants (évapotranspiration  $\text{ET}_t$  et ruissellement  $R_t$ ) du mois  $t$  :

$$\Delta\text{TWS}_t = P_t - \text{ET}_t - R_t \quad (4.4)$$

$P_t$  et  $R_t$  sont quantifiés respectivement avec les données TRMM et les sorties du modèles GLDAS-NOAH pour  $R_t$ . Le modèle est testé avec et sans ruissellement, pour évaluer si ce flux sortant est négligeable à l'échelle de la zone d'étude.  $\text{ET}_t$  est modélisé de plusieurs manières :

- par l'ET calculée par l'algorithme MOD16 à partir des données MODIS :  $\text{ET}_{\text{MODIS}}$ ,
- par le produit de l'évapotranspiration potentielle de référence (ET0) fournies par les données PET et du NDVI de MODIS :  $\text{ET0}_{\text{PET}} * \text{NDVI}_{\text{MODIS}}$ ,
- par  $Kc * \text{ET0}_{\text{PET}} * \text{NDVI}_{\text{MODIS}}$  avec  $Kc$  un coefficient cultural calibré à basse résolution par ajustement linéaire avec  $\Delta\text{TWS}$  de GRACE ( $\Delta\text{TWS}_{\text{GRACE,LR}}$ ). Le  $Kc$  vaut 1.39 (respectivement 0.95) lorsque calibré entre  $\Delta\text{TWS}_{\text{GRACE,LR}}$  et  $P_t - Kc * \text{ET0}_{\text{PET}} * \text{NDVI}_{\text{MODIS}}$  (respectivement  $P_t - Kc * \text{ET0}_{\text{PET}} * \text{NDVI}_{\text{MODIS}} - R_t$ ) avec un  $R^2$  de 0.66 (respectivement 0.52).

**Modèles empiriques.** De manière similaire aux modèles statiques, deux modèles empiriques (ML et RF) sont entraînés à LR pour prédire  $\Delta\text{TWS}$  en fonction de variables auxiliaires, les variables pertinentes et les modèles les plus parcimonieux sont sélectionnés, puis les modèles sont appliqués à HR. La table 4.2 reporte les variables sélectionnées et les performances de chaque modèle à LR.

TABLE 4.2 – Sélection de variables et performances des modèles prédictifs de  $\Delta\text{TWS}$  à LR.. Le nombre entre parenthèses est le délai optimal (en mois) par rapport au  $\Delta\text{TWS}$ . Les variables avec le suffixe “cum.” sont cumulées sur l'année précédente. La performance du modèle est évaluée contre le  $\Delta\text{TWS}$  de GRACE à LR avec le  $R^2$  et le RMSE sur les échantillons d'apprentissage et de test.

Modèle	Variables sélectionnées							Performances du modèle			
	Pluie (0)	NDVI (3)	$\Delta$ NDVI (0)	SM CCI (0)	$\Delta$ SM CCI (1)	Pluie cum.	NDVI cum.	SM CCI cum.	RMSE appren-tissage (cm)	RMSE test (cm)	$R^2$ avec $\Delta\text{TWS}_{\text{GRACE,LR}}$
ML	X	X	X	X	X	X	X	X	3.0	4.0	0.87
RF	X	X	X	X	X	X	X	X	1.3	3.7	0.95

**Résultats  $\Delta TWS$ .** Dans un premier temps la performance des  $\Delta TWS_{HR}$  modélisés (par bilan hydrique ou modèle empirique) est évaluée par rapport à la variation mensuelle du TWS dérivé des données *in situ* ( $\Delta TWS_{situ} = \Delta GWS-OW + \Delta RZSM_{GLEAM}$ ). Les gains sont calculés avec pour référence LR la variation mensuelle du TWS de GRACE ( $\Delta TWS_{GRACE,LR}$ ). Ce gain est aussi calculé pour la variation mensuelle du TWS de GRACE suréchantillonné par le facteur d'échelle à  $0.5^\circ$  ( $\Delta TWS_{SF}$ ). Les distributions des gains sont en figure 4.4. La modélisation par bilan hydrique dégrade le  $R^2$ , le R et le RMSE temporel, encore plus lorsqu'on inclut le ruissellement dans le modèle. Dans le domaine spatial, les gains de pente sont positifs, mais ceci est le plus souvent dû à une dispersion accrue des points par rapport à la basse résolution, comme le prouve le gain négatif sur le RMSE. Dans l'ensemble, les modèles empiriques ont une meilleure performance, en particulier le modèle RF avec des gains majoritairement positif dans le domaine spatial et dans le domaine temporel (pour le  $R^2$  et RMSE).

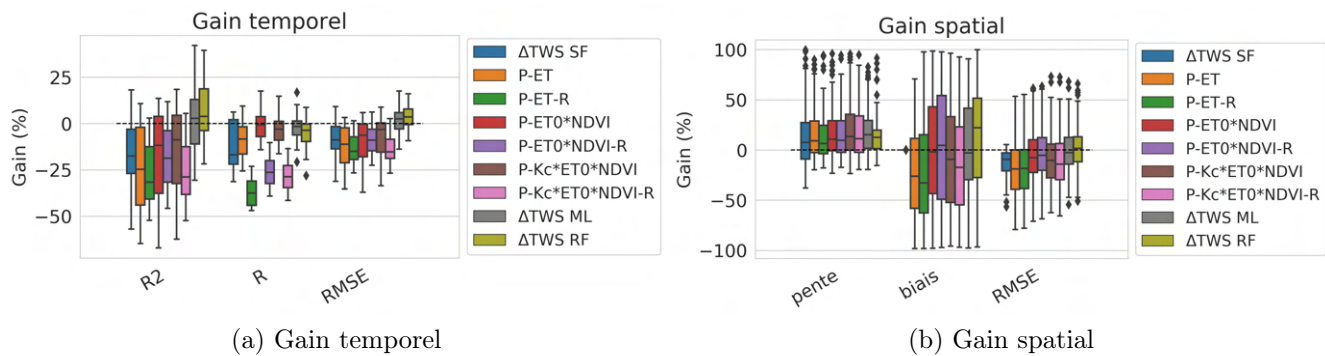


FIGURE 4.4 – Distribution des gains temporels et spatiaux de  $\Delta TWS_{HR}$ . Les gains sont calculés par rapport  $\Delta TWS_{situ}$  et la référence LR  $\Delta TWS_{GRACE}$ .

**Résultats TWS.** Le  $\Delta TWS$  modélisé par RF est le plus proche des données *in situ*, et est donc choisi pour reconstruire la série temporelle de TWS à HR. Elle est initialisée avec  $TWS_{GRACE,LR}$  au mois de mai 2003 (qui correspond à la période juste avant la mousson). Le modèle dynamique a tendance à diverger rapidement, aussi il est réinitialisé périodiquement avec la valeur de  $TWS_{GRACE,LR}$ , bien que ceci puisse avoir pour effet de briser la dynamique réelle des aquifères. La figure 4.5 montre la distribution des gains spatiaux et temporels des GWS désagrégés (calculés de la même manière que pour les modèles statiques, voir section 4.2.2) pour une réinitialisation tous les 1 à 10 ans ("0" indique que le modèle n'est initialisé qu'une fois en en mai 2003). La figure 4.6 illustre la série temporelle d'une case de  $0.5 \times 0.5^\circ$  avec réinitialisation tous les 4 ans. Le GWS désagrégé est moins proche de GWS-OW que GWS-LRref, à la fois dans le domaine spatial (gains négatifs de RMSE) et temporel (gains négatifs de  $R^2$ , R et RMSE). La qualité du GWS désagrégé est d'autant moins bonne avec la rareté de la réinitialisation avec  $TWS_{GRACE,LR}$ , ce qui est expliqué par la divergence rapide du modèle (visible en figure 4.6). Les modèles dynamiques appliqués dans cette étude, bien qu'ayant un fort potentiel dans la représentation des flux d'eau, présente de moins bons résultats que les modèles statiques. Ceci pourrait provenir des incertitudes importantes sur les champs de précipitation TRMM et d'ET0 et d'ET, limitantes pour ce type d'approche.

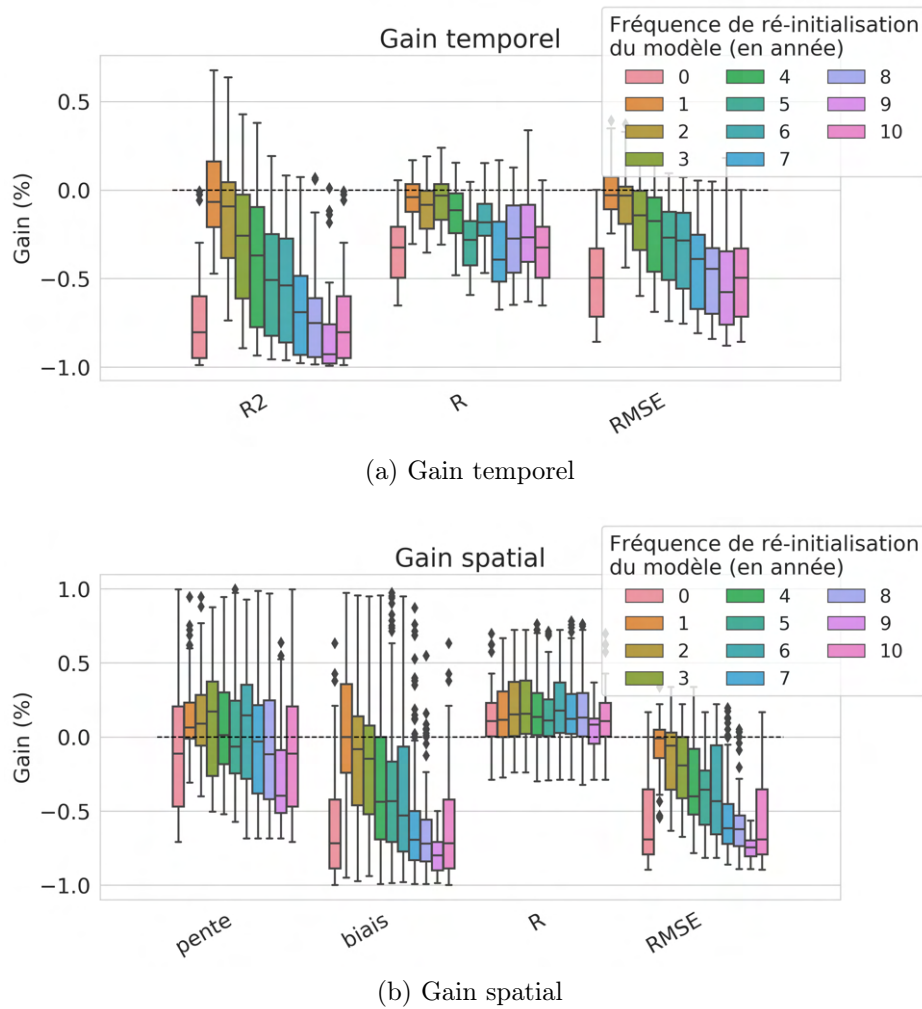


FIGURE 4.5 – Distribution des gains temporels et spatiaux de TWS pour le modèle dynamique RF initialisé par  $TWS_{GRACE}$ . Les gains sont calculés par rapport à GWS-OW et la référence LR GWS-LRref.

## 4.2.3 Limites et perspectives

### 4.2.3.1 Hétérogénéité des sources de données

Valider les produits GRACE désagrégés avec des données *in situ* est une tâche difficile, car une correspondance parfaite ne peut pas être attendue entre des sources de données si hétérogènes. Les cartes de GWS désagrégé proviennent de données satellitaires spatialisées acquises dans des configurations d'observation variées, alors que les données de validation sont des données locales acquises par des piézomètres irrégulièrement répartis sur la zone d'étude (figure 1.2). À basse résolution, la correspondance entre les données *in situ* et le GWS déconvolué est déjà limitée ( $R = 0.87$  et  $R^2 = 0.70$ , voir table 3.2).

Les données *in situ* ont leur propres incertitudes. D'abord, le GWS dérivé des mesures de hauteur d'eau sont fortement dépendantes de la valeur de  $S_y$  utilisée, qui varie à la fois horizontalement et verticalement (Dewandel et al., 2017a). La mesure de hauteur d'eau, mensuelle et ponctuelle, peut aussi être impactée par des effets à court ou long terme comme le pompage d'eau à proximité au moment de la mesure, qui ne peut être ni détecté ou rectifié avec une telle fréquence

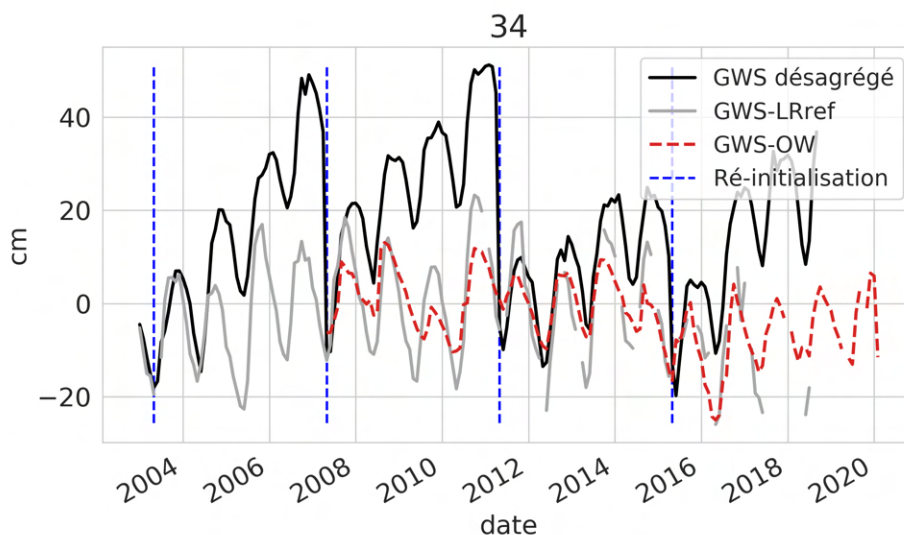


FIGURE 4.6 – Série temporelle du GWS désagrégé avec réinitialisation tous les 4 ans. Le GWS désagrégé est en noir, le GWS à basse résolution (GWS-ref) en gris, le GWS *in situ* (GSW-OW) en rouge, et les mois où la série est réinitialisée sont indiqués par des lignes verticales.

d'acquisition.

Le mode d'acquisition de GRACE diffère non seulement des mesures *in situ* instantanées, mais aussi des autres variables satellitaires qui sont des sommes ou moyennes de produits qui ont généralement une plus grande résolution spatiale et temporelle. La période de revisite de GRACE varie selon les pixels, et chaque pixel de  $300 \times 300 \text{ km}^2$  est informée par environ trois passages des satellites (Tapley et al., 2004a; Zaitchik et al., 2008). Aussi, les anomalies de GRACE sont issues d'un modèle global, et de très forts signaux peuvent parasiter d'autres zones du monde (Frappart et al., 2019; Ramillien et al., 2012). Enfin, l'échantillonnage exclusivement vertical du champ gravitationnel par les satellites GRACE nécessite un post-traitement affectant le signal.

#### 4.2.3.2 Désagrégation en milieu hétérogène

Les méthodes de désagrégation empiriques reposent sur l'hypothèse que les processus qui lient les prédicteurs et le TWS sont les mêmes à basse et haute résolution. Cependant, en moyennant les variables HR sur une zone plus grande, les dynamiques locales de ces variables sont lissées, et certaines caractéristiques HR de certains pixels ne se retrouvent pas à basse résolution. Ceci est illustré par les cycles annuels bimodaux des séries temporelles de SSM SMOS et de NDVI sur les régions fortement irriguées et unimodaux sur les régions plus naturelles à la résolution de 25 km (figure 3.9), tandis qu'à la résolution de  $113\,000 \text{ km}^2$  les cycles annuels de SSM SMOS et NDVI sont respectivement bimodal et unimodal (figure 3.6). Dans les régions irriguées comme le Telangana, les pratiques culturelles varient d'un pixel HR à l'autre, soit du fait de l'hétérogénéité géographique de la disponibilité de l'eau (gradient nord-sud dans les précipitations et présence ou non de rivières), ou d'autres facteurs anthropiques (pratiques d'irrigation, décisions locales de gestion de la ressource en eau). Mangiarotti et al. (2016) suggère dans une étude théorique qu'un signal basse résolution moyenné à partir d'un trop grand nombre de signaux HR hétérogènes asynchrones (les maxima des séries temporelles HR sont décalés dans le temps et n'ont pas lieu au même moment) peut difficilement être désagrégé spatialement. C'est typiquement le cas d'une région fortement irriguée telle que le Telangana où les comportements des signaux HR sont indépendants et non dictés

uniquement par les conditions climatiques. Dans ce cas, la modélisation du signal basse résolution est plus susceptible de capturer la dynamique de seuls quelques-uns des pixels HR (Mangiarotti et al., 2016).

#### 4.2.3.3 Perspectives

Les produits désagrégés sont généralement plus informatifs temporellement et spatialement que les données LR, comme l'indiquent les gains majoritairement positifs, mais ce n'est pas systématiquement le cas. Ceci pourrait être amélioré par le changement du contexte de l'étude. En effet, l'État du Telangana est une zone très hétérogène, au niveau de l'aquifère, de la disponibilité en eau (rivières, pluie), de l'occupation du sol et des pratiques culturelles. De plus, la zone d'étude que l'on désagrège pourrait être trop petite :  $113\,000\text{ km}^2$  correspond à une zone de  $336 \times 336\text{ km}^2$ , à la limite de la résolution de GRACE. Ainsi, des méthodes de désagrégation similaires pourraient produire de meilleurs résultats sur des zones plus grandes et moins complexes.

La méthode de validation présentée dans l'article en section 4.3 met en évidence l'hétérogénéité des performances des modèles statiques de désagrégations dans le temps. Ceci est dû notamment au changement des processus hydrologiques dominants selon les saisons, et à l'incapacité du modèle de rendre compte des hétérogénéités spatiales à HR. L'utilisation de modèles individuels pour chaque saison pourrait améliorer la performance de la désagrégation, mais réduit beaucoup le nombre de points pour l'apprentissage de ces modèles, déjà limité (174) dans le cadre de cette étude. Les premiers résultats de la désagrégation par ces modèles saisonniers sont mitigés, potentiellement du fait de la fréquence d'acquisition trop faible de GRACE (un mois) pour leur apprentissage.

## 4.3 Evaluating downscaling methods of GRACE data : a case study over a fractured crystalline aquifer in South India

Claire Pascal<sup>1</sup>, Sylvain Ferrant<sup>1</sup>, Adrien Selles<sup>2</sup>, Jean-Christophe Maréchal<sup>2</sup>, Abhilash Paswan<sup>3</sup>, Olivier Merlin<sup>1</sup>.

Publié dans *Hydrology and Earth System Sciences* (HESS).

---

1. Centre d'Étude Spatiale de la BIOSphère (CESBIO-UPS-CNRS-IRD-CNES-INRAE), 18 av. Ed. Belin, Toulouse CEDEX 9, 31401, France

2. Bureau de Recherches Géologiques et Minières (BRGM), Université de Montpellier, 1039 rue de Pinville, Montpellier, 34000, France

3. National Geophysical Research Institute, CSIR, Hyderabad, India



# Evaluating downscaling methods of GRACE (Gravity Recovery and Climate Experiment) data: a case study over a fractured crystalline aquifer in southern India

Claire Pascal<sup>1</sup>, Sylvain Ferrant<sup>1</sup>, Adrien Selles<sup>2</sup>, Jean-Christophe Maréchal<sup>2</sup>, Abhilash Paswan<sup>3</sup>, and Olivier Merlin<sup>1</sup>

<sup>1</sup>Centre d'Étude Spatiale de la BIOSphère, CESBIO-UPS-CNRS-IRD-CNES-INRAE,  
18 av. Ed. Belin, Toulouse CEDEX 9, 31401, France

<sup>2</sup>Bureau de Recherches Géologiques et Minières (BRGM), Université de Montpellier,  
1039 rue de Pinville, Montpellier, 34000, France

<sup>3</sup>National Geophysical Research Institute, CSIR, Hyderabad, India

**Correspondence:** Claire Pascal (claire.pascal@univ-tlse3.fr)

Received: 25 January 2022 – Discussion started: 8 March 2022

Revised: 18 July 2022 – Accepted: 21 July 2022 – Published: 10 August 2022

**Abstract.** GRACE (Gravity Recovery and Climate Experiment) and its follow-on mission have provided since 2002 monthly anomalies of total water storage (TWS), which are very relevant to assess the evolution of groundwater storage (GWS) at global and regional scales. However, the use of GRACE data for groundwater irrigation management is limited by their coarse ( $\simeq 300$  km) resolution. The last decade has thus seen numerous attempts to downscale GRACE data at higher – typically several tens of kilometres – resolution and to compare the downscaled GWS data with in situ measurements. Such comparison has been classically made in time, offering an estimate of the static performance of downscaling (classic validation). The point is that the performance of GWS downscaling methods may vary in time due to changes in the dominant hydrological processes through the seasons. To fill the gap, this study investigates the dynamic performance of GWS downscaling by developing a new metric for estimating the downscaling gain (new validation) against non-downscaled GWS. The new validation approach is tested over a 113 000 km<sup>2</sup> fractured granitic aquifer in southern India. GRACE TWS data are downscaled at 0.5° ( $\simeq 50$  km) resolution with a data-driven method based on random forest. The downscaling performance is evaluated by comparing the downscaled versus in situ GWS data over a total of 38 pixels at 0.5° resolution. The spatial mean of the temporal Pearson correlation coefficient ( $R$ ) and the root mean square error (RMSE) are 0.79 and 7.9 cm respec-

tively (classic validation). Confronting the downscaled results with the non-downscaling case indicates that the downscaling method allows a general improvement in terms of temporal agreement with in situ measurements ( $R = 0.76$  and  $RMSE = 8.2$  cm for the non-downscaling case). However, the downscaling gain (new validation) is not static. The mean downscaling gain in  $R$  is about +30 % or larger from August to March, including both the wet and dry (irrigated) agricultural seasons, and falls to about +10 % from April to July during a transition period including the driest months (April–May) and the beginning of monsoon (June–July). The new validation approach hence offers for the first time a standardized and comprehensive framework to interpret spatially and temporally the quality and uncertainty of the downscaled GRACE-derived GWS products, supporting future efforts in GRACE downscaling methods in various hydrological contexts.

## 1 Introduction

Groundwater is an essential resource for irrigation, especially in arid and semi-arid areas. Aquifers have suffered depletion in several areas of the world these last decades, and this resource is expected to be scarcer in the future (Wada et al., 2012). Monitoring and cautious management of this resource are therefore crucial. Groundwater monitoring is tradition-

ally achieved with networks of observation wells, but this can be challenging due to their sparse coverage, the punctual nature of the data, and the progressive abandonment of some wells or measurement difficulties and bias (Hora et al., 2019). In the meantime, new techniques for water storage monitoring have emerged with the Gravity Recovery and Climate Experiment (GRACE) satellite mission of US and German space agencies (NASA and DLR). The twin satellites of the GRACE mission were launched in 2002, and the continuity of the mission as covered by the GRACE Follow-On mission (GRACE-FO) launched in 2018. The gravimetric data retrieved from these missions have provided spatialized monthly anomalies of total water storage (TWS) for 2 decades, available worldwide. GRACE data were widely used in hydrology to study the long-term evolution of TWS or groundwater storage (GWS) by removing the contributions of other surface and sub-surface compartments from GRACE TWS at global and regional scales (Breña-Naranjo et al., 2014; Cao and Roy, 2020; Frappart et al., 2019; Papa et al., 2015; Rodell et al., 2018; Rzepecka and Birylo, 2020; Tiwari et al., 2009; Zhang et al., 2020). Nevertheless, their application at local scale for agricultural purposes remains limited due to the very low native resolution (about 400 km) of GRACE observations (Schmidt et al., 2008; Tapley et al., 2004).

During the past decade or so, several studies have proposed methods to downscale GRACE TWS data to obtain GWS maps at a spatial resolution (typically several tens of kilometres) higher than that of GRACE observations. Those downscaling approaches can be separated into two categories: model-based downscaling and data-based downscaling (also referred in the literature as “dynamic downscaling” and “statistical downscaling” respectively). The model-based downscaling approach consists in assimilating GRACE TWS data in physically based land surface or hydrological models to obtain GWS at the temporal and spatial resolutions of the model, which are generally higher than GRACE’s (Giroto et al., 2016; Houborg et al., 2012; Nie et al., 2019; Schumacher et al., 2018; Tian et al., 2017; Zaitchik et al., 2008). Yet this approach suffers from (i) the discrepancy between GRACE and model input data resolutions and (ii) the limitations inherent to models: model hypothesis and parametrization, the uncertainty of meteorological forcing, and particularly the lack of representation of anthropogenic processes such as crop irrigation (Long et al., 2013). The data-based downscaling approach consists in (i) deriving a statistical model of TWS from ancillary data available at high resolution (HR), (ii) calibrating it at low resolution (LR), (iii) applying it at HR, and (iv) removing the contribution of surface and soil moisture water stocks to isolate GWS. This data-driven approach rests on the hypothesis that the hydrological and physical processes that link those variables are identical at all resolutions (Ali et al., 2021; Jyolsna et al., 2021; Karunakalage et al., 2021; Sahour et al., 2020; Seyoum and Milewski, 2017; Vishwakarma et al., 2021; G. Zhang

et al., 2021; J. Zhang et al., 2021). In the literature, data-driven methods have been used to downscale GRACE data at various scales, either at the watershed scale for a thematic approach as in Seyoum and Milewski (2017) (5000 to 20 000 km<sup>2</sup>) or grid-based, with a downscaling resolution often limited by the coarsest resolution among the predictors (Ali et al., 2021; Jyolsna et al., 2021; Ning et al., 2014; Seyoum et al., 2019; G. Zhang et al., 2021; Zhong et al., 2021; Sahour et al., 2020).

To evaluate the GRACE data downscaled from the above approaches, different strategies have been used. Table 1 lists the validation methods used in recent papers downscaling GRACE with either model-based or data-based approaches. For both method categories, the validation of downscaled GWS mostly relies on the in situ measurements of groundwater levels (GWLs), converted or not into GWS anomalies using a specific yield (*Sy*) representative of the study area. Note that the GWS simulated by models has been occasionally used as a reference (Houborg et al., 2012; Seyoum and Milewski, 2017). In most studies, the quality of the downscaled GWS is evaluated by comparing its time series with that of GWL or GWS derived from in situ measurements for each HR unit (spatialized – HR pixel – or localized observation well) with one or several metrics, including the coefficient of determination ( $R^2$ ) or Nash–Sutcliffe efficiency coefficient (NSE), the Pearson correlation coefficient ( $R$ ), the root mean squared error (RMSE), or the mean absolute error (MAE) (Ali et al., 2021; Jyolsna et al., 2021; Karunakalage et al., 2021; Sahour et al., 2020; Yin et al., 2018; J. Zhang et al., 2021; G. Zhang et al., 2021; Zuo et al., 2021). In those studies, the downscaling procedure is considered efficient if those metrics fall within an acceptable range or if the downscaled product qualitatively restitutes the long-term trends of in situ data. The point is that any downscaling method can improve or degrade the accuracy of GRACE data at the targeted downscaling resolution depending on (i) the sub-pixel spatial variability of TWS/GWS and (ii) the uncertainties in input model parameters and forcing. Moreover, comparing the performances metrics with a “reference hypothesis” (here the “non-downscaled” case) allows us to quantitatively judge whether the downscaled product is better or worse in terms of accuracy at the targeted (fine) resolution and to evaluate whether the downscaling process is efficient. Therefore, quantifying the improvement against the GRACE data at their original resolution is crucial for properly evaluating downscaling methods. Among the 14 data-based methods listed in Table 1, only a few studies (Chen et al., 2019; Ning et al., 2014; J. Zhang et al., 2021; Zhong et al., 2021) quantify the improvement of the temporal agreement with in situ data of a downscaled product over the original LR data. Regarding the model-based approaches, all of them evaluate the temporal agreement of the downscaled GWS with in situ data against open-loop outputs (without the assimilation of GRACE data), but the results of the comparison against the LR GRACE TWS are not presented. Note that the primary



goal of the latter methods is to improve the model simulations using GRACE data and not specifically to downscale GRACE data, even though equivalence between both objectives may be argued.

For each downscaling method, Table 1 also indicates whether the evaluation of the downscaled dataset is undertaken in time or in both time and space. Zhong et al. (2021) are the only ones proposing a validation strategy combining the time and space dimensions by measuring the improvement of RMSE and  $R$  (with monthly in situ data) from LR to downscaled GWS using 42 observation wells within the GRACE pixel and for all months of the time series. This validation approach thus combines spatial and temporal evaluations but does not isolate their individual contributions. In particular, to the knowledge of the authors, none of the previous studies has specifically evaluated the capability of downscaled products to reconstitute the GWS spatial variations within the GRACE pixel at the temporal observation scale (1 month in our case).

Another issue in the application and validation of current downscaling studies is the scale at which the GRACE data are used at input. The combination of the ground tracks of the GRACE twin satellites over a period of 1 month allows a native spatial resolution of 300 to 400 km for GRACE data, both for spherical harmonic (Schmidt et al., 2008; Tapley et al., 2004) and mascon solutions from the Jet Propulsion Laboratory (JPL) (Watkins et al., 2015). The GRACE TWS grids are however provided with scaling factors with resolutions of 1 and  $0.5^\circ$  for the harmonic (Landerer and Swenson, 2012) and JPL mascon (Wiese et al., 2016) solutions respectively. Such scaling factors were originally designed to restore the lost signal of GRACE due to post-processing and to allow for averaging of the 1- or  $0.5^\circ$ -resolution oversampled TWS data over user-defined regions with a minimum extent similar to a 300–400 km-resolution GRACE pixel (Landerer and Swenson, 2012). In particular, scaling factors are not expected to efficiently downscale GRACE TWS data as neighbouring pixels are highly dependent (Landerer and Swenson, 2012). Yet many studies directly use GRACE harmonics solutions at  $1^\circ$  resolution (Ali et al., 2021; Jyolsna et al., 2021; Karunakalage et al., 2021; Ning et al., 2014; Seyoum et al., 2019; Yin et al., 2018; J. Zhang et al., 2021; G. Zhang et al., 2021; Zuo et al., 2021) or mascon solutions at  $0.5^\circ$  resolution (Karunakalage et al., 2021; Nie et al., 2019; Tian et al., 2017) as LR input data, which is far finer than their actual resolution. There is no study evaluating the uncertainty in downscaled GRACE data associated with the above assumption, i.e. neglecting the scale discrepancy between the actual resolution of GRACE observations and the grid size of the delivered oversampled GRACE data.

In this context, the objective of this study is to propose a consistent and complete validation framework covering the spatial and temporal aspects to quantify the supplementary information of downscaled GWS from GRACE compared with the LR original data. We test this framework on GRACE

data downscaled over a granitic aquifer of 113 000 km<sup>2</sup> in Telangana in southern India. We use a data-based approach to downscale GRACE mascon solution RL06M at a  $0.5^\circ$  resolution with two different models: a multilinear regression model and random forest. We also use this validation framework to evaluate the downscaling potential of the scaling factor at  $0.5^\circ$  resolution provided with the mascon solution (hence evaluating the choice of using the GRACE data oversampled at  $0.5^\circ$  resolution as a downscaled product). We compare the conclusions drawn from the classic validation techniques and the new validation framework proposed in this study.

## 2 Data and study area

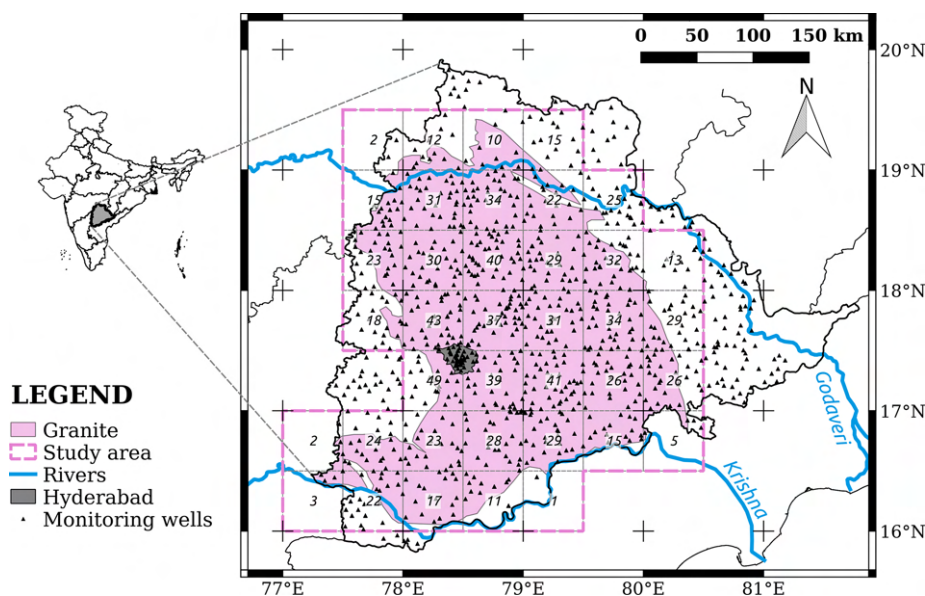
### 2.1 Study area

Telangana is a highly irrigated and densely populated (about 335 inhabitants per square kilometre in 2020 according to the Unique Identification Authority of India – UIDAI) region in southern India covering 114 800 km<sup>2</sup>. It is dominated by a semi-arid climate, where the monsoon precipitation occurs between July and October and ranges from 540 to 1300 mm with a mean of 879 mm (Indian Meteorological Department). The strong water demand in this area for domestic uses and the irrigation of two growing seasons a year is met with the surface water stored from monsoon rainfall and groundwater. The majority (67 000 km<sup>2</sup>) of the state is a shallow fractured granitic aquifer characterized by high fluctuations due to water pumping (Maréchal et al., 2006). It is usually composed of two layers: the first layer is saprolite, with a high effective porosity ( $S_y$  of 10 %), extending up to 10 to 15 m, and it is followed by a layer of fractured granite with a low capacity ( $S_y$  around 1 %) (Dewandel et al., 2017; Maréchal et al., 2006). This aquifer has a low capacity but strong dynamics as it fills and empties almost completely every year with monsoon rainfall and intense pumping. While continuous groundwater depletion has been observed with GRACE or observation wells in northern India (Asoka et al., 2017; Chen et al., 2014; Tiwari et al., 2009), northern China (Feng et al., 2013; Huang et al., 2015), Texas (Long et al., 2013), and many other parts of the world (Rodell et al., 2018), it is challenging to identify a long-term trend for groundwater storage in Telangana.

This study focuses on the granitic area of Telangana contoured with the  $0.5^\circ$ -resolution GRACE RL06M scaling factor grid. The study area is estimated at 113 000 km<sup>2</sup>, which is similar to the actual size of a GRACE pixel. Note that the GRACE RL06M pixels were extracted by selecting the  $0.5^\circ$  pixels falling within the granitic area of Telangana (pixels within the pink dotted line in Fig. 1).

**Table 1.** Validation strategies of existing – either data-based or model-based – downscaling methods of GRACE data. The downscaling method is either data-based (D) or model-based (M). The two resolutions reported are the initial resolution of GRACE data (GRACE) and the target downscaling resolution (Target). GWL: in situ groundwater level. GWS: in situ derived groundwater storage. The “Comp.” column indicates whether error statistics of the downscaled product are compared with those of another reference product: GRACE data at original low resolution (LR) or the model run in open loop (OL).

Reference	Downscaling		Resolution		Validation data		Validation metric			Validation in		Comp.
	method	method	GRACE Target	Resolution	in situ	model output	R	R <sup>2</sup>	RMSE	Trend changes	Time	
Ali et al. (2021)	D	1°	0.25°	0.25°	GWS	X	X	X		X		
Chen et al. (2019)	D	1°	0.25°	0.25°	GWL	X				X		LR
Jyotsna et al. (2021)	D	1°	0.25°	0.25°	GWS	X		X		X		
Karunakalage et al. (2021)	D	1 and 0.5°	0.25°	0.25°	GWL			X		X		
Ning et al. (2014)	D	1°	0.25°	0.25°	GWL	X				X		LR
Sahour et al. (2020)	D	13 700 to 59 200 km <sup>2</sup>	0.125°	0.125°	GWL	X				X		
Seyoum et al. (2019)	D	1°	0.25°	0.25°	GWS	X	X			X		
Seyoum and Milewski (2017)	D	500 000 km <sup>2</sup>	5000 to 20 000 km <sup>2</sup>	0.5°	GWS	X		X		X		
Vishwakarma et al. (2021)	D	62 518 to 4 672 876 km <sup>2</sup>	0.5°	0.5°	GWS	X				X		
Yin et al. (2018)	D	1°	2 km	2 km	GWL	X				X		
G. Zhang et al. (2021)	D	1°	0.25°	0.25°	GWL	X	X			X		LR
J. Zhang et al. (2021)	D	1 and 0.25°	1 km	1 km	GWL	X	X			X		LR
Zhong et al. (2021)	D	3°	5 km	5 km	GWS	X	X			X		LR
Zuo et al. (2021)	D	1°	1 km	1 km	GWL	X	X			X		
Giroto et al. (2016)	M	1°	36 km	36 km	GWS	X		X		X		OL
Houborg et al. (2012)	M	basin	4000 km <sup>2</sup>	4000 km <sup>2</sup>	GWS	X		X		X		OL
Nie et al. (2019)	M	0.125°	0.125°	0.125°	GWS	X		X		X		OL
Schumacher et al. (2018)	M	1 060 000 km <sup>2</sup>	0.5°	0.5°	GWS	X		X	X	X		OL
Tian et al. (2017)	M	0.5°	0.5°	0.5°	GWL	X				X		OL
Zaitchik et al. (2008)	M	> 500 000 km <sup>2</sup>	4000 km <sup>2</sup>	4000 km <sup>2</sup>	GWS	X		X		X		OL



**Figure 1.** Location of the study area (dotted pink line, 113 000 km<sup>2</sup>) that delineates the granitic area of Telangana (pink area, 67 000 km<sup>2</sup>, identified from Phani, 2014) with the target 0.5° resolution. The number of available observation wells (black triangles) monitored by the Groundwater Department of Telangana is indicated in the centre of each of the 38 0.5° pixels. The grey area indicates the extent of Hyderabad, the capital city of the state. The main rivers are indicated in blue.

## 2.2 Data

All data used and their sources are summarized in Table 2. Figure 2 shows time series of some of the data presented below as well as their intra-annual and interannual periodicities.

### 2.2.1 GRACE TWS

We used the state-of-the-art GRACE mascon solution from JPL (RL06M) with the Coastal Resolution Improvement (CRI) filter in this study. The mascon solution uses a priori information derived from near-global geophysical models to prevent striping. Moreover, it suffers less from leakage errors than the harmonic solution (Watkins et al., 2015) (<https://grace.jpl.nasa.gov/data/choosing-a-solution/>, last access: 19 January 2022). Each mascon is 3° × 3°, and the scaling factor grid used to restore the lost signal has a 0.5° resolution. After multiplication of the mascon grid by the scaling factor grid, all the 0.5°-resolution pixels within the study area are spatially averaged over the study area to produce a LR TWS time series at 113 000 km<sup>2</sup> scale. The baseline of TWS anomalies was modified by subtracting the long-term mean of the 2007–2015 period.

### 2.2.2 Ancillary data

We use ancillary variables from three different datasets to predict GRACE TWS: the monthly rainfall from the TRMM mission at a 0.25° resolution, the normalized difference vegetation index (NDVI) from MODIS at 1 km and the remotely sensed surface soil moisture data from the ESA CCI product

(combining passive microwave-derived soil moisture products) at 0.25°. All these datasets provide monthly data except the CCI soil moisture (SM CCI) product, which was temporally aggregated at a monthly scale. The temporal window aggregation varies for GRACE TWS and is not always the same as those of ancillary data, but we assumed that the effects of slightly varying windows were negligible. All datasets were aggregated with bilinear resampling both at the downscaling target resolution (0.5°) and at regional scale (113 000 km<sup>2</sup>). The values were converted into anomalies by subtracting the long-term mean of the 2007–2015 period.

### 2.2.3 Deconvolution of GRACE TWS with GLEAM

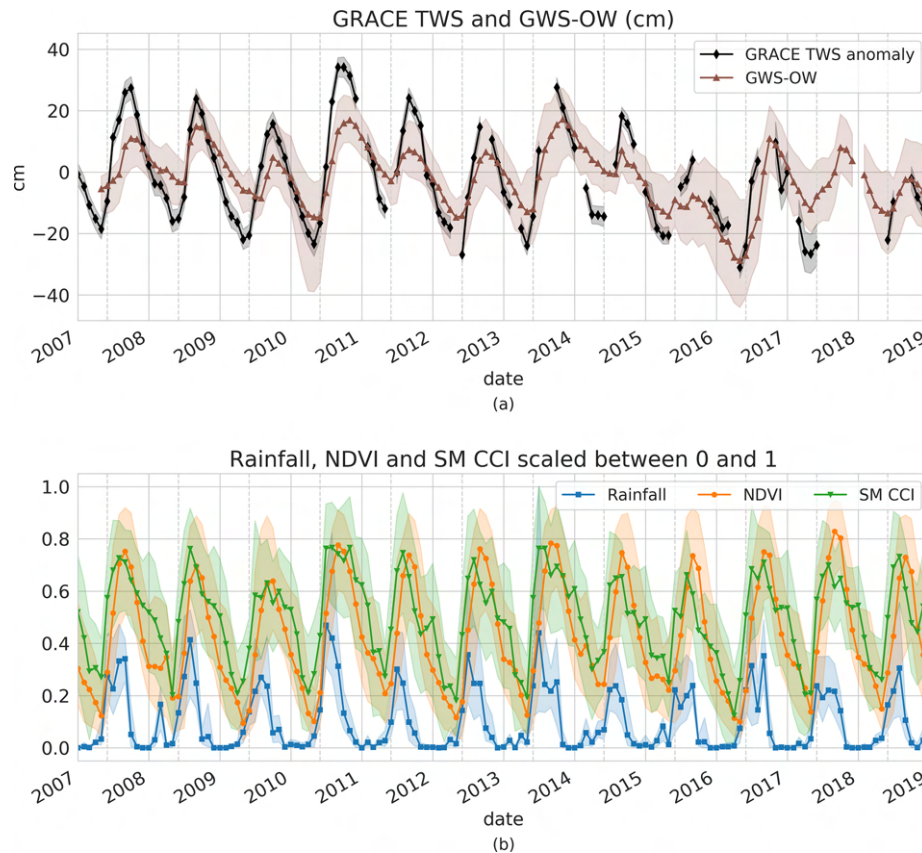
GWS is a sub-compartment of TWS, and hence the down-scaled GRACE TWS is not directly comparable with in situ derived GWS. In semi-arid areas, a common assumption is generally made that the essential contributions to TWS are GWS and soil moisture (SM) storage, thus neglecting canopy, snow and surface water storage (Eq. 1):

$$\Delta\text{TWS} = \Delta\text{GWS} + \Delta\text{SM}, \quad (1)$$

with  $\Delta$  representing the anomalies regarding a baseline, the 2007–2015 average in our case. In Telangana, the rivers (except the major rivers Godavari and Krishna; see Fig. 1) are not perennial and only flow for a few months during and after the monsoon. Surface water stocks are composed of large dams built on major rivers, with a cumulative capacity estimated at 113 mm (Indian National Register of Large Dams) and small reservoirs in the upstream part, with a

**Table 2.** Summary of all data used.

Variable	Source	Native spatial resolution	Usage
TWS	GRACE RL06M	3°	Target variable
Rainfall	TRMM 3B43 V7	0.25°	Predictor
NDVI	MOD13A3v006	1 km	Predictor
Surface soil moisture	ESA CCI v06.1 passive product	0.25°	Predictor
RZSM	GLEAM v3.5	0.25°	Deconvolution of GRACE TWS
GWS	Telangana State Groundwater Board	Punctual data	Validation



**Figure 2.** Time series of low-resolution (a) GRACE TWS with its uncertainty envelope (average of the mascon uncertainty resampled at 0.5° provided with GRACE data) and GWS anomalies derived from in situ measurements (GWS-OW) in centimetres and (b) rainfall, NDVI and SM CCI. Those three predictors were scaled between 0 and 1 to compare their temporal cycles. The envelope for GRACE TWS is the uncertainty provided with the mascon solution, and those for GWS-OW, rainfall, NDVI and SM CCI correspond to the lowest and highest values found in high-resolution (0.5°) pixels at each time step. The months of June (beginning of the monsoon) are marked by dotted vertical lines.

capacity estimated at 30 mm in a previous study (Pascal et al., 2021). This potential reservoir of 143 mm represents 24 % of GRACE TWS annual fluctuation in this area during the 2002–2021 time period (600 mm), yet the reservoirs are rarely simultaneously full and, most of the time, surface water storage can be neglected. Most studies use model outputs of SM to deconvolute GRACE TWS (Ali et al., 2021; Chen et al., 2019; Jyolsna et al., 2021; Karunakalage et al., 2021; Yin et al., 2018; J. Zhang et al., 2021; Zuo

et al., 2021; Sahour et al., 2020; Seyoum and Milewski, 2017; Zhong et al., 2021). We used the Global Land Evaporation Amsterdam Model (GLEAM) v3.5b monthly root zone soil moisture (RZSM) dataset to simulate SM storage, which we transform into anomalies to the baseline 2007–2015. GLEAM v3.5b is a model driven by satellite data that estimates evapotranspiration and soil moisture over a 0.25°-resolution grid for the period 2003–2020 (Martens et al.,

2017; Miralles et al., 2011). RZSM anomalies were computed by retrieving the 2007–2015 mean.

### 2.2.4 Validation GWS data (GWS-OW)

We use GWL data from the Groundwater Department of Telangana (India Water Resources Information System, <https://indiawris.gov.in/wris>, last access: 19 January 2022) for the period 2007–2019. These data provide monthly surveys of instantaneous GWL of 1006 wells distributed over the study area (see Fig. 1). Maps of GWL at 0.5° were produced from the interpolation of well data with the inverse distance weighting (IDW) method (which avoids kriging bias and provides more accurate values on data points) and were converted into a GWL anomaly by retrieving the long-term mean of the 2007–2015 period. These maps were converted to GWS maps by multiplying by a  $S_y$  that was calibrated with a linear fitting between GRACE TWS deconvoluted with GLEAM RZSM and the GWL anomaly at regional scale. The  $S_y$  was estimated at 4.7 %, which is an intermediate and consistent value between the  $S_y$  of both layers (saprolite at 10 % and fractured granite at 1 %) composing the aquifer in the study region. In the following, we designate these computed GWS anomalies as GWS-OW.

## 3 Downscaling and validation methods

This section details the validation method developed in this study (Sect. 3.1) that consists of a validation against an LR reference in both spatial and temporal aspects. This framework is tested on state-of-the-art statistical downscaling methods that are detailed in Sect. 3.2.

### 3.1 Evaluation of downscaled data

#### 3.1.1 Gain against the “null hypothesis”

As highlighted in the introduction, a lack in the majority of publications on GRACE downscaling is the comparison of the downscaled GWS with a null hypothesis. In particular, current evaluation methods check whether metrics fall within an acceptable range that is qualitatively defined but do not quantify the improvement provided by the downscaling process from a reference hypothesis. To fill the gap of current validation strategies of the downscaling methods of GRACE data, new metrics are proposed herein to quantitatively assess the accuracy of the downscaled data compared with the data at the original GRACE resolution (null hypothesis). In this case, two LR TWS references are possible: either the spatially averaged TWS value (produced as explained in Sect. 2.2.1) or the product of the mascon solution and its scaling factor grid at 0.5° resolution. In both cases, the contribution of SM to TWS is removed (using GLEAM RZSM estimates used at the 0.5° target resolution) to obtain GWS, comparable with in situ data. We chose to use the aver-

aged TWS deconvoluted with the 0.5° GLEAM RZSM (further called GWS-LRref) as the LR reference. The 0.5°-scale factor-based product (further called SF) is used as the downscaling first guess whose performance will be compared with the downscaling techniques proposed in this paper.

We chose to compute a relative gain similarly to Merlin et al. (2015). For a given metric  $M$  measuring the agreement with the validation data (e.g.  $R$ , RMSE), the gain  $G$  is computed as follows (Eq. 2):

$$G = \frac{|M_{\text{opt}} - M_{\text{LR}}| - |M_{\text{opt}} - M_{\text{HR}}|}{|M_{\text{opt}} - M_{\text{LR}}| + |M_{\text{opt}} - M_{\text{HR}}|}, \quad (2)$$

with  $M_{\text{LR}}$  the value of the metric for the GWS-LRref,  $M_{\text{HR}}$  its value for the downscaled GWS, and  $M_{\text{opt}}$  the optimal value of this metric (e.g. 1 for  $R$ , 0 for RMSE). The gain of Eq. (2) can be computed in time or in space.

#### 3.1.2 Temporal gain at high spatial resolution

For the temporal analysis, we compute this gain in the time series of GWS on all HR pixels and for three metrics:  $R$ ,  $R^2$ , and RMSE (Eqs. 3–5). These are temporal gains, as they measure the improvement of the agreement of the time series on each HR (0.5°) pixel where in situ measurements are available.

$$G_R = \frac{|1 - R_{\text{LR}}| - |1 - R_{\text{HR}}|}{|1 - R_{\text{LR}}| + |1 - R_{\text{HR}}|} \quad (3)$$

$$G_{R^2} = \frac{|1 - R_{\text{LR}}^2| - |1 - R_{\text{HR}}^2|}{|1 - R_{\text{LR}}^2| + |1 - R_{\text{HR}}^2|} \quad (4)$$

$$G_{\text{RMSE}} = \frac{\text{RMSE}_{\text{LR}} - \text{RMSE}_{\text{HR}}}{\text{RMSE}_{\text{LR}} + \text{RMSE}_{\text{HR}}} \quad (5)$$

#### 3.1.3 Spatial gain at monthly scale

For the spatial analysis, we compare the monthly maps of downscaled GWS with reference maps of GWS-OW. For each time step, we compute a gain over the LR reference on four metrics: the slope  $S$  of the linear regression (Eq. 6), the mean bias  $B$  (Eq. 7),  $R$  (Eq. 3), and RMSE (Eq. 5).

$$G_S = \frac{|1 - S_{\text{LR}}| - |1 - S_{\text{HR}}|}{|1 - S_{\text{LR}}| + |1 - S_{\text{HR}}|} \quad (6)$$

$$G_B = \frac{|B_{\text{LR}}| - |B_{\text{HR}}|}{|B_{\text{LR}}| + |B_{\text{HR}}|} \quad (7)$$

We expect  $S$  and  $R$  to be closer to 1 and  $B$  and RMSE closer to 0 for the downscaled product than for the LR reference. The slope is a common indicator to evaluate downscaled products, in particular for soil moisture downscaling (Merlin et al., 2015; Sabaghy et al., 2020). Indeed, the variability of GWS is expected to be higher at HR and closer to that of in situ measurements than at LR. Computing metrics for each time step rather than on the whole time series (all time steps and all HR pixels mixed) allows us for the first

time to eliminate the contributions of intra-annual and inter-annual variations and to specifically isolate the contribution of GWS spatial variability in the GRACE sampling period.

### 3.2 Statistical downscaling method

We use a data-based downscaling method that consists in training a model at LR between TWS and ancillary variables resampled at LR (113 000 km<sup>2</sup>). This model and ancillary variables at HR (0.5°) are then used to predict TWS at 0.5°. An additive correction is applied at LR to force the average of HR TWS predicted by the model to be equal to the TWS observed at LR (GRACE observation). The corrected TWS at 0.5° is finally deconvoluted into GWS with the GLEAM RZSM. We compare the two models often used in the literature: the multilinear regression model and the random forest model. The downscaling process is summarized in the flowchart of Fig. 3.

#### 3.2.1 Variable selection

For this data-driven approach, we selected remote sensing predictors that have a hydrological meaning. Also, we avoided model outputs, as irrigation is often not well represented in models. The predictors considered herein are precipitation (TRMM), surface SM (CCI), NDVI (MODIS) as an indicator of crop fraction, and the monthly variation of NDVI ( $\Delta$ NDVI). We also used as predictors the cumulative sum over the past year for all variables (except  $\Delta$ NDVI) by considering that it provides information about the state of the aquifer before the start of the irrigation season. Note that some predictors lag behind GRACE TWS due to the time that hydrological processes take. We determined the optimal time lag between TWS and each variable from 0 to 3 months by maximizing their temporal correlation coefficients (Sahour et al., 2020; Seyoum and Milewski, 2017). For both multilinear regression and random forest approaches, parsimonious models are obtained by selecting the optimal number of the most meaningful variables that allow prediction of the TWS. We used the RFECV (recursive feature elimination with cross-validation) algorithm, which is a greedy feature elimination algorithm similar to sequential backward selection.

#### 3.2.2 Multilinear regression model

The multilinear (ML) regression model fits a linear relationship between the target variable  $Y$  (here TWS) and  $p$  predictors  $X_1, X_2, \dots, X_p$  (Eq. 8):

$$Y = \beta_0 + \beta_1 X_1 + \dots + \beta_p X_p + E. \quad (8)$$

The  $\beta_0, \beta_1, \dots, \beta_p$  are determined by minimizing the mean squared error between the data and the model predictions. This model has the advantage of being easily interpretable but is limited by the assumptions that relationships between

variables are linear. Before training the ML model, the issue of multicollinearity (the existence of linear relationships between variables) was addressed. The elimination of redundant variables increases the precision of the coefficients of the regression and helps to properly identify the contribution of the remaining variables to the target variable (here TWS), and especially the signs of the coefficients. We used the variance inflation factor (VIF) (Alin, 2010) as in Sahour et al. (2020) to detect multicollinearity, and predictors with  $VIF > 10$  were removed.

#### 3.2.3 Random forest regressor

The random forest (RF) algorithm (Breiman, 2001) is a supervised ensemble learning algorithm composed of independent decision trees. Each decision tree learns with a subset of the predictors (here the square root of the maximum number of predictors) using a bootstrap sampling. This method softens the relationship constraints between variables but loses in interpretability. There is no need to remove some variables before training the model as the RF algorithm deals well with collinearity.

#### 3.2.4 Additive correction

After predicting HR TWS with the ML or RF model, we corrected the TWS values so that the spatial average of HR TWS at each time step would be equal to LR TWS. We add an offset value to correct the HR TWS at each month of the time series that corresponds to the difference between LR GRACE TWS and the spatial average of HR TWS predictions at the same date (Eq. 9):

$$TWS_{HR,t,i}^{corr} = TWS_{HR,t,i} + TWS_{LR,t} - \frac{\sum_i TWS_{HR,t,i}}{n_{pix}}, \quad (9)$$

with  $TWS_{HR,t,i}$  the HR TWS predicted by the model for month  $t$  and pixel  $i$ ,  $TWS_{HR,t,i}^{corr}$  the bias-corrected TWS,  $TWS_{LR,t}$  the LR TWS at date  $t$ , and  $\frac{\sum_i TWS_{HR,t,i}}{n_{pix}}$  the spatial average of HR TWS at date  $t$ .

## 4 Results

This section aims at evaluating the efficiency of the two data-based downscaling methods, i.e. ML and RF models against GWS-OW. In each case, we compare these results with the first-guess downscaling product, i.e. the product of the mascon solution and its scaling factors at 0.5° resolution (SF). After commenting on the results of the model calibration at LR in Sect. 4.1, we analyse the conclusions drawn from the classic evaluation methods found in the literature (Sect. 4.2) and then from the new validation method proposed in this study (Sect. 4.3). The synthesis of the different conclusions is presented in Sect. 4.4.

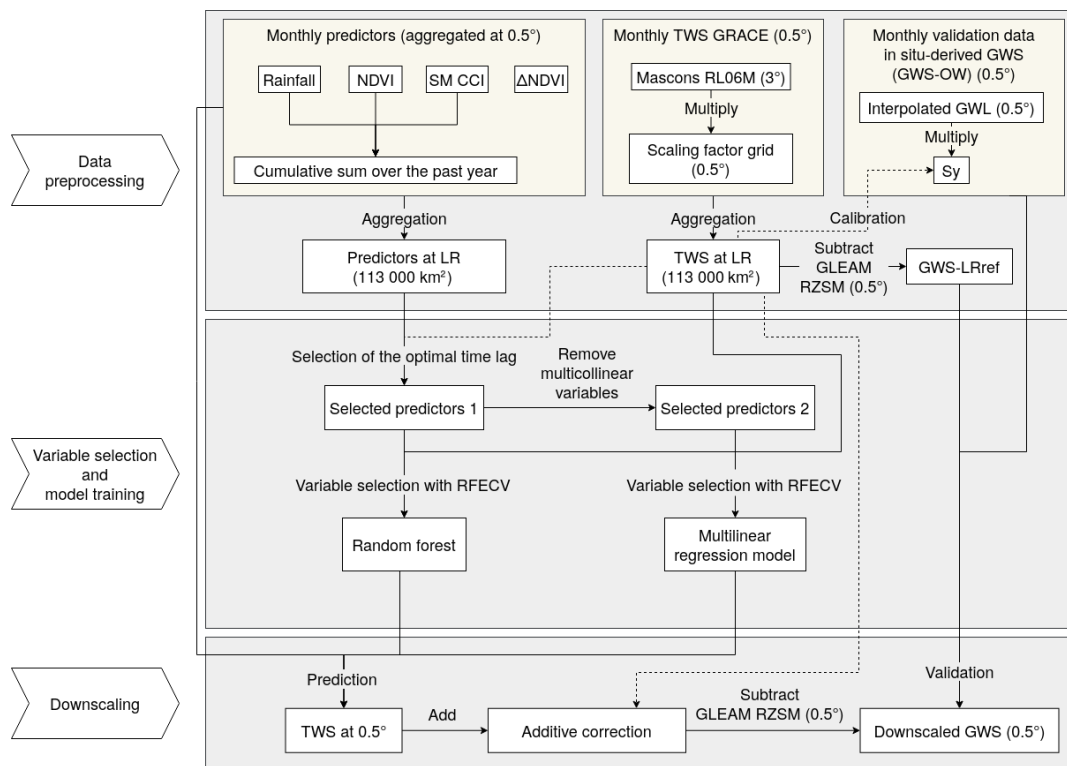


Figure 3. Flowchart of the downscaling method.

Table 3. Correlation coefficients of ancillary variables with GRACE TWS. The optimal time lag is indicated by bold correlations. The underlined correlations are not statistically significant.

Lag	Rainfall	NDVI	SM CCI	ΔNDVI
0	<b>0.90</b>	0.30	0.79	<u>0.10</u>
-1	0.85	0.68	<b>0.91</b>	0.51
-2	0.54	<b>0.79</b>	0.71	<b>0.68</b>
-3	0.16	0.68	0.34	0.63

#### 4.1 Variable selection and model calibration at LR

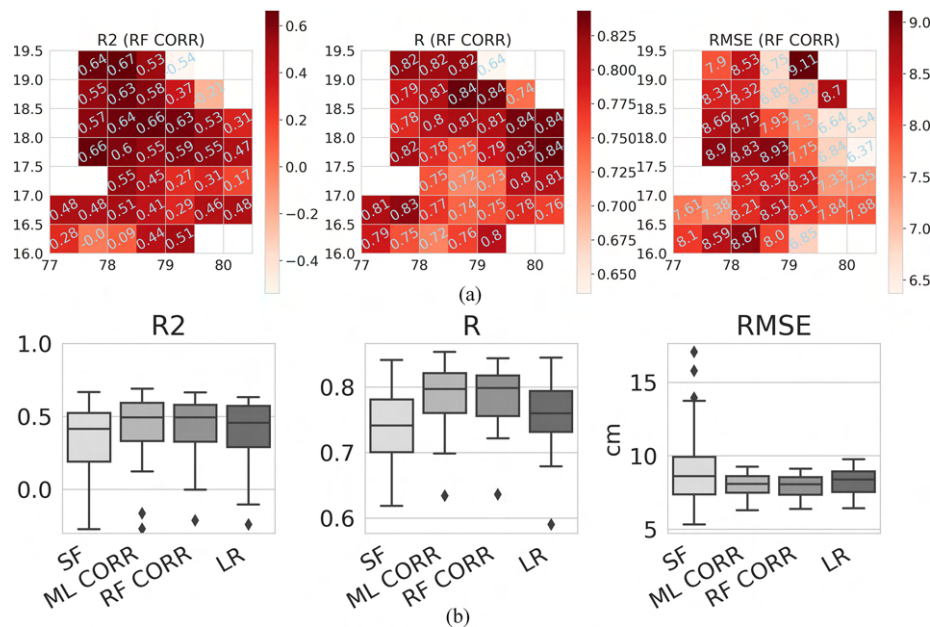
The correlation coefficients of the ancillary variables with GRACE TWS at LR are reported in Table 3. The bold correlations indicate which time lag was chosen for each variable: no lag for NDVI, 1-month lag for CCI soil moisture, 2-month lag for ΔNDVI and rainfall. The selected variables for each model are indicated in Table 4. Four variables were selected for the ML model: ΔNDVI, NDVI, SM CCI, and SM CCI accumulated over a year. The RF model selected two additional variables: monthly rainfall and rainfall accumulated over a year.

ML and RF models are trained on a random sample of 80 % of the whole time series (174 points in total). The selected variables and the model performances are reported in Table 4. The RF model has a better  $R^2$  than the ML model

(0.97 against 0.90), yet the RMSE and  $R^2$  on the test set are far larger (lower) than on the train set (4.6 cm against 1.9 cm and 0.93 against 0.98). This reveals that the RF model suffers from overfitting due to the quality of the data and the small amount of data (139 points) used to train the model, resulting in poor generalization. The RMSE on the train set is respectively 5.0 and 4.6 cm for the ML model, which represents 7 % and 6 % of the GRACE TWS total amplitude over the region during the study period (71 cm). Both models seem to be able to predict GRACE TWS with good performance. However, the performance is lower when compared with in situ data. As an example, the  $R^2$  between in situ derived TWS (sum of GWS-OW and RZSM GLEAM) aggregated at LR and GRACE TWS is 0.80. This shows that only limited agreement can be expected between satellite data (or modelled from satellite data) and in situ data, because of (i) the inherent uncertainties of the data, (ii) the interpolation of in situ data, and more generally (iii) the diversity of data sources. All those uncertainty sources also apply to the TWS predicted by models at both low and high resolutions. The  $R^2$  with in situ derived TWS falls from 0.90 and 0.97 to 0.78 and 0.82 for the ML and RF predictions respectively. This can be due to the existing uncertainty mentioned earlier but also to the possible lack of representativeness of in situ measurements at the GRACE spatial resolution.

**Table 4.** Variable selection and model performance at LR. The number in parentheses is the number of lag months. Variables with the suffix “cum.” are cumulated over the last year. The underlined variables were eliminated with the VIF selection method. The model performance is evaluated against GRACE TWS with the  $R^2$  and the RMSE on train and test sets. The  $R^2$  with in situ derived TWS (sum of GWS-OW and GLEAM RZSM at LR) is also shown.

Model	Variable selection							Model performances					
	Rainfall (2)	NDVI (0)	SM CCI (1)	$\Delta$ NDVI (2)	Rainfall cum.	NDVI cum.	SM CCI cum.	RMSE train (cm)	RMSE test (cm)	$R^2$ train	$R^2$ test	$R^2$ with GRACE TWS (train + test)	$R^2$ with in situ derived TWS (train + test)
ML		X	X	X			X	5.2	5.0	0.89	0.91	0.90	0.78
RF	X	X	X	X	X		X	1.9	4.6	0.98	0.93	0.97	0.82



**Figure 4.** (a) Spatial distribution of  $R^2$ ,  $R$ , and RMSE for the downscaling with the random forest model with bias correction (RF CORR). The numerical value of the metric is indicated in the grid. The abscissa is the east longitude and the ordinate the north latitude. (b) Boxplot (median and quartiles) of  $R^2$ ,  $R$ , and RMSE between GWS-OW and the scaling factor product (SF), linear (ML), and random forest (RF) model-downscaled products with bias correction (CORR) and the low-resolution reference GWS-LRref (LR). The RMSE is an equivalent water height in centimetres.

## 4.2 Classic evaluation

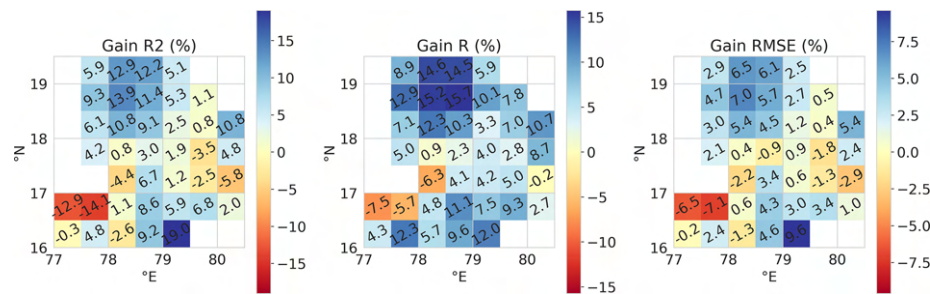
The temporal agreement between GWS-OW and downscaled products was evaluated on every HR pixel with  $R^2$ ,  $R$ , and RMSE for the SF downscaling and both the ML and RF models with correction by the LR offset value (CORR). Figure 4a shows the spatial distribution of these three metrics for the RF CORR-downscaled GWS for visualization and Fig. 4b the distribution of the three metrics on all pixels for all the downscaling methods. The temporal agreement of the SF product with GWS-OW seems to be the worst given the wide distribution of  $R^2$  with an average of 0.21 and some outliers in negative values and an average RMSE of 9.1 cm. The SF product appears to perform less well than the LR reference GWS-LRref (average  $R^2$ ,  $R$ , and RMSE

of 0.38, 0.76, and 8.2 cm). The  $R$  and  $R^2$  are better on average for ML CORR (0.79 and 0.42) and RF CORR (0.79 and 0.43), and the reduced variability of  $R$  and RMSE for ML CORR and RF CORR suggests that the bias correction produces results with a more uniform quality. The RMSE is still relatively large, ranging from 6.3 to 9.3 cm (6.4 to 9.3 cm) for the ML (RF) model. As a reference, the amplitude of GWS-OW in this area during the 2007–2019 time period ranges from 33 to 70 cm on all 38 HR pixels.

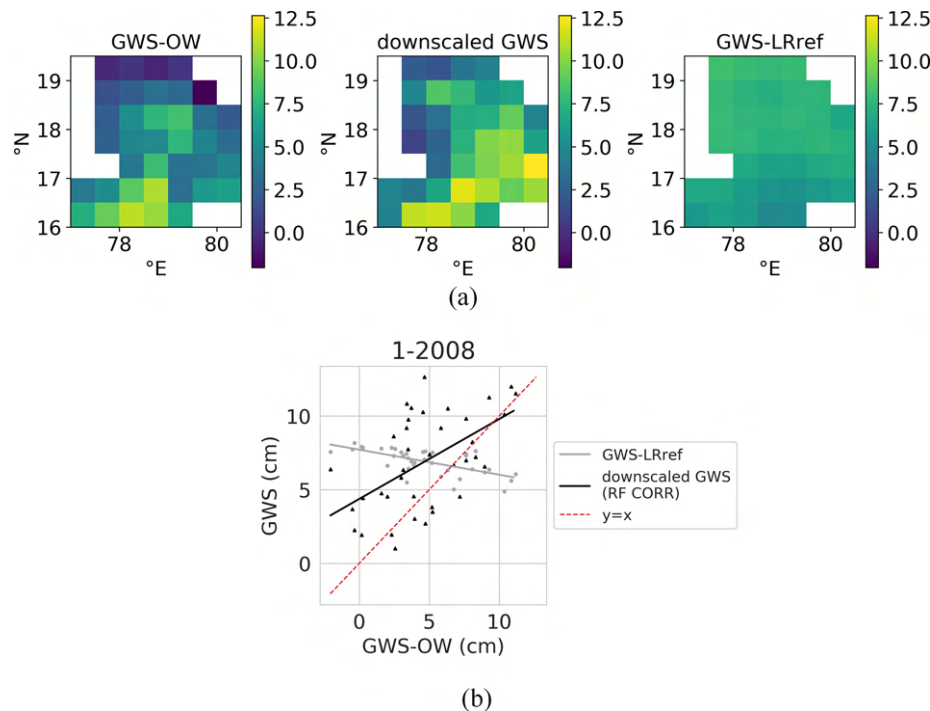
## 4.3 Evaluation with temporal and spatial gains

The temporal gains are computed as explained in Sect. 3.1 and are shown for the particular case of RF CORR in Fig. 5 for visualization. The spatial gains are computed at each





**Figure 5.** Spatial distribution of the gains of  $R^2$ ,  $R$ , and RMSE for the downscaling with RF CORR. The numerical value of the metric is indicated on the grid. The abscissa is the east longitude and the ordinate the north latitude.

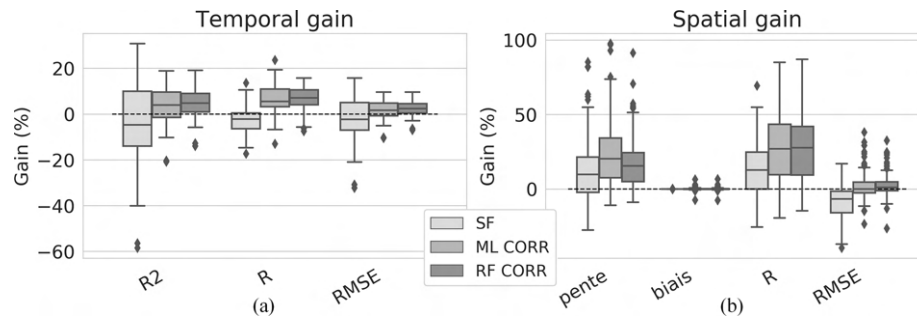


**Figure 6.** Illustration of the spatial gain for the month of January 2008. (a) Maps of HR in situ derived GWS (GWS-OW), RF CORR-downscaled GWS, and LR reference GWS (GWS-LRref). The abscissa is the east longitude and the ordinate the north latitude. (b) Scatterplot of the GWS-LRref (grey points) and RF-downscaled GWS (black points) against GWS-OW. The identity function is indicated by a dashed red line. The slope of the two linear regression fits on grey and black points are used to compute the gain in the slope. The differences in dispersion, uncertainty, and bias of the two point clouds are evaluated with gains in  $R$ , RMSE, and mean bias.

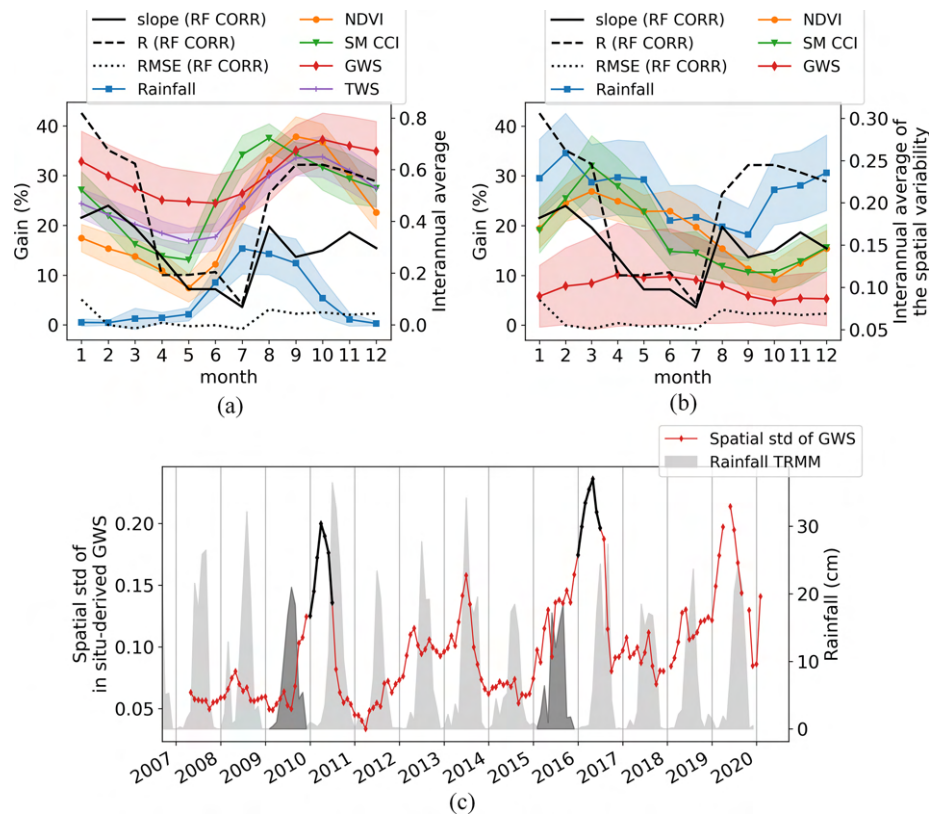
time step between the two point clouds of GWS-OW and GWS-LRref or downscaled GWS, as illustrated in Fig. 6. The boxplots of temporal gains in all HR pixels and the boxplot of spatial gains in the whole time series are shown in Fig. 7. ML CORR and RF CORR show the best results: average gains in  $R^2$ ,  $R$ , and RMSE are respectively 3.2%, 6.5%, and 1.55% for ML CORR and 4.0%, 6.7%, and 1.9% for RF CORR. In particular, the temporal gains for the RF CORR product seem to be positive in the north and south of the study area (cf. Fig. 5), which coincides with the two main river basins of the state but also concerns pixels with mixed geology and where the least number of obser-

vation wells are available (see Fig. 1). The pixel at 17° N, 78° E contains the major part of the capital city of the state, Hyderabad, a heavily urbanized area where natural hydrological processes as well as observation well measurements are highly perturbed by domestic water use, explaining the negative gains in  $R^2$ ,  $R$ , and RMSE (−4.4%, −6.3%, and −2.2% respectively).

In the spatial domain (see Fig. 7b), the quality of the SF-downscaled GWS is questionable. The quasi-null gain in bias was expected, as the only difference between LR and SF TWS is a multiplicative factor generally close to 1. The SF-downscaled GWS shows positive gains for slope and



**Figure 7.** Boxplots (median and quartiles) of (a) temporal and (b) spatial gains for the SF, ML and RF model downscaling approaches. The CORR designation indicates that the downscaled TWS is corrected for the LR bias from GRACE data.



**Figure 8.** Monthly medians of spatial gains (black curves) on slope,  $R$ , and RMSE for downscaled GWS with the RF CORR model (left axis) with, on the right axis, (a) average  $\pm$  standard deviation (SD) of low-resolution rainfall, NDVI, SM from the CCI dataset, and in situ derived GWS and TWS scaled between 0 and 1 and (b) interannual average  $\pm$  SD of the monthly spatial variability divided by the grid maximum of rainfall, NDVI, SM from the CCI dataset, and in situ derived GWS, scaled between 0 and 1. (c) Time series of the monthly spatial variability of in situ derived GWS (red curve) with monthly cumulative rainfall (grey). The abnormally dry monsoons of 2009 and 2015 and the high GWS spatial variability during the following dry season are highlighted in black.

$R$  on most of the time series (12.9 % and 13.9 % respectively on average) but at the cost of higher uncertainty ( $-8.5$  % gain in RMSE). Monthly scatterplots (results not shown) indicate that the slope getting closer to 1 is most of the time a consequence of an increased dispersion due to what appears to be additional noise at each time step brought by the scaling factor grid. The improvement of spatial representativity of GWS with data-based downscaling methods

(ML CORR and RF CORR) is shown by overall positive gains in slopes and  $R$  (22.9 % and 28.8 % for ML CORR, 18.2 % and 27.2 % for RF CORR) while maintaining a general positive gain in RMSE as well (2.0 % and 2.2 % for ML CORR and RF CORR respectively). The bias correction at LR adjusts the HR TWS predicted by the model to the GRACE TWS amplitude, explaining the null gain in bias for both models.

Complementary to the spatial analysis of temporal metrics (e.g. in Fig. 5), the spatial gains can be analysed in time. Figure 8 shows the monthly medians of the gains in slope,  $R$ , and RMSE for RF CORR downscaled GWS. It appears that gains in slope and  $R$  are lowest during the month of July (beginning of the monsoon). Gains in both slope and  $R$  increase until January–February (beginning of the second crop season that ends in April) and decrease again. The monthly gains of RMSE have lower amplitudes than those of slope and  $R$  but have a similar pattern. The periodicity in downscaling performances is due to the capacity or incapacity of the model trained at LR to reconstitute the spatial variability of some intermittent processes. In this paper, the tested downscaling methods are empirical, like the majority of existing methods (Ali et al., 2021; Jyolsna et al., 2021; Sahour et al., 2020; Seyoum and Milewski, 2017). Therefore, we are not able to represent explicitly the underlying hydrological processes that explain (in the downscaling procedure) the spatial variability of GWS at a given time. However, the performance of the downscaling methods essentially relies on their capability to represent implicitly the discharge and recharge of the aquifer at  $0.5^\circ$  resolution. This is the reason why the temporal variability in the downscaling performance can be interpreted in terms of taking into account the dominant hydrological processes and their seasonal dynamics.

In Telangana, the year can be divided into several periods given their dominant hydrological processes. The month of August marks the start of aquifer recharge by the rainfall that occurs 2 to 3 months after the beginning of the monsoon (which lasts generally from August to October; see Fig. 8a). It is also the beginning of the growing season (which typically lasts from July to November), when the monsoon rainfall stored at the surface and in the aquifer are used for irrigation. The higher spatial gains in slope,  $R$ , and RMSE during this period show that the recharge process in space is correctly represented with the precipitation data at  $0.5^\circ$  (having mainly a north–south gradient). The period between January and March, during the dry season, is marked by the heavy pumping and use of surface water for crop irrigation. This process is relatively well represented by the downscaling model from SM CCI and NDVI data, which provide indirect information on irrigation and crop stage respectively. During this period, the spatial variability of both predictors (illustrated by Fig. 8b, which represents the interannual average of the monthly spatial variability) is relatively large, accounting for the differences in crop fraction and type that highly depend on surface water availability. By April–May, irrigation stops and groundwater reaches its lowest level. The downscaling gains obtained at that time of year are relatively low. The model probably fails to reconstitute the diversity of HR GWS when the water availability and thus water exchanges are very scarce and hardly inferable from the chosen predictors. At the beginning of the monsoon in June–July, heavy rainfall occurs and fills rivers and reservoirs. However, at this early stage of the monsoon, rainfall has not reached the

aquifers and GWS remains low as in April and May. Also, surface water is an important component of the water column at this time of year (potentially up to 24 % of the annual fluctuation of TWS; see Sect. 2.2.3), yet runoff is not directly modelled by any of the variables of the RF model, which could also mislead the model into attributing surface water stocks to groundwater.

The use of a spatial gain also highlights the difficulties of state-of-the-art “static” downscaling methods (calibrated with constant parameters) to reconstitute an interannual variability. This is illustrated by Fig. 8b, which represents the interannual average (curve) and variability (envelope) of the monthly spatial variability of GWS. The interannual variability, which is lowest from August to January and highest from April to July for GWS, is inversely proportional to the downscaling performance. This result indicates that this kind of method is unable to represent the interannual variation of the dominant hydrological processes. Such a difference in interannual variability during the end of the dry season can be explained by the succession of drier and wetter periods dictated by El Niño and La Niña phenomena (Asoka et al., 2017; Vissa et al., 2019). This involves differences in yearly cumulative rainfall that determine the types of crops according to their water needs. During the driest years in particular, differences in water availability widen the gap between  $0.5^\circ$  regions, explaining higher spatial variabilities of GWS. This is illustrated in Fig. 8c by the abnormally high GWS spatial variability following the dry monsoons of 2009 and 2015.

#### 4.4 Comparative analysis of both validation methods

The thresholds to decide whether temporal metrics are poor, satisfactory, or good are often arbitrarily decided and are different with the context of the study and the authors’ choices. In our case, all downscaled GWS products have correlation coefficients with GWS-OW systematically larger than 0.57 on each of the 38 HR pixels, which can be considered a quite satisfactory result. With the  $R^2$  criteria, ML CORR and RF CORR seem to have the best performance, with at least half of the HR pixels having  $R^2$  larger than 0.5. For ML CORR and RF CORR, the RMSE does not go below 6.3 and 6.7 cm respectively, with a median RMSE of 8.1 cm for both methods, which still represent a non-negligible 18 % (16 %) error against GWS-OW (GWS-LRref) amplitude at LR.

Those above appreciations of temporal metrics do not indicate the superiority and the downscaling capacity of these downscaled GWS maps over GWS-LRref. With spatial and temporal gains, it was shown that ML CORR and RF CORR products are able to improve the temporal agreement with in situ data for most of the HR pixels. In the spatial aspect, ML CORR and RF CORR both improve the spatial representativity of GWS for most of the time series (positive gains in slope and  $R$ ), with a slightly lower uncertainty (gain in RMSE mostly positive). In addition to the compar-

ison against the LR reference, the validation in both time and space allows a better understanding of the downscaling strengths and flaws depending on local characteristics of some HR pixels (e.g. presence of rivers, agricultural practices, climatological variability, or large cities) or the time of year (e.g. wet or dry season). In the temporal domain, the RF CORR downscaled GWS seems to be better correlated with in situ data in the north and south of the study area near large rivers. This suggests that the model trained at LR has difficulties in modelling certain processes, e.g. the hydrological response to anthropic pressure that is localized and thus smoothed when averaged over a larger region. In the spatial domain, our validation method shows that the RF CORR downscaled GWS performed less during the dry season and the beginning of the wet season, supposedly because the model fails to represent the spatial variability of GWS when GWS is low or when surface water represents an important fraction of TWS. This highlights the weaknesses of a static model trained on a whole time series with no regards for the specificities of the hydrologically dominant processes at several times of the year.

## 5 Discussion

Here we discuss how downscaled products can be impacted by (i) the resolution at which GRACE data are used (Sect. 5.1) and (ii) the uncertainty issues when combining data from heterogeneous sources (Sect. 5.2).

### 5.1 Impact of the GRACE actual resolution on downscaled results

The validation framework proposed in this article was used to evaluate the downscaling potential of the scaling factor built from mascon solution RL06M. The scaling factor is built by fitting a unique factor between the TWS from the GLDAS CLM model at  $0.5^\circ$  and aggregated at mascon scale ( $3^\circ$ ) to evaluate the signal loss over the entire time series. Although it is not meant to downscale the mascon solution, several recent articles in the literature have used the oversampled TWS (spherical harmonics at  $1^\circ$  or mascon solution at  $0.5^\circ$ ) as the LR input data. Our validation framework clearly showed that the product of GRACE TWS mascons and their scaling factor grid (SF method) degrades the temporal agreement with in situ data and is noisy at a monthly timescale. Such results indicate that this product should not be used at the  $0.5^\circ$  resolution.

### 5.2 Other uncertainty sources in the validation exercise

Another point that should be highlighted is the difficulty in validating downscaled products with in situ data. Downscaled GWS is built from remote sensing data with various acquisition processes, while validation data are derived from water-level depth acquired by local piezometers with a het-

erogenous distribution in the study area (see Fig. 1). Each methodological step before a possible comparison between spatialized in situ and remote-sensing-derived GWS adds uncertainties at LR, illustrated by a low  $R^2$  (0.63) and high RMSE (6.1 cm) between GWS-OW and GWS-LRref.

The in situ data have their own uncertainties. First, the GWS derived from GWL measurements is highly dependent on the value of the  $S_y$  used. Here we used a horizontally and vertically homogeneous  $S_y$ , obtained with a linear adjustment between LR GWL and GWS-LRref. Some authors avoid this issue by directly comparing GRACE-downscaled GWS with GWL measurements (Karunakalage et al., 2021; Ning et al., 2014; Tian et al., 2019, 2017; Yin et al., 2018; J. Zhang et al., 2021; G. Zhang et al., 2021; Zuo et al., 2021). Another issue is that instantaneous GWL can be impacted by short- or long-term effects such as the neighbourhood pumping intensity at the moment of piezometer measurement, which cannot be detected or rectified with the monthly temporal frequency of acquisition.

The acquisition mode for GRACE is also different from in situ data – unique instantaneous measurement – and other remote sensing predictors – the average or sum of higher temporal frequency products. GRACE has a heterogeneous revisit frequency, where each  $300 \times 300 \text{ km}^2$  pixel is informed by approximately three overpasses of the GRACE satellites during the month (Tapley et al., 2004; Zaitchik et al., 2008). This can lead for example to smoothing of the GRACE anomaly by skipping extremes. Another issue in GRACE acquisition is the exclusively vertical sampling of the gravitational field that produces striping in the solution and requires post-processing that alters the signal.

## 6 Conclusions

To date, validation strategies for GRACE-derived downscaled products have rested essentially on the appreciation of temporal metrics or trends between downscaled products and localized in situ measurements. Yet such a validation approach is insufficient to fully assess the usefulness of the downscaling method as it suffers from a lack of (i) appropriate validation of the spatial distribution of the downscaled GRACE-derived GWS within the GRACE pixel and (ii) comparison with the results that would be obtained without downscaling (by directly using GRACE TWS at the fine scale). This article reviews the validation methods of existing downscaling methods of GRACE data, both model-based and data-based, and proposes a more extensive validation framework. In particular, a set of gains is used to evaluate the improvement of downscaled products against a low-resolution (LR) reference, including both temporal and, for the first time, spatial aspects. Such gains aim at fully determining the quality and uncertainty of downscaled GRACE-derived GWS products in a more comprehensive way.

The new validation framework is tested to evaluate the performance of two data-based downscaling approaches with multilinear (ML) and random forest (RF) models over a 113 000 km<sup>2</sup> fractured aquifer in southern India to the target resolution of 0.5°. The HR TWS predicted by each model is bias-corrected at each time step from the difference between the LR GRACE TWS and the average of HR TWS. We use GRACE TWS from the RL06M mascon solution for this study, which is multiplied by its 0.5° scaling factor grid and averaged over the study area to produce the LR reference series. A secondary objective of the paper is to also assess the downscaling potential of the scaling factor by considering the product of mascon and scaling factor (SF) at the 0.5° resolution to be a downscaled product. The comparison of the two data-based downscaling methods (bias-corrected ML and RF) with the LR reference shows an improvement in terms of correlation with in situ measurements. In the temporal domain, the spatial average gains in Pearson correlation coefficients (*R*) and root mean squared error (RMSE) are +6.5 % and +1.6 % (+6. % and +1.9 %) for ML (RF). In the spatial domain, the gains in *R* and RMSE are +28.8 % and +2 % (+27.2 % and +2.2 %) for ML (RF) respectively. The new validation method also confirms that the SF product cannot be used at 0.5° resolution. Although the average *R* in HR pixels is similar for all methods (0.74, 0.74, and 0.76 for SF, ML, and RF respectively), the SF product degrades both temporal and spatial accuracies at 0.5° resolution compared with the LR (without downscaling) case, showing that it cannot be used as a valid downscaling approach. The spatial analysis of temporal gains reveals a spatial heterogeneity in downscaling performances, which are particularly poor over urbanized areas. The spatial evaluation originally proposed in this study is also able to analyse the seasonality of the downscaling performance. The RF downscaling performance is lower (gains in *R* below +10 %) during the end of the dry season when GWS is at its lowest and at the beginning of the monsoon when surface flow, not included in the RF model, is a major process. In particular, the spatial validation presented in this study highlights, for the first time, the flaws of static GRACE downscaling methods in contexts where the dominant hydrological processes are not the same throughout the year (such as a highly irrigated semi-arid region with a wet season and a dry season as in the case of this study). This shows how complete and comprehensive validation approaches are an essential tool to interpret spatially and temporally the quality and uncertainty of the downscaled GRACE-derived GWS products and hence to better understand and improve downscaling models and their hypotheses in the future.

While the GRACE-FO mission provides continuity of spaceborne gravity change measurements, upcoming similar missions (MARVEL, Lemoine and Mandea, 2020; Lemoine et al., 2020; MAGIC, Massotti et al., 2021) plan to significantly improve the precision and quality of gravimetric estimates by proposing new configurations for the satellite con-

stellations. Nevertheless, the specified spatial resolution for those future data still undergo strong technical limitations. Therefore, the recourse to downscaling techniques will be, at least in the medium term, the only way to obtain TWS products at a finer scale useful for basin-scale water management.

*Data availability.* GRACE/GRACE-FO Mascon data are available at <http://grace.jpl.nasa.gov> (NASA, 2022a), soil moisture CCI data at <https://www.esa-soilmoisture-cci.org> (ESA, 2022), and MODIS data at <https://appears.earthdatacloud.nasa.gov> (NASA, 2022b). Water level data are not publicly available.

*Author contributions.* The conceptualization of this work and the methodology were developed by CP and OM. Experiments and analysis were done by CP and supervised by OM and SF. The manuscript was written by CP with contributions from OM, SF, AS, and JCM. Water-level data were provided by AP and cleansed by AS.

*Competing interests.* The contact author has declared that none of the authors has any competing interests.

*Disclaimer.* Publisher's note: Copernicus Publications remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

*Financial support.* This research was realized as part of a PhD thesis funded by a French national research ministry doctoral fellowship. The article processing charges for this open-access publication were covered by the Horizon 2020 ACCWA project (grant agreement no. 823965) in the context of the Marie Skłodowska-Curie research and innovation staff exchange (RISE) programme.

*Review statement.* This paper was edited by Zhongbo Yu and reviewed by two anonymous referees.

## References

- Ali, S., Liu, D., Fu, Q., Cheema, M. J. M., Pham, Q. B., Rahaman, M. M., Dang, T. D., and Anh, D. T.: Improving the Resolution of GRACE Data for Spatio-Temporal Groundwater Storage Assessment, *Remote Sens.*, 13, 3513, <https://doi.org/10.3390/rs13173513>, 2021.
- Alin, A.: Multicollinearity, *WIREs Comput. Stat.*, 2, 370–374, <https://doi.org/10.1002/wics.84>, 2010.
- Asoka, A., Gleeson, T., Wada, Y., and Mishra, V.: Relative contribution of monsoon precipitation and pumping to changes in groundwater storage in India, *Nat. Geosci.*, 10, 109–117, <https://doi.org/10.1038/ngeo2869>, 2017.
- Breiman, L.: Random Forests, *Mach. Learn.*, 45, 5–32, <https://doi.org/10.1023/A:1010933404324>, 2001.

- Breña-Naranjo, J. A., Kendall, A. D., and Hyndman, D. W.: Improved methods for satellite-based groundwater storage estimates: A decade of monitoring the high plains aquifer from space and ground observations, *Geophys. Res. Lett.*, 41, 6167–6173, <https://doi.org/10.1002/2014GL061213>, 2014.
- Cao, Y. and Roy, S. S.: Spatial patterns of seasonal level trends of groundwater in India during 2002–2016, *Weather*, 75, 123–128, <https://doi.org/10.1002/wea.3370>, 2020.
- Chen, J., Li, J., Zhang, Z., and Ni, S.: Long-term groundwater variations in Northwest India from satellite gravity measurements, *Global Planet. Change*, 116, 130–138, <https://doi.org/10.1016/j.gloplacha.2014.02.007>, 2014.
- Chen, L., He, Q., Liu, K., Li, J., and Jing, C.: Downscaling of GRACE-Derived Groundwater Storage Based on the Random Forest Model, *Remote Sens.*, 11, 2979, <https://doi.org/10.3390/rs11242979>, 2019.
- Dewandel, B., Caballero, Y., Perrin, J., Boisson, A., Dazin, F., Ferrant, S., Chandra, S., and Maréchal, J.-C.: A methodology for regionalizing 3-D effective porosity at watershed scale in crystalline aquifers, *Hydrol. Process.*, 31, 2277–2295, <https://doi.org/10.1002/hyp.11187>, 2017.
- ESA: Climate Change Initiative, <https://www.esa-soilmoisture-cci.org>, last access: 9 August 2022.
- Feng, W., Zhong, M., Lemoine, J.-M., Biancale, R., Hsu, H.-T., and Xia, J.: Evaluation of groundwater depletion in North China using the Gravity Recovery and Climate Experiment (GRACE) data and ground-based measurements, *Water Resour. Res.*, 49, 2110–2118, <https://doi.org/10.1002/wrcr.20192>, 2013.
- Frappart, F., Papa, F., Güntner, A., Tomasella, J., Pfeiffer, J., Ramilien, G., Emilio, T., Schietti, J., Seoane, L., da Silva Carvalho, J., Medeiros Moreira, D., Bonnet, M. P., and Seyler, F.: The spatio-temporal variability of groundwater storage in the Amazon River Basin, *Adv. Water Resour.*, 124, 41–52, <https://doi.org/10.1016/j.advwatres.2018.12.005>, 2019.
- Giroto, M., Lannoy, G. J. M. D., Reichle, R. H., and Rodell, M.: Assimilation of gridded terrestrial water storage observations from GRACE into a land surface model, *Water Resour. Res.*, 52, 4164–4183, <https://doi.org/10.1002/2015WR018417>, 2016.
- Hora, T., Srinivasan, V., and Basu, N. B.: The Groundwater Recovery Paradox in South India, *Geophys. Res. Lett.*, 46, 9602–9611, <https://doi.org/10.1029/2019GL083525>, 2019.
- Houborg, R., Rodell, M., Li, B., Reichle, R., and Zaitchik, B. F.: Drought indicators based on model-assimilated Gravity Recovery and Climate Experiment (GRACE) terrestrial water storage observations, *Water Resour. Res.*, 48, W07525, <https://doi.org/10.1029/2011WR011291>, 2012.
- Huang, Z., Pan, Y., Gong, H., Yeh, P. J.-F., Li, X., Zhou, D., and Zhao, W.: Subregional-scale groundwater depletion detected by GRACE for both shallow and deep aquifers in North China Plain, *Geophys. Res. Lett.*, 42, 1791–1799, <https://doi.org/10.1002/2014GL062498>, 2015.
- Jyolsna, P. J., Kambhammettu, B. V. N. P., and Gorugantula, S.: Application of random forest and multi-linear regression methods in downscaling GRACE derived groundwater storage changes, *Hydrolog. Sci. J.*, 66, 874–887, <https://doi.org/10.1080/02626667.2021.1896719>, 2021.
- Karunakalage, A., Sarkar, T., Kannaujiya, S., Chauhan, P., Pranjali, P., Taloor, A. K., and Kumar, S.: The appraisal of groundwater storage dwindling effect, by applying high resolution downscaling GRACE data in and around Mehsana district, Gujarat, India, *Groundwater Sustain. Dev.*, 13, 100559, <https://doi.org/10.1016/j.gsd.2021.100559>, 2021.
- Landerer, F. W. and Swenson, S. C.: Accuracy of scaled GRACE terrestrial water storage estimates, *Water Resour. Res.*, 48, W04531, <https://doi.org/10.1029/2011WR011453>, 2012.
- Lemoine, J.-M. and Manda, M.: The MARVEL gravity and reference frame mission proposal, in: EGU General Assembly Conference Abstracts, p. 13359, <https://ui.adsabs.harvard.edu/abs/2020EGUGA..2213359L> (last access: 8 August 2022), 2020.
- Lemoine, J. M., Meyssignac, B., Manda, M., Samain, E., Bourgoigne, S., Blazquez, A., Balmino, G., Louise, L., and Michaud, J.: MARVEL Mission Proposal: The Latest Update, 2020, in: AGU Fall Meeting Abstracts, G020-08, <https://ui.adsabs.harvard.edu/abs/2020AGUFMG020...08L> (last access: 8 August 2022), 2020.
- Long, D., Scanlon, B. R., Longuevergne, L., Sun, A. Y., Fernando, D. N., and Save, H.: GRACE satellite monitoring of large depletion in water storage in response to the 2011 drought in Texas, *Geophys. Res. Lett.*, 40, 3395–3401, <https://doi.org/10.1002/grl.50655>, 2013.
- Maréchal, J. C., Dewandel, B., Ahmed, S., Galeazzi, L., and Zaidi, F. K.: Combined estimation of specific yield and natural recharge in a semi-arid groundwater basin with irrigated agriculture, *J. Hydrol.*, 329, 281–293, <https://doi.org/10.1016/j.jhydrol.2006.02.022>, 2006.
- Martens, B., Miralles, D. G., Lievens, H., van der Schalie, R., de Jeu, R. A. M., Fernández-Prieto, D., Beck, H. E., Dorigo, W. A., and Verhoest, N. E. C.: GLEAM v3: satellite-based land evaporation and root-zone soil moisture, *Geosci. Model Dev.*, 10, 1903–1925, <https://doi.org/10.5194/gmd-10-1903-2017>, 2017.
- Massotti, L., Siemes, C., March, G., Haagmans, R., and Silvestrin, P.: Next Generation Gravity Mission Elements of the Mass Change and Geoscience International Constellation: From Orbit Selection to Instrument and Mission Design, *Remote Sens.*, 13, 3935–3966, <https://doi.org/10.3390/rs13193935>, 2021.
- Merlin, O., Malbêteau, Y., Nofri, Y., Bacon, S., Khabba, S. E.-R. S., and Jarlan, L.: Performance Metrics for Soil Moisture Downscaling Methods: Application to DISPATCH Data in Central Morocco, *Remote Sens.*, 7, 3783–3807, <https://doi.org/10.3390/rs70403783>, 2015.
- Miralles, D. G., Holmes, T. R. H., De Jeu, R. A. M., Gash, J. H., Meesters, A. G. C. A., and Dolman, A. J.: Global land-surface evaporation estimated from satellite-based observations, *Hydrology and Earth System Sciences*, 15, 453–469, <https://doi.org/10.5194/hess-15-453-2011>, 2011.
- NASA: Measuring Earth's Surface Mass and Water Changes, <http://grace.jpl.nasa.gov> (last access: 9 August 2022), 2022a.
- NASA: Welcome to AρρEEARS!, <https://appears.earthdatacloud.nasa.gov> (last access: 9 August 2022), 2022b.
- Nie, W., Zaitchik, B. F., Rodell, M., Kumar, S. V., Arsenault, K. R., Li, B., and Getirana, A.: Assimilating GRACE Into a Land Surface Model in the Presence of an Irrigation-Induced Groundwater Trend, *Water Resour. Res.*, 55, 11274–11294, <https://doi.org/10.1029/2019WR025363>, 2019.

- Ning, S., Ishidaira, H., and Wang, J.: Statistical Downscaling of Grace-Derived Terrestrial Water Storage Using Satellite and Gldas Products, *J. Jpn. Soc. Civ. Eng. Ser. B1*, 70, I\_133–I\_138, [https://doi.org/10.2208/jscjehe.70.I\\_133](https://doi.org/10.2208/jscjehe.70.I_133), 2014.
- Papa, F., Frappart, F., Malbeteau, Y., Shamsudduha, M., Vuruputur, V., Sekhar, M., Ramillien, G., Prigent, C., Aires, F., Pandey, R. K., Bala, S., and Calmant, S.: Satellite-derived surface and sub-surface water storage in the Ganges–Brahmaputra River Basin, *J. Hydrol.: Reg. Stud.*, 4, 15–35, <https://doi.org/10.1016/j.ejrh.2015.03.004>, 2015.
- Pascal, C., Ferrant, S., Selles, A., Maréchal, J.-C., Gascoin, S., and Merlin, O.: High-Resolution Mapping of Rainwater Harvesting System Capacity from Satellite Derived Products in South India, in: 2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS, July 2021, Brussels, 7011–7014, <https://doi.org/10.1109/IGARSS47720.2021.9553131>, 2021.
- Phani, R. C.: Mineral Resources of Telangana State, India: The Way Forward, *Int. J. Innov. Res. Sci. Eng. Tech.*, 3, 15450–15459, <https://doi.org/10.15680/IJRSET.2014.0308052>, 2014.
- Rodell, M., Famiglietti, J. S., Wiese, D. N., Reager, J. T., Beaudoing, H. K., Landerer, F. W., and Lo, M.-H.: Emerging trends in global freshwater availability, *Nature*, 557, 651–659, <https://doi.org/10.1038/s41586-018-0123-1>, 2018.
- Rzepecka, Z. and Birylo, M.: Groundwater Storage Changes Derived from GRACE and GLDAS on Smaller River Basins – A Case Study in Poland, *Geosciences*, 10, 124, <https://doi.org/10.3390/geosciences10040124>, 2020.
- Sabaghy, S., Walker, J. P., Renzullo, L. J., Akbar, R., Chan, S., Chaubell, J., Das, N., Dunbar, R. S., Entekhabi, D., Gevaert, A., Jackson, T. J., Loew, A., Merlin, O., Moghaddam, M., Peng, J., Peng, J., Piepmeier, J., Rüdiger, C., Stefan, V., Wu, X., Ye, N., and Yueh, S.: Comprehensive analysis of alternative downscaled soil moisture products, *Remote Sens. Environ.*, 239, 111586, <https://doi.org/10.1016/j.rse.2019.111586>, 2020.
- Sahour, H., Sultan, M., Vazifedan, M., Abdelmohsen, K., Karki, S., Yellich, J. A., Gebremichael, E., Alshehri, F., and Elbayoumi, T. M.: Statistical Applications to Downscale GRACE-Derived Terrestrial Water Storage Data and to Fill Temporal Gaps, *Remote Sens.*, 12, 533, <https://doi.org/10.3390/rs12030533>, 2020.
- Schmidt, R., Flechtner, F., Meyer, U., Neumayer, K.-H., Dahle, C., König, R., and Kusche, J.: Hydrological Signals Observed by the GRACE Satellites, *Surv. Geophys.*, 29, 319–334, <https://doi.org/10.1007/s10712-008-9033-3>, 2008.
- Schumacher, M., Forootan, E., van Dijk, A. I. J. M., Müller Schmied, H., Crosbie, R. S., Kusche, J., and Döll, P.: Improving drought simulations within the Murray-Darling Basin by combined calibration/assimilation of GRACE data into the WaterGAP Global Hydrology Model, *Remote Sens. Environ.*, 204, 212–228, <https://doi.org/10.1016/j.rse.2017.10.029>, 2018.
- Seyoum, W., Kwon, D., and Milewski, A.: Downscaling GRACE TWSA Data into High-Resolution Groundwater Level Anomaly Using Machine Learning-Based Models in a Glacial Aquifer System, *Remote Sens.*, 11, 824, <https://doi.org/10.3390/rs11070824>, 2019.
- Seyoum, W. M. and Milewski, A. M.: Improved methods for estimating local terrestrial water dynamics from GRACE in the Northern High Plains, *Adv. Water Resour.*, 110, 279–290, <https://doi.org/10.1016/j.advwatres.2017.10.021>, 2017.
- Tapley, B. D., Bettadpur, S., Watkins, M., and Reigber, C.: The gravity recovery and climate experiment: Mission overview and early results, *Geophys. Res. Lett.*, 31, L09607, <https://doi.org/10.1029/2004GL019920>, 2004.
- Tian, S., Tregoning, P., Renzullo, L. J., van Dijk, A. I. J. M., Walker, J. P., Pauwels, V. R. N., and Allgeyer, S.: Improved water balance component estimates through joint assimilation of GRACE water storage and SMOS soil moisture retrievals, *Water Resour. Res.*, 53, 1820–1840, <https://doi.org/10.1002/2016WR019641>, 2017.
- Tian, S., Renzullo, L. J., van Dijk, A. I. J. M., Tregoning, P., and Walker, J. P.: Global joint assimilation of GRACE and SMOS for improved estimation of root-zone soil moisture and vegetation response, *Hydrol. Earth Syst. Sci.*, 23, 1067–1081, <https://doi.org/10.5194/hess-23-1067-2019>.
- Tiwari, V. M., Wahr, J., and Swenson, S.: Dwindling groundwater resources in northern India, from satellite gravity observations, *Geophys. Res. Lett.*, 36, L18401, <https://doi.org/10.1029/2009GL039401>, 2009.
- Vishwakarma, B. D., Zhang, J., and Sneeuw, N.: Downscaling GRACE total water storage change using partial least squares regression, *Scient. Data*, 8, 95, <https://doi.org/10.1038/s41597-021-00862-6>, 2021.
- Vissa, N. K., Anandh, P. C., Behera, M. M., and Mishra, S.: ENSO-induced groundwater changes in India derived from GRACE and GLDAS, *J. Earth Syst. Sci.*, 128, 115, <https://doi.org/10.1007/s12040-019-1148-z>, 2019.
- Wada, Y., van Beek, L. P. H., and Bierkens, M. F. P.: Nonsustainable groundwater sustaining irrigation: A global assessment, *Water Resour. Res.*, 48, W00L06, <https://doi.org/10.1029/2011WR010562>, 2012.
- Watkins, M. M., Wiese, D. N., Yuan, D.-N., Boening, C., and Landerer, F. W.: Improved methods for observing Earth’s time variable mass distribution with GRACE using spherical cap mascons, *J. Geophys. Res.-Solid*, 120, 2648–2671, <https://doi.org/10.1002/2014JB011547>, 2015.
- Wiese, D. N., Landerer, F. W., and Watkins, M. M.: Quantifying and reducing leakage errors in the JPL RL05M GRACE mascon solution, *Water Resour. Res.*, 52, 7490–7502, <https://doi.org/10.1002/2016WR019344>, 2016.
- Yin, W., Hu, L., Zhang, M., Wang, J., and Han, S.-C.: Statistical Downscaling of GRACE-Derived Groundwater Storage Using ET Data in the North China Plain, *J. Geophys. Res.-Atmos.*, 123, 5973–5987, <https://doi.org/10.1029/2017JD027468>, 2018.
- Zaitchik, B. F., Rodell, M., and Reichle, R. H.: Assimilation of GRACE Terrestrial Water Storage Data into a Land Surface Model: Results for the Mississippi River Basin, *J. Hydrometeorol.*, 9, 535–548, <https://doi.org/10.1175/2007JHM951.1>, 2008.
- Zhang, G., Zheng, W., Yin, W., and Lei, W.: Improving the Resolution and Accuracy of Groundwater Level Anomalies Using the Machine Learning-Based Fusion Model in the North China Plain, *Sensors*, 21, 46, <https://doi.org/10.3390/s21010046>, 2021.
- Zhang, J., Liu, K., and Wang, M.: Seasonal and Interannual Variations in China’s Groundwater Based on GRACE Data and Multisource Hydrological Models, *Remote Sens.*, 12, 845, <https://doi.org/10.3390/rs12050845>, 2020.
- Zhang, J., Liu, K., and Wang, M.: Downscaling Groundwater Storage Data in China to a 1-km Resolution Using Machine Learning Methods, *Remote Sens.*, 13, 523, <https://doi.org/10.3390/rs13030523>, 2021.

Zhong, D., Wang, S., and Li, J.: A Self-Calibration Variance-Component Model for Spatial Downscaling of GRACE Observations Using Land Surface Model Outputs, *Water Resour. Res.*, 57, e2020WR028944, <https://doi.org/10.1029/2020WR028944>, 2021.

Zuo, J., Xu, J., Chen, Y., and Li, W.: Downscaling simulation of groundwater storage in the Tarim River basin in northwest China based on GRACE data, *Phys. Chem. Earth Pt. A/B/C*, 123, 103042, <https://doi.org/10.1016/j.pce.2021.103042>, 2021.



---

## 4.4 Ce qu'il faut retenir

- Les méthodes de désagrégation empiriques statiques avec un modèle *Random Forest* permettent de produire des cartes de GWS majoritairement meilleures que la basse résolution, mais pas systématiquement.
  - Une méthode de validation exhaustive est nécessaire pour correctement quantifier l'intérêt des produits désagrégés par rapport à la basse résolution, à la fois dans le temps et dans l'espace, pour comprendre et améliorer les modèles de désagrégation futurs.
  - Les méthodes de désagrégation statiques ont des performances hétérogènes dans le temps dans des contextes où les processus hydrologiques dominants changent au cours de l'année.
  - Le GWS désagrégé par des modèles dynamiques donne de moins bons résultats qu'à basse résolution.
  - La désagrégation fonctionne mal pour des zones très hétérogènes où les cycles annuels des prédicteurs sont déphasés.
  - Utiliser le TWS de GRACE rééchantillonné à l'échelle du facteur d'échelle plutôt qu'une valeur moyennée à la résolution augmente l'erreur spatiale et temporelle par rapport aux données *in situ*.
  - Seule une correspondance limitée peut être attendue entre les données *in situ* et des produits désagrégés à partir de données satellitaires.
-

# Chapitre 5

## Cartographie des volumes d'eau de surface avec Pléiades et Sentinel-2

### Sommaire

---

<b>5.1</b>	<b>Introduction</b>	<b>100</b>
<b>5.2</b>	<b>Quantification de la capacité du RHS et suivi temporel avec des images haute résolution</b>	<b>101</b>
5.2.1	État de l'art	101
5.2.2	Méthodes et application sur le RHS du Telangana	102
5.2.3	Résultats	107
<b>5.3</b>	<b>Monitoring the volume of South Indian Rainwater Harvesting System : sensitivity of the bathymetry retrieval to DEM features</b>	<b>110</b>
<b>5.4</b>	<b>Discussion et perspectives</b>	<b>136</b>
5.4.1	Précision du MNT	136
5.4.2	Précision des surfaces en eau détectées sur Sentinel-2	136
5.4.3	Délimitation des tanks	137
5.4.4	Futures applications	138
<b>5.5</b>	<b>Ce qu'il faut retenir</b>	<b>139</b>

---

## 5.1 Introduction

Beaucoup de régions du monde souffrent d'une pénurie d'eau croissante, à la fois pour l'irrigation des cultures et les usages domestiques en zones urbaines (Wada et al., 2011). Beaucoup d'infrastructures ont été développées pour gérer l'eau de pluie et des rivières et lutter contre ce stress hydrique croissant, comme des grands lacs de barrage ou de petits réservoirs. En Inde, l'augmentation du nombre de petits réservoirs (aussi appelés "tanks") est une stratégie adoptée par le gouvernement pour s'adapter au stress hydrique dans les régions en amont des rivières (Giordano and Villholth, 2007; Krol et al., 2011). Ces réservoirs stockent en surface les pluies de mousson et constituent une ressource d'eau directement accessible pour l'agriculture ou le bétail. Ces tanks ont également été promus pour favoriser la recharge (Glendenning et al., 2012), mais la perte associée à l'évaporation pourrait être importante (Boisson et al., 2014; Massuel et al., 2014; Perrin et al., 2012), et leur impact semble être minime par rapport à la recharge par la pluie et les flux de retour de l'irrigation (Brauns et al., 2022).

Ce système de récupération des eaux de pluie indien (RHS, *Rainwater Harvesting System*) est composé de plusieurs centaines de milliers de tanks répartis le long d'un réseau de drainage non pérenne. Le nombre de tanks fonctionnels en Inde est estimé entre 200 000 (Palanisami et al., 2010; Sharma, 2003) et 500 000 (Reddy, 2015). Dans l'État du Telangana, un tank typique est une dépression plate d'un kilomètre de long, avec une digue de 3 à 10 mètres de haut et de 20 mètres d'épaisseur, qui se vide presque entièrement pendant la saison sèche. En 2015 dans cet État, un programme coûteux de cinq ans, la mission Kakatiya, a initié la réhabilitation de 46 000 tanks afin d'en retirer le dépôt de sédiments et favoriser la recharge, ainsi que la construction de nouveaux tanks. Ceci montre l'importance socio-politique du RHS face aux besoins grandissants en eau, pourtant, les tanks du RHS demeurent très peu documentés. Contrairement aux grands lacs de barrage, ils ne font pas l'objet d'un inventaire exhaustif ni d'un suivi volumétrique du fait de leur grand nombre et de leur petite taille, et leur étude ne s'est faite jusqu'à présent qu'à l'échelle individuelle (Boisson et al., 2014; Massuel et al., 2014; Perrin et al., 2012). L'impact du RHS dans les études hydrologiques à large échelle est souvent négligé ou grossièrement estimé, alors qu'il pourrait avoir un impact significatif (Habets et al., 2018).

Le suivi volumétrique du RHS à large échelle et l'étude de son impact hydrologique régional est limité par l'absence de données satellitaires très haute résolution, mais ce verrou pourrait être levé par la future mission CO3D de cartographie de la topographie de la Terre (Lebègue et al., 2020) avec la mise à disposition de Modèles Numériques de Terrain (MNT) globaux très haute résolution avec un taux de revisite élevé. À ce jour, seuls des produits commerciaux fournissent de tels MNT, et à des échelles très locales. En particulier, des MNT à 2 m de résolution sur des emprises de  $20 \times 20 \text{ km}^2$  peuvent être fabriqués à partir d'acquisitions optiques stéréoscopiques Pléiades à 50 cm de résolution (Gleyzes et al., 2012).

Ce chapitre propose une application concrète des MNT haute résolution dérivés de Pléiades et des images optiques Sentinel-2 pour le suivi volumétrique du RHS sur quatre emprises de  $20 \times 20 \text{ km}^2$ . Une première partie détaille les méthodes utilisées pour quantifier le volume du RHS, et leur application pour (i) la quantification de la capacité maximale du RHS du Telangana (étude présentée lors de la conférence IGARSS 2021 (Pascal et al., 2021)) et (ii) le suivi volumétrique du RHS à une échelle de  $20 \times 20 \text{ km}^2$ . La seconde partie présente un article soumis à *Remote Sensing* (MDPI) qui démontre l'utilité des modèles numériques de terrain (MNT) très haute résolution et des séries temporelles Sentinel-2 pour le suivi du volume du RHS, et qui quantifie l'apport de ces MNT par rapport aux MNT globaux actuels dans ce type d'études.

## 5.2 Quantification de la capacité du RHS et suivi temporel avec des images haute résolution

### 5.2.1 État de l’art

#### 5.2.1.1 Estimation du volume d’eau de surface

L’estimation du volume d’eau dans les réservoirs repose généralement sur une combinaison de mesures altimétriques (e.g ICESAT, ENVISAT, Topex/Poseidon) et de surfaces en eau dérivées d’images satellites radar (TERRASAR-X, RADARSAT-2) ou optiques multispectrales (Landsat, Formosat) (Baup et al., 2014; Crétaux et al., 2011; Duan and Bastiaanssen, 2013; Frappart et al., 2006). Ces estimations se font généralement sur des lacs isolés plutôt qu’à l’échelle régionale. La multiplication des missions satellites pour la cartographie de la topographie terrestre a mis à disposition plusieurs MNT globaux comme ASTER (*Advanced Spaceborne Thermal Emission and Reflection Radiometer*) (Tachikawa et al., 2011) (30 m de résolution), TanDEM-X (Krieger et al., 2013) (12 m de résolution, produit commercial), ALOS (*Advanced Land Observing Satellite “DAICHI” World 3D-30m*) (Tadono et al., 2014) et le MNT de Copernicus à 30 m (*WorldDEM GLO-30*). De telles données ont encouragé le développement de méthodes d’estimation de volume à partir de la bathymétrie dérivée des MNT, combinée avec des données altimétriques (Bonnema and Hossain, 2017), de lidar (Li et al., 2019) ou de séries temporelles de surfaces en eau dérivées de Landsat (Avisse et al., 2017; Bhagwat et al., 2019; Bonnema and Hossain, 2017) ou de la mission plus récente Sentinel-1 (Vanthof and Kelly, 2019).

#### 5.2.1.2 Intérêt des MNT haute résolution à fort taux de revisite

Dans la plupart des cas, les dates d’acquisition des MNT ne sont pas choisies par l’utilisateur, et la présence d’eau ou de végétation rend inaccessible la bathymétrie des réservoirs. Pour compenser ces imperfections, la géométrie des réservoirs est souvent simplifiée par une forme trapézoïdale prédéfinie (Baup et al., 2014; Bonnema and Hossain, 2017; Magome et al., 2003), interpolée par continuité topographique (Liu et al., 2020) ou en modélisant la relation surface-volume (relation ou courbe hypsométrique) (Avisse et al., 2017; Baup et al., 2014; Bhagwat et al., 2019; Vanthof and Kelly, 2019).

Plusieurs études ont appliqué ces méthodes d’estimation de volume à échelle régionale sur de grands réservoirs (Bhagwat et al., 2019; Bonnema and Hossain, 2017; Duan and Bastiaanssen, 2013) et plus récemment sur des structures plus petites (Arendt et al., 2021; Vanthof and Kelly, 2019). En particulier, Vanthof and Kelly (2019) ont démontré l’intérêt des données satellitaires haute résolution pour extraire des relations hypsométriques sur les tanks du RHS en Inde du Sud dans l’État du Tamil Nadu, en utilisant le MNT TanDEM-X commercial à 12 m. Cette étude a estimé une fluctuation de 8.2 millions de mètres cubes (MCM, *million cubic meter*) d’eau de surface sur 559 tanks dans une région de 5647 km<sup>2</sup> avec une cartographie dynamique des surfaces en eau à partir de Sentinel-1. Cette étude souffre néanmoins (i) des incertitudes de détection d’eau inhérentes aux images radar Sentinel-1 et (ii) de la simplification de la géométrie des réservoirs par une relation puissance modélisant la relation hypsométrique pour compenser les incertitudes du MNT.

Cependant, à ce jour, les MNT globaux ont une résolution n’excédant pas 12 m (pour la solution commerciale de TanDEM-X) ou 30 m pour d’autres MNT globaux (ALOS, WorldDEM GLO-30). Ces résolutions complexifient leur utilisation pour l’estimation des volumes de tanks de quelques

hectares comme ceux du RHS (un hectare contient onze pixels de  $30 \times 30 \text{ m}^2$ ).

De récents MNT très haute résolution (2 m) sur demande comme ceux produits à partir des images Pléiades permettent de contrôler la date d’acquisition et donnent accès à la bathymétrie précise des réservoirs. Cependant, ce sont des produits commerciaux disponibles localement, chaque image Pléiades couvrant environ  $20 \times 20 \text{ km}^2$  (Gleyzes et al., 2012).

## 5.2.2 Méthodes et application sur le RHS du Telangana

Cette section détaille les données et méthodes utilisées pour la quantification du volume du RHS dans l’État du Telangana. Des cartes d’eau Sentinel-2 sont utilisées pour identifier les tanks actifs du RHS, puis une courbe hypsométrique est dérivée à partir des MNT Pléiades pour chaque tank. Ces courbes hypsométriques sont utilisées pour deux études : (i) l’estimation de la capacité volumétrique maximale du RHS à l’échelle du Telangana pour l’évaluation de son importance en tant que réservoir hydrologique, et (ii) le suivi temporel du volume du RHS à l’échelle d’une emprise Pléiades. Les étapes de la méthode sont détaillées en figure 5.1.

### 5.2.2.1 Zone d’étude et données

Cette étude est réalisée sur les emprises de quatre MNT Pléiades à 2 m de résolution (emprises (a), (b), (c) et (d) sur la figure 5.2) couvrant respectivement 444, 322, 463 et  $584 \text{ km}^2$ . Les MNT (a), (b) et (c) sont produits à partir d’acquisition de juin 2016 et le MNT (d) de juin 2019 (la section 2.3.4.2 détaille la production des MNT à partir des images stéréoscopiques). L’acquisition des MNT en saison sèche lorsque la plupart des tanks sont vides donne accès à leur bathymétrie. Des cartes d’eau sont produites à partir des séries temporelles d’images Sentinel-2 couvrant les quatre emprises sur une période temporelle proche des acquisitions Pléiades (entre 2016 et 2020, voir section 2.3.4.3 pour une description détaillée des cartes d’eau Sentinel-2), qui permettent la cartographie des surfaces en eau en temps quasi-réel (taux de revisite de cinq jours) à haute résolution (10 m).

### 5.2.2.2 Identification des tanks

En raison de leur petite taille, les tanks du RHS ne sont pas référencés dans les inventaires nationaux (NRLD - Indian National Register of Large Dams) ou globaux (GRanD (Lehner et al., 2011), dont les tanks identifiés sont en magenta en figure 5.2). Les inventaires locaux de ces petits réservoirs sont rares et leurs contours peuvent être imprécis (Vanthof and Kelly, 2019). En l’absence de ces inventaires, les tanks peuvent être identifiés à partir d’images satellitaires. L’algorithme automatique *Blue Spot detection* (Arendt et al., 2021; Nielsen et al., 2011) permet l’identification des dépressions naturelles à partir d’un MNT, mais elles ne sont pas garanties d’être des tanks actifs. Les tanks actifs peuvent aussi être identifiés par des méthodes de détection d’eau à partir d’images satellitaires : Vanthof and Kelly (2019) ont identifié les contours de 1503 tanks dans l’État indien du Tamil Nadu en combinant un inventaire local et la base de données *Global Surface Water* (Pekel et al., 2016) (surfaces bleues dans la figure 5.2).

Ici, les tanks du RHS sont identifiés à partir de cartes d’eau dérivées des images Sentinel-2, qui permettent d’identifier les contours des tanks actifs à haute résolution (10 m). Les cartes d’eau de la série temporelle sont superposées et fusionnées pour obtenir une carte d’étendue d’eau maximale (MWAE, *Maximum Water Area Extent*) (polygones rouges dans la figure 5.2). La carte est raffinée par inspection visuelle en retirant (i) les surfaces correspondant aux lacs de barrage et lits de

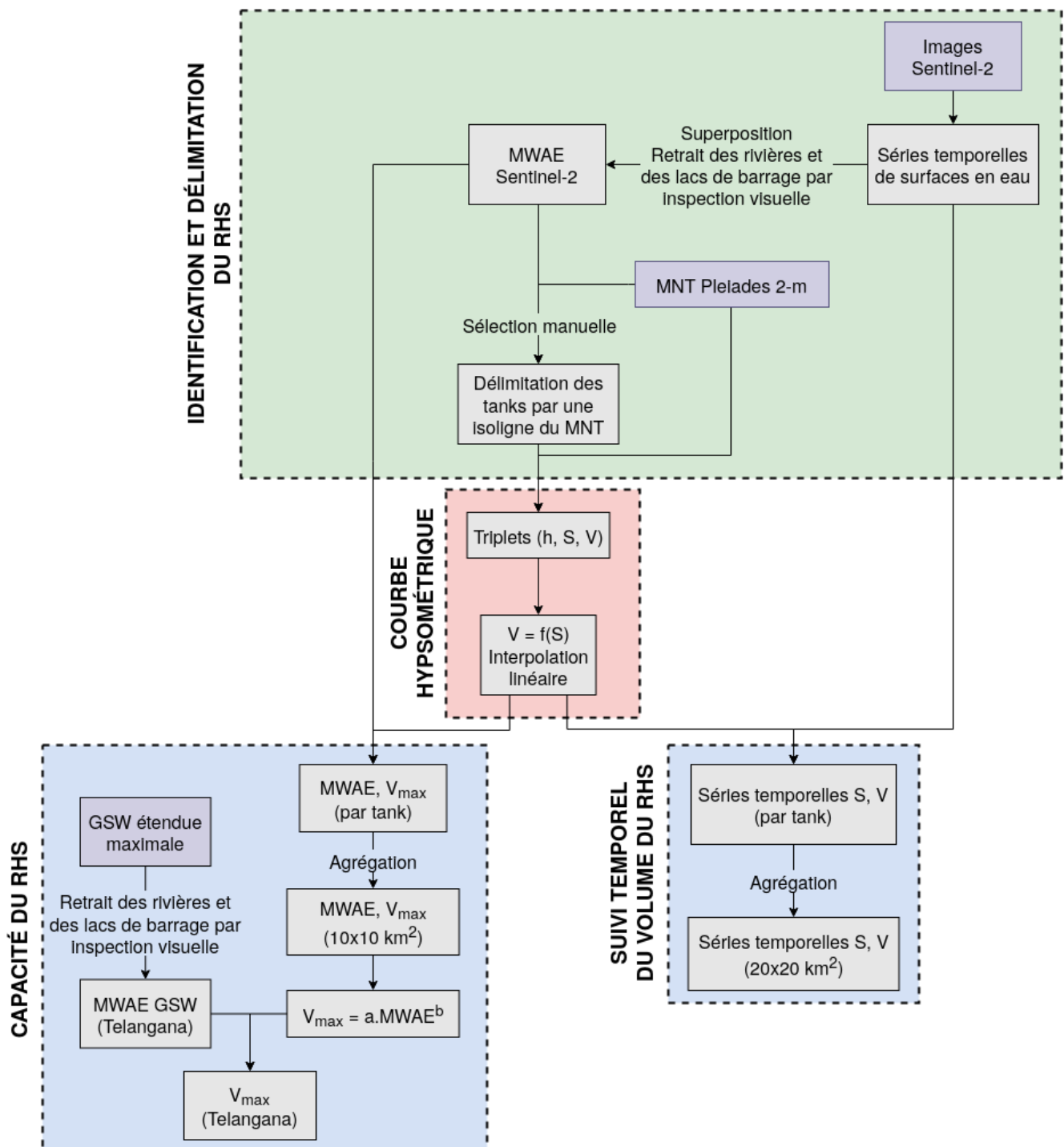


FIGURE 5.1 – Étapes de l'estimation du volume avec Pléiades et Sentinel-2. S : surface, V : volume, MWAE : surface maximale observée,  $V_{max}$  : capacité volumétrique, GSW : *Global Surface Water*.  $V = f(S)$  : relation hypsométrique.

rivières, (ii) les tanks contenant plus de 5% de données manquantes du MNT et (iii) les tanks avec un MWAE inférieur à 1000 m<sup>2</sup>, seuil arbitraire en dessous duquel ils ne sont pas pris en compte. 508 tanks ont été identifiés sur les quatre emprises Pléiades : 112 sur la scène (a), 91 sur la scène (b), 114 sur la scène (c) et 191 sur la scène (d). La majorité des tanks ont une taille inférieure à 5 ha (voir la distribution en figure 5.3) et leur taille peut atteindre 185 ha. Le MWAE des tanks

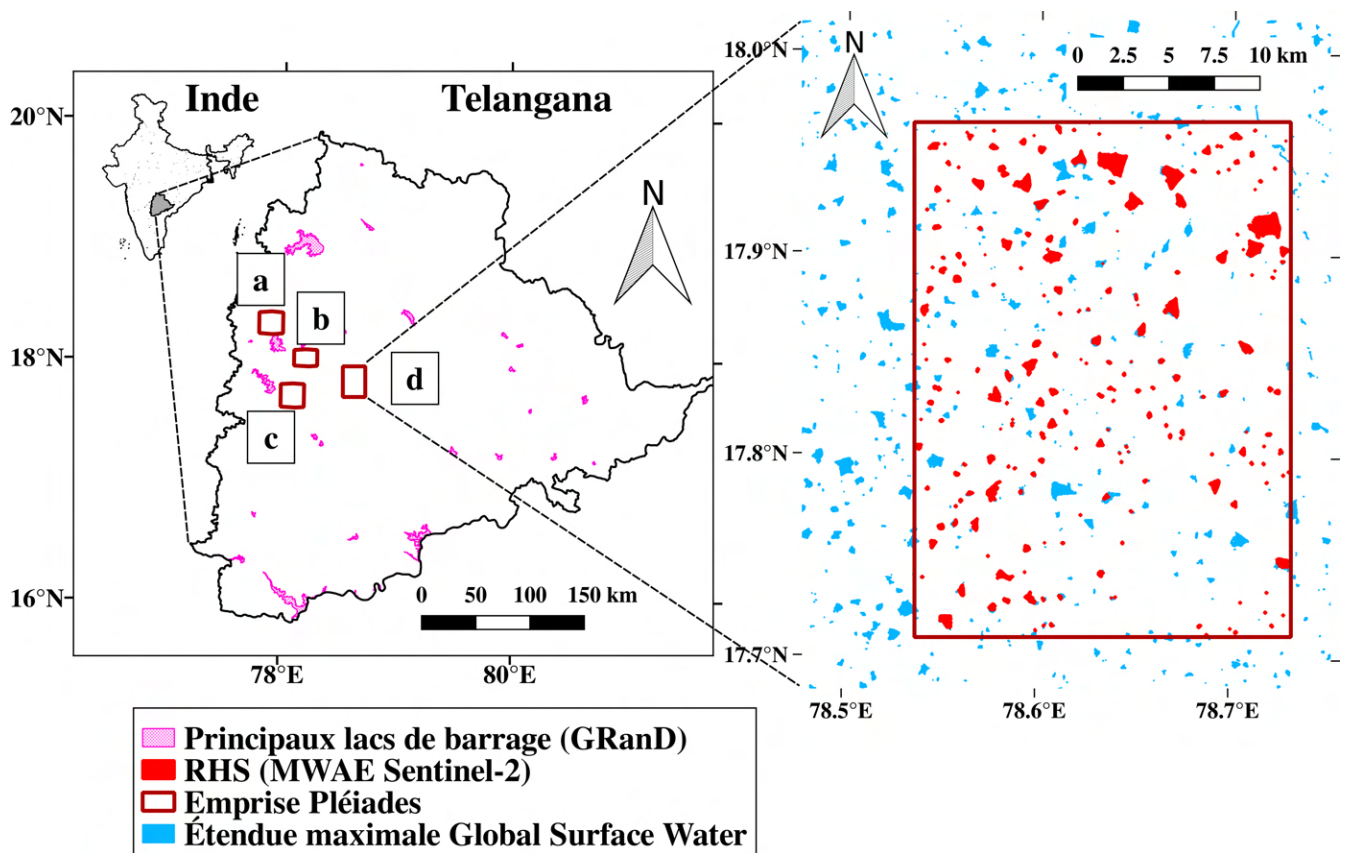


FIGURE 5.2 – Localisation de la zone d'étude et inventaire des données. Les quatre emprises Pléiades (rectangles rouges) sont localisées dans l'ouest du Telangana. Les 25 principaux lacs de barrage identifiés dans l'inventaire GRanD (Lehner et al., 2011) sont en magenta. L'image de droite présente un zoom sur le MNT (d). La localisation des tanks du RHS est indiquée par l'étendue d'eau maximale (MWAE - *Maximum Water Area Extent*) détectée sur les images Sentinel-2 en rouge. L'étendue d'eau maximale de la base de données *Global Surface Water* (Pekel et al., 2016) est indiquée en bleu.

cumulé représente 4100 ha sur les 1813 km<sup>2</sup> de la zone d'étude, soit 2,3%. La superposition du MNT avec le MWAE de Sentinel-2 est confirmée par inspection visuelle, chaque tank identifié par Sentinel-2 correspondant effectivement à une dépression observée sur le MNT.

### 5.2.2.3 Relation hypsométrique basée sur la bathymétrie

Pour chaque tank identifié grâce au MWAE Sentinel-2, des paires de surface et volume ont été estimés à partir du MNT pour plusieurs hauteurs d'eau avec un incrément de 10 cm, de 0 jusqu'à la hauteur maximale atteinte dans le polygone délimitant le tank ( $h_{max}$ ). Pour une hauteur d'eau  $h$ , le volume est calculé à partir de la bathymétrie en intégrant les volumes de chaque pixels, comme dans Vanthof and Kelly (2019) et Arendt et al. (2021) (Equation (5.1)) :

$$V_h = S_{pix} \sum_{pix \in S_h} (z_{min} + h - z_{pix}) \quad (5.1)$$

avec  $h$  la hauteur d'eau variant de 0 à  $h_{max}$ ,  $S_h$  l'aire de tous les pixels inclus avec une altitude inférieure à  $z_{min} + h$ ,  $S_{pix}$  la surface d'un pixel (4 m<sup>2</sup>) et  $h_{max}$  la hauteur d'eau maximale. ( $z_{min} +$

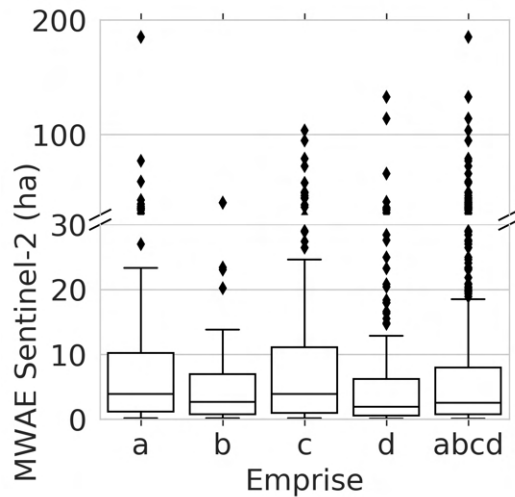


FIGURE 5.3 – Distribution du MWAE des tanks pour chaque emprise Pléiades et pour l’ensemble des emprises (abcd).

$h - z_{pix}$ ) est la profondeur d’eau pour un pixel donné, avec  $z_{pix}$  l’altitude du pixel et  $z_{min}$  l’altitude minimale du tank.

Une courbe hypsométrique continue est construite pour chaque tank par interpolation linéaire entre les points.

#### 5.2.2.4 Délimitation des tanks par une isoligne Pléiades

Malgré la bonne superposition du MNT Pléiades et du MWAE Sentinel-2, la délimitation des tanks par le MWAE de Sentinel-2 perturbe le calcul de la courbe hypsométrique du fait de la divergence entre les deux jeux de données. La figure 5.4 illustre ce phénomène en comparant les courbes hypsométriques calculées sur un tank délimité par le MWAE de Sentinel-2 (en rouge sur la figure 5.4a) ou une courbe de niveau du MNT sélectionnée à la main (en jaune), qui correspond à une surface en eau théorique. Les deux courbes hypsométriques correspondantes sont en figure 5.4b, et montrent que délimiter les tanks par le MWAE Sentinel-2 surestime les volumes d’eau pour les surfaces les plus grandes, et induit un artefact en fin de courbe (courbe rouge).

En effet, les surfaces en eau Sentinel-2 sont souvent sous-estimées du fait de (i) la résolution, (ii) les imperfections connues des méthodes de détection d’eau sur des images multispectrales, en particulier pour les lacs saisonniers peu profonds, sur les bords des lacs ou à cause de la présence de végétation (Doña et al., 2016). Délimiter les tanks par le MWAE rogne les courbes de niveau à partir d’un certain niveau d’eau (comme celle correspondant à une hauteur d’eau de 4.5 m en figure 5.4) et élimine les pixels de plus faible profondeur, ce qui résulte en une surestimation de  $\sum_{pix \in S_h} (z_{min} + h - z_{pix})$  et donc du volume pour les surfaces les plus grandes.

De plus, le  $h_{max}$  d’un polygone Sentinel-2 est généralement plus haut que le véritable niveau d’eau maximal atteint puisque quelques pixels de la digue sont englobés dans le MWAE en raison de la résolution de Sentinel-2 plus grossière que Pléiades. Ceci résulte en un artefact quasi-vertical en fin de courbe lorsque la hauteur (et donc le volume) augmente artificiellement alors que la surface reste quasiment la même. Ainsi, sur le tank de la figure 5.4a, l’eau ne semble pas dépasser les 5 mètres selon le MNT, mais  $h_{max}$  est bien supérieur ( $> 12$  m).

Pour éviter ces imprécisions, les courbes hypsométriques sont calculées avec des tanks délimités par une isoligne du MNT sélectionnée manuellement de telle sorte à ce qu’elle englobe le MWAE



de Sentinel-2 (comme illustré en figure 5.4a), dont la hauteur vaut  $z_{min} + h_{max}$ .

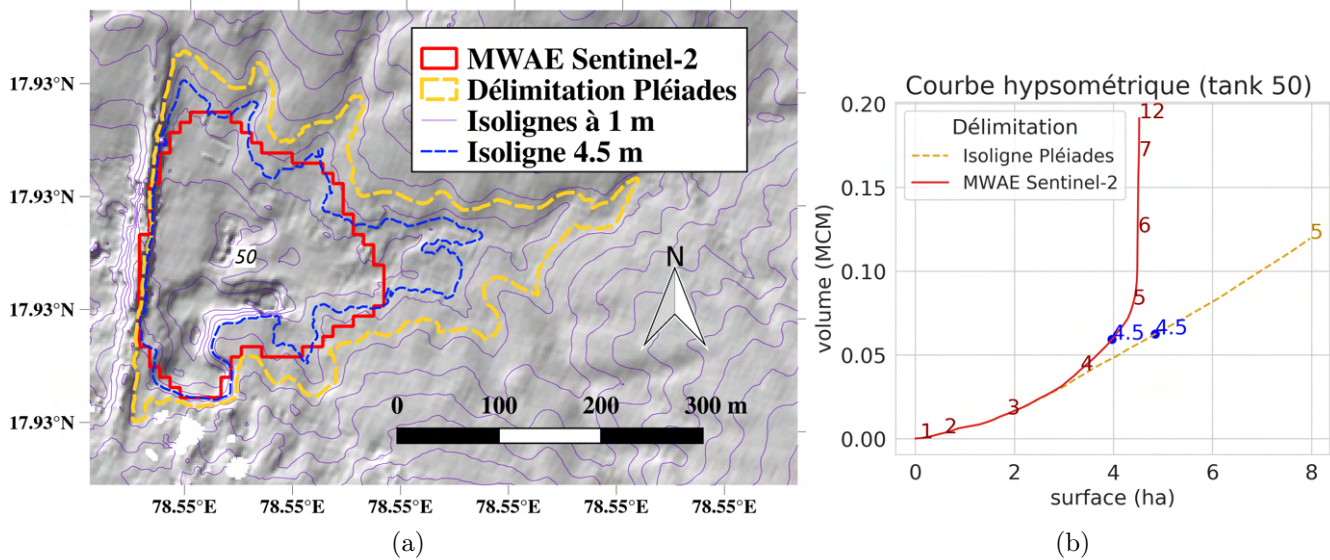


FIGURE 5.4 – Courbes hypsométriques d’un tank délimité par Sentinel-2 et une isoligne Pléiades. (a) Tank délimité par le MWAE Sentinel-2 (rouge) et par une isoligne Pléiades (jaune). Les isolignes à 1 m sont en traits fins violets, et le trait épais bleu en pointillé correspond à une hauteur d’eau de 4.5 m. (b) Courbes hypsométriques du tank délimité par Sentinel-2 (rouge) et l’isoligne Pléiades (jaune). Les deux points bleus correspondent aux surfaces et volumes calculés avec les deux courbes hypsométriques pour une hauteur d’eau de 4.5 m .

### 5.2.2.5 Quantification de la capacité du RHS à l’échelle du Telangana

L’étendue maximale (MWAE) du RHS à l’échelle du Telangana ( $114\,800\text{ km}^2$ ) est quantifiable par imagerie satellitaire optique ou radar, cependant l’estimation du volume des tanks est limitée aux zones couvertes par les quatre MNT Pléiades, soit  $1813\text{ km}^2$ . L’estimation de la capacité maximale du RHS du Telangana ( $V_{max}$ ) se fait à partir de (i) l’inférence d’une relation empirique entre MWAE et  $V_{max}$  sur les emprises Pléiades, (ii) l’identification du MWAE à l’échelle du Telangana à partir de l’étendue maximale de *Global Surface Water* (GSW) dont les rivières et lacs de barrages sont préalablement éliminés, et (iii) le calcul de  $V_{max}$  avec la relation empirique et le MWAE de GSW.

### 5.2.2.6 Dynamique temporelle du stock d’eau à l’échelle d’une emprise Pléiades

La série temporelle de surfaces en eau dérivée de Sentinel-2 est utilisée pour extraire le volume d’eau stockée dans les tanks pour chaque date d’acquisition. Chaque polygone des cartes d’eau est associée à un tank par intersection avec la carte de MWAE Sentinel-2 (produite comme en 5.2.2.2), et son aire permet de retrouver le volume à partir de la relation hypsométrique. Seule la valeur de l’aire est utilisée, aussi une superposition précise des cartes Sentinel-2 à 10 m et des isolignes du MNT Pléiades n’est pas nécessaire. Les valeurs de surface et de volume pour les tanks individuels sont cumulées à l’échelle de l’emprise Pléiades pour chaque date d’acquisition, afin d’obtenir une série temporelle de volume du RHS pour les quatre emprises Pléiades.

TABLE 5.1 – Étendue d’eau maximale observée (MWAE) dérivée de *Global Surface Water* (GSW) ou des séries temporelles Sentinel-2 (SENT) exprimée en  $\text{km}^2$  et en pourcentage par rapport à l’aire de l’emprise Pléiades. La capacité correspondante ( $V_{max}$ ) est exprimée en millions de mètres cubes (MCM) et en mm.

Emprise Pléiades et date d’acquisition	Aire de l’em- prise ( $\text{km}^2$ )	MWAE du RHS ( $\text{km}^2$ / %)		$V_{max}$ estimé sur le RHS (MCM / mm)	
		GSW	SENT	GSW	SENT
a - 01/06/2016	443.6	10.6 / 2.4	11.6 / 2.6	16.1 / 36.2	18.1 / 40.7
b - 14/06/2016	322.3	4.0 / 1.2	4.7 / 1.5	4.2 / 12.9	5.4 / 16.9
c - 14/06/2016	463.5	12.5 / 2.7	13.4 / 2.9	14.4 / 31.0	16.1 / 34.7
d - 16/06/2019	583.9	14.9 / 2.6	12.2 / 2.1	20.1 / 34.5	18.7 / 32.0

## 5.2.3 Résultats

### 5.2.3.1 Capacité du RHS

**Différences entre Sentinel-2 et GSW.** Les volumes des tanks individuels sont calculés à partir de leur MWAE et de leur relation hypsométrique, et sont cumulés sur chaque emprise Pléiades. La table 5.1 présente les MWAE (surfaces en  $\text{km}^2$  et proportion d’eau sur l’image en %) et les  $V_{max}$  correspondants (en MCM et en mm) dérivés de Sentinel-2 et de GSW. Sur les emprises (a), (b) et (c), le MWAE de GSW sous-estime celui de Sentinel-2 en raison de sa plus faible résolution (30 m contre 10 m, phénomène illustré en figure 5.5). Ce n’est pas le cas pour l’emprise (d), puisque la série temporelle d’images Sentinel-2 n’a pas capturé le MWAE qui a eu lieu en septembre 2016 (Ferrant et al., 2017). Pour les emprises (a), (b) et (c), le MWAE de GSW est sous-estimé de 6.7 à 14.9% (pour une moyenne de 10%) et le volume de 10.6 à 22% par rapport à Sentinel-2.

**Relation entre MWAE et  $V_{max}$ .** Les couples (MWAE,  $V_{max}$ ) (en % de la surface de l’emprise et en mm) pour les quatre emprises Pléiades sont reportés en figure 5.6a (losanges gris). Les quatre emprises Pléiades sont divisées en quatre zones géographiques de surface similaire (environ  $10 \times 10 \text{ km}^2$ , en noir sur la figure 5.6b) pour augmenter la taille du jeu de données. Une relation puissance, correspondant légèrement mieux aux données qu’un modèle linéaire, est calibrée sur les 16 couples (MWAE,  $V_{max}$ ) correspondants (points noirs sur la figure 5.6a), de la forme :

$$V_{max} = a.MWAE^b \quad (5.2)$$

avec a et b les coefficients évalués à 10.88 et 1.20 pour un  $R^2$  de 0.94 et un RMSE de 9 mm. À partir de ce modèle, le  $V_{max}$  du RHS à l’échelle du Telangana est déterminé à partir (i) du MWAE de GSW (en bleu sur la figure 5.6a) sur le Telangana couvrant 1.6% de la surface de l’État et (ii) du cumul des volumes estimés par le MWAE GSW sur une grille de  $10 \times 10 \text{ km}^2$  (grille en magenta sur la figure 5.6b). Les couples (MWAE,  $V_{max}$ ) correspondants sont reportés en figure 5.6a (points bleu et magenta respectivement). Les volumes estimés sont respectivement de 18.4 et 20.7 mm, et de 20.9 et 23.5 mm en prenant en compte une sous estimation moyenne de 10% du MWAE GSW par rapport à Sentinel-2.

Cette capacité de 23.5 mm est bien plus faible que la capacité cumulée des grands lacs de barrage recensés par le NRLD (113 mm). Cependant ce réservoir a son importance comparé à la

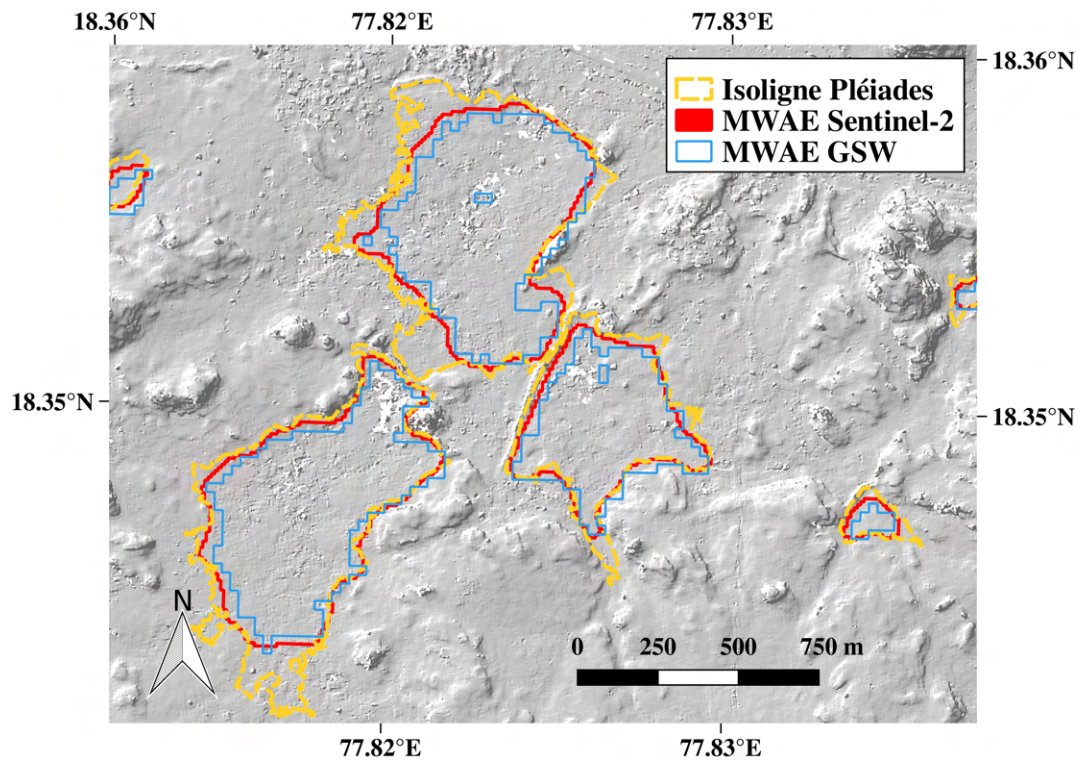


FIGURE 5.5 – Différence entre les MWAE dérivées de Landsat (bleu), Sentinel-2 (rouge) et la délimitation par une isoligne Pléiades (jaune).

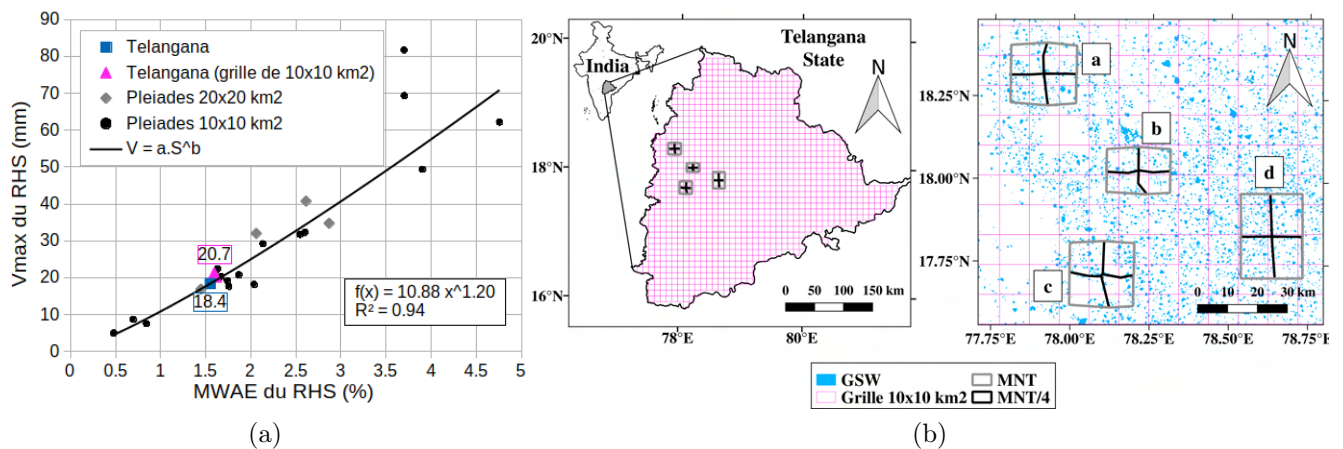


FIGURE 5.6 – Quantification de la capacité théorique du RHS à l'échelle du Telangana. (a) Relation empirique entre la capacité maximale du RHS (mm) et son MWAE (%). ( $MWAE$ ,  $V_{max}$ ) pour les emprises Pléiades natives (losanges gris, emprises en gris sur la figure (b)) et pour les emprises divisées en quatre zones géographiques similaires (points noirs, emprises en noir sur la figure (b)). Une relation puissance est calibrée sur les points noirs. La capacité du RHS à l'échelle du Telangana est calculée à partir (i) du MWAE de GSW à l'échelle du Telangana (triangle bleu, en bleu sur la figure (b)) et (ii) par le cumul des volumes calculés à partir du MWAE sur une grille de  $10 \times 10 \text{ km}^2$  (carré magenta, en magenta sur la figure (b)).

demande en irrigation. Sur l'aquifère de la ville de Gajwell (emprise Pléiades (d)), une étude à

partir de cartes de riz dérivées d'images satellite et d'enquêtes de terrain a estimé la demande pendant la saison sèche entre 62 à 295 mm pour les saisons sèches de 2016 et 2017 (Ferrant et al., 2017).

La capacité du RHS du Telangana peut être comparée avec de précédentes études sur le volume des tanks à partir du MNT TanDEM-X. Vanthof and Kelly (2019) ont estimé cette capacité à 1.5 mm sur un bassin de 5647 km<sup>2</sup> dans l'État indien du Tamil Nadu en utilisant le MWAE de Sentinel-1 de 2015 et un sous-échantillon de 274 tanks sur 1503 identifiés. Dans le bassin de Cuvelai (18 370 km<sup>2</sup>), localisé entre l'Angola et la Namibie, la capacité de 190 000 dépression naturelles est estimée à 103 mm (Arendt et al., 2021). La première étude sous-estime clairement la capacité maximum car réalisée sur un sous-échantillon de tanks sur une unique date d'acquisition avec des données radar, qui sous-estiment les surfaces en eau du fait de la difficulté du radar à détecter les bords des étendues d'eau. La seconde étude dépasse les estimations faites sur des tuiles de 10 × 10 km<sup>2</sup> sur le Telangana, qui varient de 5 à 82 mm (voir figure 5.6a).

### 5.2.3.2 Suivi temporel du volume du RHS

Des séries temporelles de volume du RHS sont produites à partir des cartes d'eau dérivées de Sentinel-2, à l'échelle des quatre emprises Pléiades (a), (b), (c) et (d) (figure 5.7). Les séries temporelles optiques Sentinel-2 sont interrompues pendant la mousson en raison de la couverture nuageuse, ce qui n'entrave pas le suivi de la vidange des réservoirs pendant la saison sèche et ses usages pour la compréhension des impacts hydrologiques du RHS ou le développement de systèmes d'alerte de sécheresse.

L'étude en section 5.3 applique cette méthode de suivi temporel avec la bathymétrie des MNT Pléiades et d'autres MNT dans une étude comparative. Elle démontre l'intérêt des MNT très haute résolution à fort taux de revisite par rapport aux MNT globaux actuellement disponibles, et présente un exemple d'application future de tels MNT qui seront fournis dans le cadre de la mission CO3D (Lebègue et al., 2020).

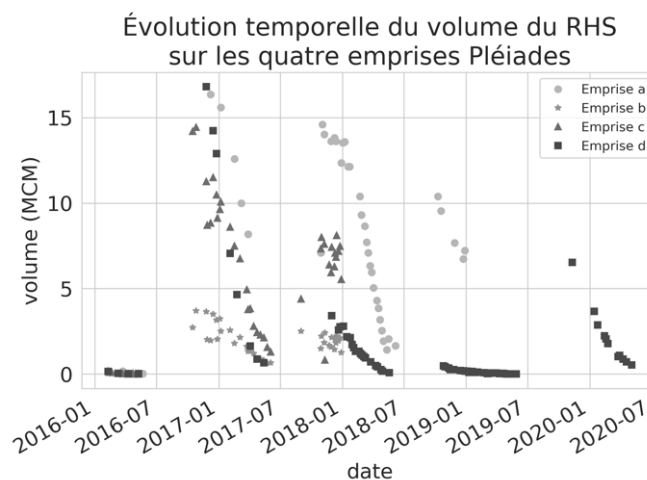


FIGURE 5.7 – Séries temporelles de volume d'eau (en millions de mètres cubes - MCM) à l'échelle des quatre emprises Pléiades.

### 5.3 Monitoring the volume of South Indian Rainwater Harvesting System : sensitivity of the bathymetry retrieval to DEM features

Claire Pascal<sup>1</sup>, Sylvain Ferrant<sup>1</sup>, Santiago Peña-Luque<sup>2</sup> Adrien Selles<sup>3</sup>, Jean-Christophe Maréchal<sup>3</sup>, Simon Gascoin<sup>1</sup>, Olivier Merlin<sup>1</sup>.

Soumis à *Remote Sensing* (MDPI) en novembre 2022.

---

1. Centre d'Étude Spatiale de la BIOSphère (CESBIO-UPS-CNRS-IRD-CNES-INRAE), 18 av. Ed. Belin, Toulouse CEDEX 9, 31401, France

2. Centre National d'Études Spatiales (CNES), 18 Avenue Edouard Belin, Toulouse, 31400, France

3. Bureau de Recherches Géologiques et Minières (BRGM), Université de Montpellier, 1039 rue de Pinville, Montpellier, 34000, France

# Monitoring the volume of southern Indian Rainwater Harvesting System: sensitivity of the bathymetry retrieval to DEM features

Claire Pascal <sup>1,†\*</sup>, Sylvain Ferrant <sup>1,†</sup>, Santiago Peña-Luque <sup>2</sup>, Adrien Selles <sup>3</sup>, Jean-Christophe Maréchal <sup>3</sup>, Simon Gascoïn <sup>1</sup> and Olivier Merlin <sup>1</sup>

<sup>1</sup> Centre d'Etudes Spatiales de la BIOSphère (CESBIO), UPS-Université de Toulouse, CNRS, CNES, IRD, INRAE, 18 Av. Edouard Belin, bpi 2801, 31401 Toulouse, France

<sup>2</sup> Centre National d'Etudes Spatiales (CNES), 18 Avenue Edouard Belin, 31400 Toulouse, France

<sup>3</sup> Bureau de Recherches Géologiques et Minières (BRGM), Université de Montpellier, 1039 rue de Pinville, Montpellier, 34000, France

\* Correspondence: claire.pascal@univ-tlse3.fr

† These authors contributed equally to this work.

**Abstract:** The Indian Rainwater Harvesting System (RHS) is composed of hundreds of thousands of Small Reservoirs (SR < 10 ha) disconnected from any perennial river, empty a part of the year. Crucial for irrigation during the dry season in upstream areas, this regional water resource is not monitored, even with remote sensing, given the small size and depth of those reservoirs. Nevertheless, nowadays high to very high resolution satellite images can be an opportunity to monitor the surface area and volume content at regional scale. We quantified for the first time this resource in the Deccan Plateau, south India, and the related systematic error based on the choice of available DEMs. To do so, we used a set of 4 very high (2-m) resolution Digital Elevation Model (DEM) derived from stereoscopic acquisitions of Pleiades Satellites, covering a total area of 1,813 km<sup>2</sup>. Located in the Telangana State (114,789 km<sup>2</sup>), 505 SR geometries from this ungauged hydrological system have been acquired when SR were empty. We used those SR geometries as a surface-volume hypsometric relationship to evaluate the change of hypsometric relationship with the resolution (from native 2 m to 90 m). Since those VHR data are not yet available everywhere, we used the 2 m data-set as a reference to quantify the volume retrieval drift using coarser but freely available global and regional DEMs: ALOS World 3D-30 m, WorldDEM GLO-30 at 30 m TanDEM-X DEM at 90 m and one Indian DEM (CartoDEM at 30 m). Time series of Sentinel-2-derived surface water area extent were used to retrieve volumes from any hypsometric relationship derived from the five DEMs. The differences in volume retrievals are quantified using the Mean Absolute Percentage Error (MAPE) using the 2m-DEM based geometries as a reference. MAPE of global DEMs range from 47% to 78% whereas the MAPE are 17%, 29% and 46% for Pleiades DEM aggregated at 12, 30 and 90 m, respectively. Aggregated volumes over the studied 505 SR were estimated from global DEMs. Among them, ALOS World 3D-30m and CartoDEM overestimate the small water storage and underestimate volume for large filling rates contrary to GLO-30 and TanDEM-X DEM that generally underestimate water volumes by more than 40%. The relative uncertainty using ALOS reaches 26%. Very high-resolution DEMs acquired in the dry period are a promising source of data for improving our capacity to monitor water storage at SR scale, with acceptable MAPE (<20%) after resampling at 12 m to easier the data distribution. New satellite missions such as CO3D will soon democratize such high resolution DEM, at lower costs and with higher coverage and revisit, to catch empty tanks geometries throughout India.

**Keywords:** Very high-resolution DEM; Pleiades; ungauged small reservoirs; Sentinel-2 water detection; semi-arid agro-hydrology; reservoir bathymetry

**Citation:** Lastname, F.; Lastname, F.; Lastname, F. Monitoring the volume of South Indian Rainwater Harvesting System: sensitivity of the bathymetry retrieval to DEM features. *Remote Sens.* **2022**, *1*, 0. <https://doi.org/>

Received:

Accepted:

Published:

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Copyright:** © 2022 by the authors. Submitted to [Remote Sens.](https://www.mdpi.com/journal/remotesensing) for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Many semi-arid regions are facing an increasing need for water, both for irrigation water to sustain food production and for domestic water, leading to an increased water

stress [1]. To cope with these increasing stressors, many infrastructures have been built to manage rain and river waters, such as big dams and reservoirs. Contrary to big surface reservoirs, resulting from river dam installations, small reservoirs (SR), or tanks, are less inventoried and monitored. Yet, they are known to have a strong impact on hydrology worldwide [2]. Increasing the number of SR, especially in upstream regions, is a common adaptation to water stressors [3,4]. In South India, SR are used to catch monsoon rainfall runoff in upstream areas, to sustain crop irrigation and maintain open water for cattle during the following dry season. SR have also been promoted to enhance aquifer recharge [5], which has been questioned in terms of efficiency as the loss associated with direct evaporation could be high [6–8].

This Indian traditional Rainwater Harvesting System (RHS) is composed of hundreds of thousands of SR, spread along a non perennial drainage network. In the Telangana State, a typical SR consists in a flat depression of around one kilometer wide, dammed with an earthen dyke of 3 to 10 meters high and 20 meters wide, that often completely dries up during the dry season. In this Indian state, a costly 5-years program named Mission Kakatiya started in 2015 to rejuvenate more than 46,000 SR, i.e. excavating for a few meters to desilt the bottom part and thus enhance the aquifer recharge and create new ones. This program shows the socio-political importance of the RHS, but also its fast structural changes following a growing water demand.

Quantifying the storage capacity and monitoring the volume of this RHS is hence crucial to estimate water stocks and to forecast stressors on local to regional water needs, yet is challenging due to their large number and the scarcity of in-situ monitoring systems. These monitoring difficulties could be overcome with the increasing volume of available satellite data.

Usually, water volumes in reservoirs are estimated based on combinations of altimeter (e.g. ICESAT, ENVISAT, Topex/Poseidon) measurements and surface water area extents derived from radar (TERRASAR-X, RADARSAT-2) or multispectral optical (Landsat, Formosat) satellite data [9–12]. Those retrievals were performed for single isolated lakes and not at a regional scale. More recently, the multiplication of satellite missions for Earth topography mapping led to the release of free access global Digital Elevation Model (DEM) such as Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) [13], TanDEM-X DEM [14], Advanced Land Observing Satellite “DAICHI” (ALOS) World 3D-30m [15] and Copernicus’s WorldDEM GLO-30, with a resolution up to 30 m. This encouraged the development of methods for surface water volume retrieval using the bathymetry derived from DEM combined with either altimeter data [16], satellite lidar transects [17] or time series of surface water extent derived from Landsat [16,18,19] or the more recent mission Sentinel-1 [20]. However, surface retrieval from either radar or optical data has its own uncertainties. Landsat data covering 40 years have been processed by several research teams to produce monthly water maps [21,22]. Recent studies have explored the capacities and accuracies of Sentinel missions to retrieve regional surface water dynamics [23,24] with a focus on radar and optical related bias to this retrieval [25].

In most cases, the acquisition date of the DEMs cannot be chosen by the user which results in impossible retrieval of the bottom geometry of reservoirs due to the presence of vegetation or water. To cope with these imperfections, the reservoir geometry is often simplified with a pre-definite trapezoidal shape [9,16,26], by topographic continuity along channels [27] or by modelling the volume-area relationship [9,18–20].

Several approaches of volume estimates methods at regional scale have been tested for large reservoirs [11,16,19] and more recently for SR [20,28]. In particular, a previous study has reported the interest of high resolution remote sensing data to retrieve the volume-area relationships for SR of the RHS in South India, Tamil Nadu state, using the TanDEM-X DEM commercial product at 12-m resolution [20]. This study has estimated a fluctuation of 8.2 million cubic meter (MCM) of surface water in 559 SR over a region of 5,647 km<sup>2</sup> using dynamic surface water mapping. According to the authors, the limitations of this study are (i) the surface water detection accuracy based on Sentinel-1 radar data and (ii)

the simplification of reservoirs' geometry with power volume-area relationships due to the DEM precision.

However, to this date, the freely available global DEMs have a resolution coarser than that of TanDEM-X DEM commercial product. For instance, it is 30 m for ALOS World 3D-30m, WorldDEM GLO-30 and CartoDEM or 90 m for TanDEM-X non-commercial DEM. A coarse resolution complexifies the use of remote sensing for water volume estimation of the SR of RHS.

Recent high-resolution on-demand DEMs like the ones produced from Pleiades very-high (70 cm) resolution (VHR) stereoscopic images allow controlling the acquisition date and the accurate retrieval of reservoir bathymetry. Yet they are commercial products only available locally, as Pleiades scenes typically only cover a  $20 \times 20$  km<sup>2</sup> area [29]. The need for large scale inland water bodies monitoring have motivated the forthcoming SWOT (Surface Water and Ocean Topography) mission [30], yet its resolution will not be sufficient for the majority of the SR of the RHS, in particular those with an area below 10 ha. On the other hand, upcoming very high resolution stereoscopy satellite missions such as CO3D [31] will provide VHR DEM cheaper than Pleiades with a high temporal frequency. These DEMs will be crucial to monitor a high number of these SR, as it has been done only for individual SR so far [6–8]. Regional studies on SR have not been attempted with freely available DEMs because of their coarse ( $\geq 30$ m) resolution, although there has been two attempts with a commercial 12-m DEM [20,28]. The objective of this study is to quantify the efficiency of VHR DEM acquired during dry periods (such as Pleiades or those from CO3D available in the near future), compared to currently available global coarse DEM, for the retrieval of surface water volumes of SR composing the Indian RHS.

In this study, we use VHR DEM to extract empty SR geometries, taken as a reference, to explore the drifts in volume estimates induced by coarser resolution DEMs, upscaled from the reference or extracted from global to national available DEMs. In particular, we use four VHR (2-m) Pleiades DEM, validated with ground control points, as a true reference for 505 SR bathymetries over a small area covering 1,813 km<sup>2</sup> in the Telangana state. Pleiades images were selected because all SR were empty at this time and the cloud cover was low. Surface water volumes at DEM scale are computed from a time series of surface water extent areas derived from Sentinel-2 optical data with a 10-m resolution. We use volumes derived from VHR recent Pleiades DEM as a reference to quantify the difference with volumes retrieved from four of the most recent large-scale freely available DEM : CartoDEM (30-m Indian DEM), Copernicus WorldDEM GLO-30, ALOS World 3D-30 m (30-m global DEMs) and TanDEM-X DEM (90-m global DEM). We analyze the impact of DEM features on volume estimates, namely the spatial resolution, the numerical precision used for data dissemination (float or integer format), the SR water filling ratio at the time of acquisition, defined as the percentage of RHS MWAEE that was flooded during DEM acquisition, and the time lapse between their acquisition period and the reference Pleiades DEM.

## 2. Study area and data sets

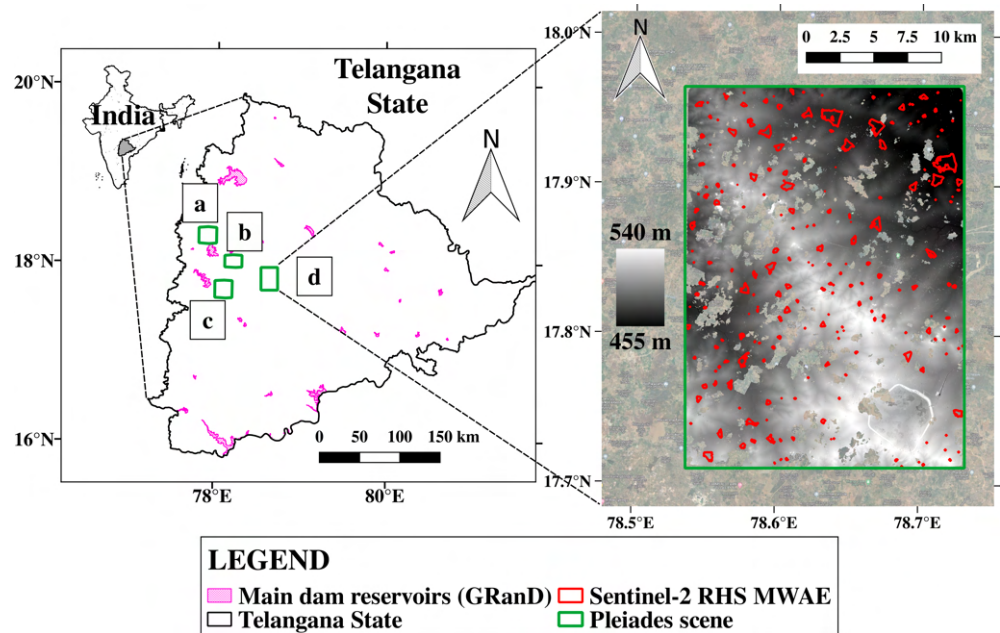
### 2.1. Study area

The Telangana State (114,789 km<sup>2</sup>) is a highly irrigated region in South India with a high population density (about 335 inhabitants per km<sup>2</sup> in 2020 according to the Unique Identification Authority of India (UIDAI)), dominated by a semi-arid climate. The monsoon precipitation occurs between July and October and ranges from 540 to 1,300 mm with a mean of 879 mm (Indian Meteorological Department). In this area, groundwater is complementing the surface water resource to sustain one or two growing seasons of irrigated crops in a year. The irrigated water demand strongly varies with the available water resource, from 51 to 310 mm per growing season in upstream areas [32,33].

The majority of the State (57%) is composed of a shallow fractured granitic aquifer characterized by high fluctuations due to water pumping [34]. Monsoon rainfall is stored on the surface by large dam reservoirs built along rivers. The Indian National Register of



Large Dams (NRLD) reports 126 dam lakes with a cumulative capacity of 113 mm. The 25 largest reservoirs, referenced in the global GRanD inventory [35], are shown in magenta in Figure 1. In upstream areas, surface water storage is ensured by the RHS, a network of non-perennial SR which fill during the monsoon season and empty quickly during the dry season. These structures are also promoted for aquifer recharge enhancement [7]. A previous study estimated the total capacity of the SR of the RHS of the Telangana State at 30 mm using Global Surface Water database [21], Pleiades DEM and Sentinel-2-derived MWAE [36].



**FIGURE 1.** Location of the study area. The four Pleiades scenes (a), (b), (c) and (d) (green rectangles) are located in the western part of the Telangana State. The 25 main dam lakes identified in the GRanD inventory (Lehner et al., 2011) are represented in magenta. The image on the right shows a zoom on the DEM on scene (d) with Pleiades elevations in shades of grey. The location of the small reservoirs (SR) of the Rainwater Harvesting System (RHS) is indicated by the Sentinel-2 Maximum Water Area Extent (MWAE) in red.

## 2.2. DEM

In the following, we present the characteristics of available global DEMs that we have selected to explore the possibility to estimate volumes stored in the Telangana ungauged RHS. The main information on the DEMs used in this study are summarized in Table 1.

### 2.2.1. Pleiades 2-m DEM

The Pleiades stereoscopic images used in this study include three stereo-pairs that were acquired in June 2016 (available in Airbus Defense and Space catalogue before this study, images (a), (b), (c) in Figure 1) and one stereo-pair acquired in June 2019 (specifically tasked for the purpose of this study, image (d) in Figure 1), and cover respectively 444, 322, 463 and 584 km<sup>2</sup>. The images were acquired at the end of the dry season to get the bathymetry of empty reservoirs. In those four areas, a 2-m resolution DEM was generated using the panchromatic stereoscopic images and the NASA Ames Stereo Pipeline (ASP) software V2.6.2 [37]. The photogrammetric processing was done using the rational polynomial coefficient provided by the satellite operator. First, we generated point fields from the panchromatic images with the block matching algorithm of the stereo utility of ASP Marti et al. [38]. Then, we rasterized the point fields on a 2-m resolution UTM grid using the

point2dem tool of ASP [39]. Some areas covered by water (river bed for example) or bright soils for which the reflectances were saturated have no elevation estimates, explaining No-Data values.

### 2.2.2. CartoDEM

The Cartosat-1 Digital Elevation Model (CartoDEM) is provided by the Indian Space Research Organization (ISRO). It is produced from stereo panchromatic images from Cartosat-1 at a 2.5-m resolution and the water surfaces were flattened. It is provided with a horizontal resolution of 30 m and elevation values are rounded to the meter. The accuracy of CartoDEM was estimated on three test areas against control points, with a RMSE ranging from 3.4 to 4.72 m and a linear error at a 90% confidence level ranging from 4.7 and 7.3 m [40].

### 2.2.3. TanDEM-X DEM

TanDEM-X DEM is a global DEM generated from bistatic X-Band interferometric SAR acquisitions of TerraSAR-X and TanDEM-X satellites. The TanDEM-X mission was financed by a Public-Private Partnership between DLR (German Aerospace Centre) and Airbus Defence & Space. The DEM is produced with acquisitions ranging from December 12, 2010 to January 16, 2015. It has a native resolution of 0.4 arcseconds (<10 m) and is provided with a horizontal resolution of 3 arcseconds (90 m) and a vertical accuracy <2 m for slopes <20%, and <4 m for slopes <40% [41].

### 2.2.4. Copernicus WorldDEM GLO-30

The Copernicus DEM is a Digital Surface Model (DSM) produced with the data from the TanDEM-X mission. The instance used for this study is GLO-30, a global DEM with a resolution of 30 m and the same vertical accuracy as TanDEM-X DEM.

### 2.2.5. ALOS World 3D-30m

ALOS World 3D-30m (AW3D30) Version 3.2. is a global DSM processed by the Japan Aerospace Exploration Agency (JAXA). It is produced from stereoscopic panchromatic optical images from the Panchromatic Remote-sensing Instrument for Stereo Mapping (PRISM) sensor of the Advanced Land Observing Satellite "DAICHI" (ALOS) which operated from 2006 to 2011. It is provided with a spatial resolution of 30 m and elevation values are rounded to the meter. The accuracy of AW3D30 is expressed as a RMSE of 5 m in both vertical and horizontal directions [42].

### 2.2.6. Pleiades at lower resolution

The 2-m Pleiades DEMs were aggregated at 12 m (resolution of TanDEM-X DEM commercial product), 30 m and 90 m using bilinear resampling. This allows to investigate the isolated effect of DEM resolution on surface water storage estimates and to quantify the impact of a lower resolution more suitable for data dissemination.

## 2.3. Sentinel-2 images

Sentinel-2 is a constellation of two satellites launched in June 2015 and March 2017 that deliver multispectral images at 10- and 20-m resolution and a revisit time of five days. We used Sentinel-2 images processed by the MAJA processing chain [43] which detects clouds, clouds and topographic shadows, and corrects reflectance levels from atmospheric phenomena (absorption by atmospheric gases and scattering by air molecules and aerosols). These S2-L2A images were downloaded from Theia data platform (<https://www.theia-land.fr/>, accessed on August 5th, 2021). A total of 42, 33, 76 and 76 cloud-free Sentinel-2 images over a period of 2 years close to the Pleiades acquisition date were downloaded for scenes (a), (b), (c) and (d) respectively.

TABLE 1. DEM used, their resolution, sensors and acquisition date. "Composite" indicate that the DEM was produced using multiple acquisitions.

DEM	Horizontal resolution	Data type	Acquisition date	Sensor (resolution)
Pleiades (Airbus)	2 m	float	(a) : 01-06-2016; (b) : 14-06-2016; (c) : 16-06-2016; (d) : 16-06-2019	Pleiades optical stereoscopy (50 cm)
CartoDEM (ISRO)	30 m	integer	2006-2016 (composite)	Stereo panchromatic images Cartosat-1 (2.5 m)
TanDEM-X DEM (EOC)	90 m	float	2010-2015 (composite)	X-Band interferometric SAR TanDEM-X (2.4 m - 4 m)
WorldDEM GLO-30 (Copernicus)	30 m	float	2010-2015 (composite)	X-Band interferometric SAR TanDEM-X (2.4 m - 4 m)
ALOS World 3D-30m (JAXA)	30 m	integer	2006-2011 (composite)	PRISM panchromatic optical sensor ALOS (2.5 m)

### 3. Methodology

This section details the method used to retrieve water volumes from DEMs and Sentinel-2 data. This process is split in three main steps, illustrated in Figure 2. First, we preprocessed all DEMs, that is Pleiades 2-m DEM, the four large-scale DEMs, as well as Pleiades DEM aggregated at 12 m, 30 m and 90 m. Then, we used Sentinel-2 images to produce time series of surface water areas and to identify the SR on the study area. Finally, water volumes were estimated using the ground geometry of SR extracted from the DEMs and time series of surface water areas. All volume estimates were compared against the reference Pleiades 2 m DEM estimates.

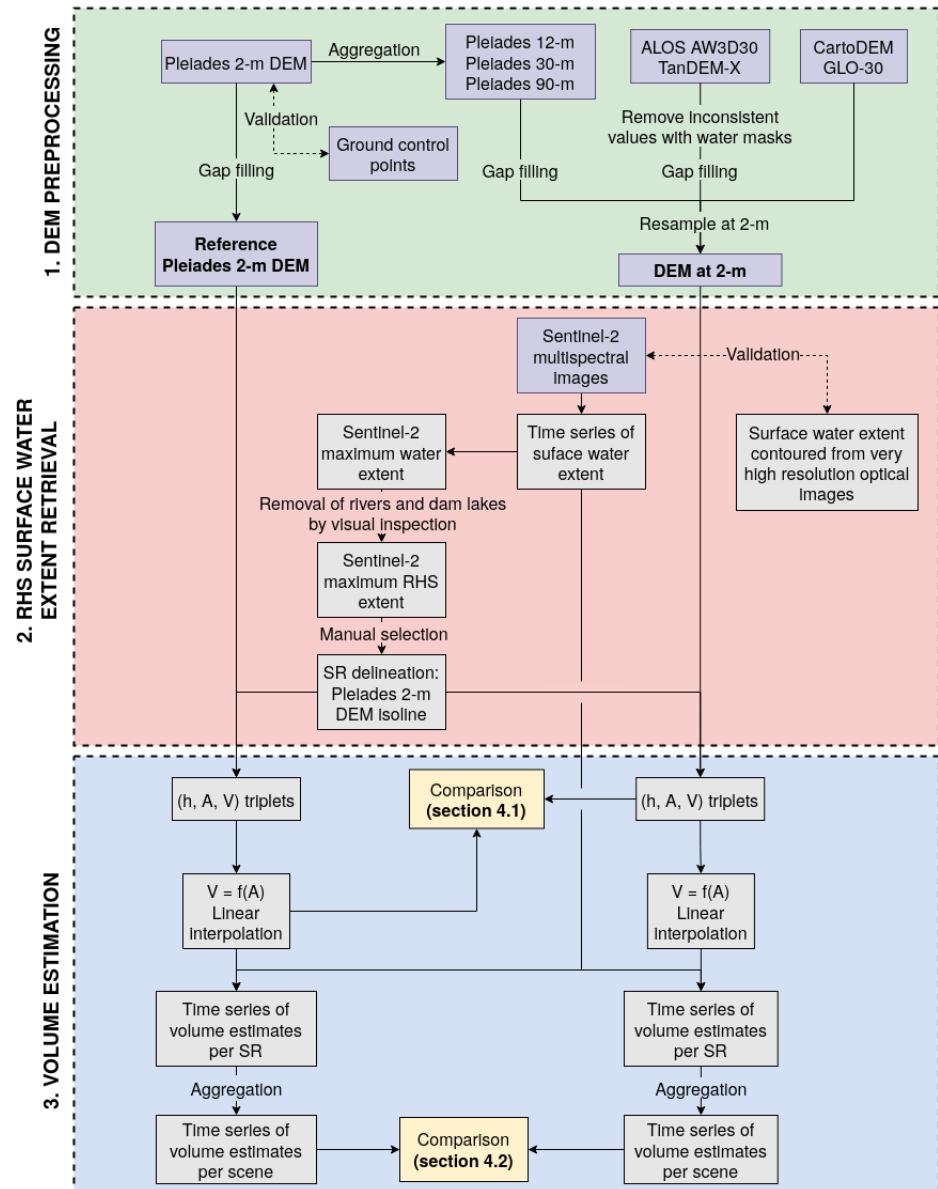


FIGURE 2. Schematic flow diagram of the proposed method for volume estimation. SR : small reservoir. RHS : Rainwater Harvesting System. h, A, V : height, area and volume.  $V=f(A)$  : volume-area relationship.

### 3.1. Step 1 : DEM preprocessing

#### 3.1.1. Validation of Pleiades DEM

The evaluation of Pleiades DEM accuracy would require field elevation data surveyed close to Pleiades acquisition date. Although we do not have such data, we have access to field data acquired in 2010 from Perrin et al. [8]. These data consists in 1,146 differential-GPS (DGPS) measurements mapping the bottom of four empty tanks located in Pleiades scene (d) (see Figure 3).

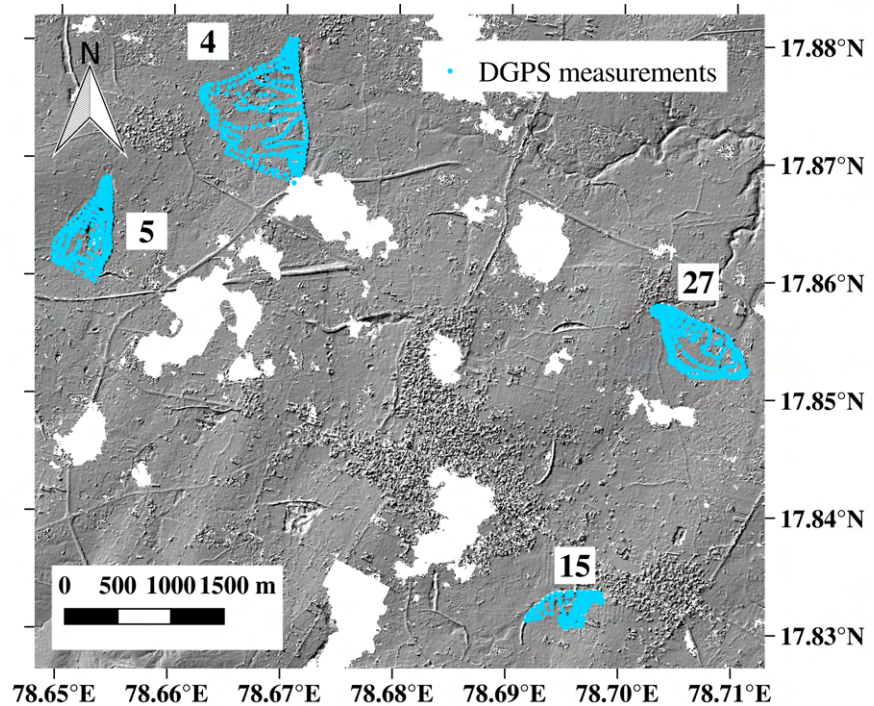


FIGURE 3. Location of the four surveyed tanks in Pleiades scene (d) (numbered as in [8]), superimposed with Pleiades DEM. Tanks 4, 5, 15 and 27 were surveyed with DGPS for a total of 1,146 measurements (blue dots) distributed over the four tanks.

We use three metrics to characterize the error of Pleiades elevation at the location of the field measurement with respect to DGPS :

- An estimation of the pixel level accuracy by the standard deviation of the distribution of the differences between DGPS measurements and Pleiades elevation,
- The RMSE between DGPS measurements and Pleiades elevations, which corresponds to the error provided with CartoDEM and AW3D30,
- The absolute vertical accuracy, which is the uncertainty in Pleiades elevation with respect to DGPS measurements caused by random and uncorrected systematic errors, expressed as a linear error at 90% confidence level. This corresponds to the error provided with the 12-m commercial TanDEM-X DEM and CartoDEM.

#### 3.1.2. Preprocessing

CartoDEM, TanDEM-X DEM, GLO-30 and AW3D30 products are provided with water masks, indicating the presence of water at the acquisition dates. The elevation values for areas corresponding to surface water can be very noisy, therefore DEMs must be corrected for water surfaces. Elevation values within water bodies are already flattened for Copernicus GLO-30 and CartoDEM, contrary to AW3D30 and TanDEM-X DEM. For these last two, water masks were used to remove the inconsistent values and the DEMs were then gap-filled as in Arendt et al. [28]. All DEMs were then resampled at 2 m with

bilinear interpolation to ensure (i) comparable results with Pleiades 2-m DEM and (ii) a sufficient number of pixels is available to extract the geometries of SR.

### 3.2. Step 2 : RHS surface water extent retrieval from Sentinel-2 data

#### 3.2.1. Time series of surface water area extents

Surface water maps were produced from the time series of Sentinel-2 images at 10-m resolution. The surface water maps were obtained from a supervised classification at pixel level using the random forest algorithm [44], commonly used in surface water detection [45,46]. A training data set was selected for each Sentinel-2 L2A product using a randomized selection of 10,000 pixels, including 2,000 water samples. The water pixels were selected from permanent water identified from the Global Surface Water occurrence dataset [21] and the HAND (Height Above Nearest Drainage) dataset [47] is used to select non water pixels. We build the model on 100 trees and used Sentinel-2 bands B3, B4, B8, B11, B12, as well as the NDWI [48] and MNDWI [49] indices for classification.

#### 3.2.2. Validation of surface water maps

The accuracy of Sentinel-2-derived water maps was evaluated against very high resolution historical images from Google Earth Pro (Maxar Technologies or CNES/Airbus images) between 2016 and 2020. We manually contoured water extents for 16 SR at several time steps that matched Sentinel-2 acquisition date down to five days, for a total of 29 points.

#### 3.2.3. SR identification and delineation

Because of their small size, the SR of the RHS are not referenced in existing national (NRLD) or global (GRAND [35]) inventories of water reservoirs. Local inventories for these small reservoirs are rare and their contours can be geometrically imprecise [20]. While such inventories do not exist, SR can also be identified from remote sensing data. The automatic Blue Spot detection [28,50] allows the retrieval of natural depressions extents directly from a DEM, yet they are not guaranteed to be indeed active SR. Active SR can also be identified with surface water detection methods from satellite images : Vanthof and Kelly [20] used the combination of a local authority inventory shapefile and the Global Surface Water dataset based on 30-m resolution Landsat data [21] to identify and retrieve the boundaries of 1503 SR in the Tamil Nadu State in India.

In this study, we identified the SR of the RHS with surface water maps derived from Sentinel-2. All maps of the time series were superimposed and merged to obtain a polygon shapefile of the Maximum Water Area Extent (MWAE). The shapefile was refined through visual inspection by removing (i) the areas corresponding to river bed and dammed river reservoirs, (ii) SR containing more than 5% of Pleiades 2-m DEM no-data and (iii) areas inferior than 1,000 m<sup>2</sup>, threshold below which they are considered temporary puddles and not dug structures. This led to the identification of 505 SR over the 4 Pleiades scenes, composing the RHS : 112 on scene (a), 91 on (b), 114 on (c) and 188 on (d). Their size varies from 0.1 ha to 185 ha, and they cumulatively represent 4,100 ha over the 1,813 km<sup>2</sup> study area.

The Sentinel-2 MWAE of the 505 SR was validated visually with Pleiades DEM by ensuring that each polygon corresponded to a depression in Pleiades DEM. Despite a very good superimposition of the two datasets, a preliminary analysis showed that the discrepancy between Sentinel-2 derived MWAE and the isolines derived from Pleiades DEM leads to non negligible differences between volume estimates for largest area extents. Indeed, besides the good registration between the two datasets, Sentinel-2 resolution and optical water detection algorithm leads to an erosion of water bodies, i.e. undetected water on the water body contour within mixels (mixed pixels of water and non water) or for shallow depths. This erosion leads to an imperfect superposition of both MWAE and isolines. Due to those geographical discrepancies, the maximum elevation within the MWAE is not realistic. Thus, each SR MWAE were contoured using the closest elevation isoline selected

manually from the Pleiades 2-m DEM (step illustrated in Figure 4). The maximum water level reached within the polygon  $h_{max}$  was then defined using the difference between the elevation of this isoline and the minimum elevation within the MWAE.

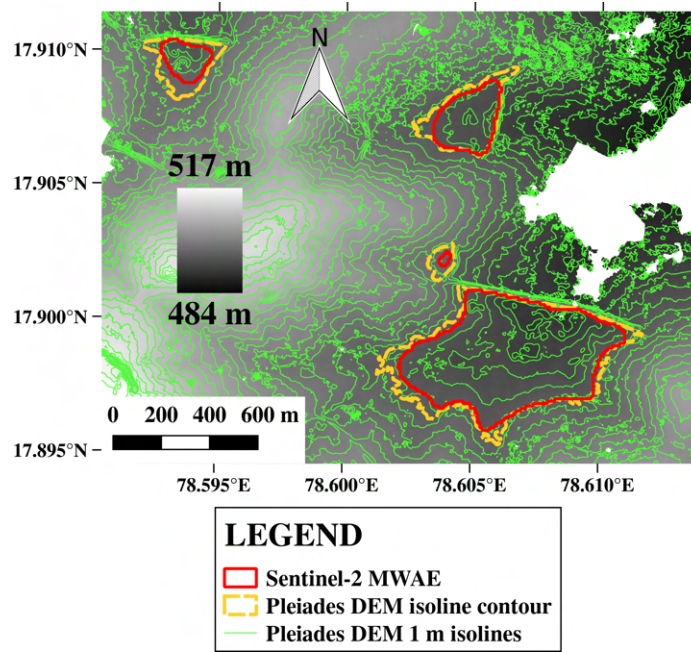


FIGURE 4. Comparison of Sentinel-2 MWAE (red line) and Pleiades DEM isolines (green lines). The yellow dashed line is the isoline that was manually selected for the construction of the hypsometry relationship of the 505 small reservoirs of the study. The background map is Pleiades DEM elevation in meters.

### 3.3. Step 3 : Volume estimation

#### 3.3.1. DEM-based volume estimation method

For each SR delimited by an isoline from Pleiades DEM, paired volume and area estimates were extracted from each DEM at 10 cm height increment, from 0 to the maximum water level reached within the polygon (the height of the isoline  $h_{max}$ ). For a given water height  $h$ , we compute the volume with the true bathymetry by integrating water volumes pixel-wise, as in Vanthof and Kelly [20] and Arendt et al. [28] (Equation (1)).

$$V_h = A_{pix} \sum_{pix \in A_h} (z_{min} + h - z_{pix}) \quad (1)$$

with  $h$  the water level ranging from 0 to  $h_{max}$ ,  $A_h$  the cumulative area of all the pixels within the SR maximum extent and with an elevation below  $h$ ,  $A_{pix}$  the surface of one pixel ( $4 \text{ m}^2$ ),  $h_{max}$  the water level at maximum capacity.  $(z_{min} + h - z_{pix})$  is the water depth for a given pixel, with  $z_{pix}$  the altitude of the pixel and  $z_{min}$  the lowest altitude of the SR.

#### 3.3.2. SR hypsometric relationships

The method described in section 3.3.1 only allows the retrieval of discrete Volume-Area (V-A) data for each tank. Previous studies derived continuous relationships by regression on V-A points with polynomial models [8,9,19] or power models [20,51]. In this study, we had access to very high resolution DEMs that resolved fine details of the RHS geometries, such as pits of a few meters deep at the bottom of dykes and narrow dyke summits often covered by a road of 20-m large. Pits are dug from year to year to remove clay, to access

the remaining water late in the dry season and to improve the seepage into the aquifer. From this very high resolution dataset, we have preferred to build a “true” hypsometry relationship using a linear interpolation between a subsampling of elevation each 10 cm, which accounts for the SR heterogeneous geometries found in the study site, rather than a global interpolation of volume-area relationship using power or polynomial models. This relationship was used for Pleiades 2-m DEM and all the other DEMs.

### 3.3.3. Temporal dynamic of surface water storage at scene scale

Time series of surface water area extents derived from Sentinel-2 were used to retrieve the volume stored for each Sentinel-2 acquisition date using hypsometric relationships obtained for each SR. Each water body detected is associated to a SR by intersection with the Sentinel-2 MWAE maps (produced as explained in section 3.2.3), and the surface water area is used to retrieve the volume with the hypsometry relationship. As a consequence of this volume retrieval method, there is no need of a perfect spatial match between both data sets, i.e. Pleiades DEM isolines and Sentinel-2 10-m water maps. V-A values of individual SR were summed at scene scale to obtain time series of volume estimates of the RHS for the four Pleiades scenes.

### 3.4. Validation metrics

Three metrics were used to evaluate the fit between the series of predicted values  $(\hat{y}_i)_{1 \leq i \leq N}$  and reference values (derived from Pleiades)  $(y_i)_{1 \leq i \leq N}$  for each SR. The coefficient of determination ( $R^2$ ) is an indicator of the match between observed and predicted values :

$$R^2 = 1 - \frac{\sum_{i=1}^N (\hat{y}_i - y_i)^2}{\sum_{i=1}^N (y_i - \bar{y})^2}, \quad (2)$$

where  $\bar{y}$  is the average of observed values. The mean absolute percentage error (MAPE) is a measure of the relative error :

$$\text{MAPE} = \frac{1}{N} \sum_{i=1}^N \frac{|y_i - \hat{y}_i|}{y_i}, \quad (3)$$

and the mean percentage error (MPE) is a measure of the relative bias :

$$\text{MPE} = \frac{1}{N} \sum_{i=1}^N \frac{y_i - \hat{y}_i}{y_i}. \quad (4)$$

A positive (respectively negative) MPE reflects an average underestimation (respectively overestimation) of predicted values against true values. Note that high relative uncertainty for smaller volumes can increase MAPE and MPE despite a low absolute error, and therefore these two metrics must be carefully interpreted.

## 4. Results

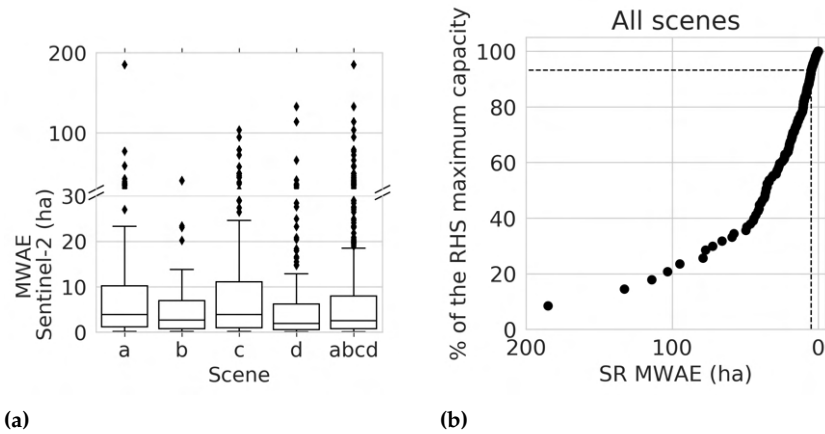
This section is articulated in two parts. Section 4.1 compares the SR geometries and hypsometry relationships derived from each DEMs. Section 4.2 focuses on the accuracy of aggregated water volume retrieval from Sentinel-2 water area extents time series in each DEM footprint.

### 4.1. Comparison of SR hypsometry relationships from selected available DEMs

We used the bathymetry from Pleiades 2-m DEM to estimate the potential capacity of the RHS : it is estimated at 57.6 million cubic meters (32 mm) over the 1,813 km<sup>2</sup> study area. The majority of the RHS is composed of SR with a MWAE lower than 5 ha (as shown in the distribution in Figure 5a), yet these very small reservoirs represent only between 4 to 11%

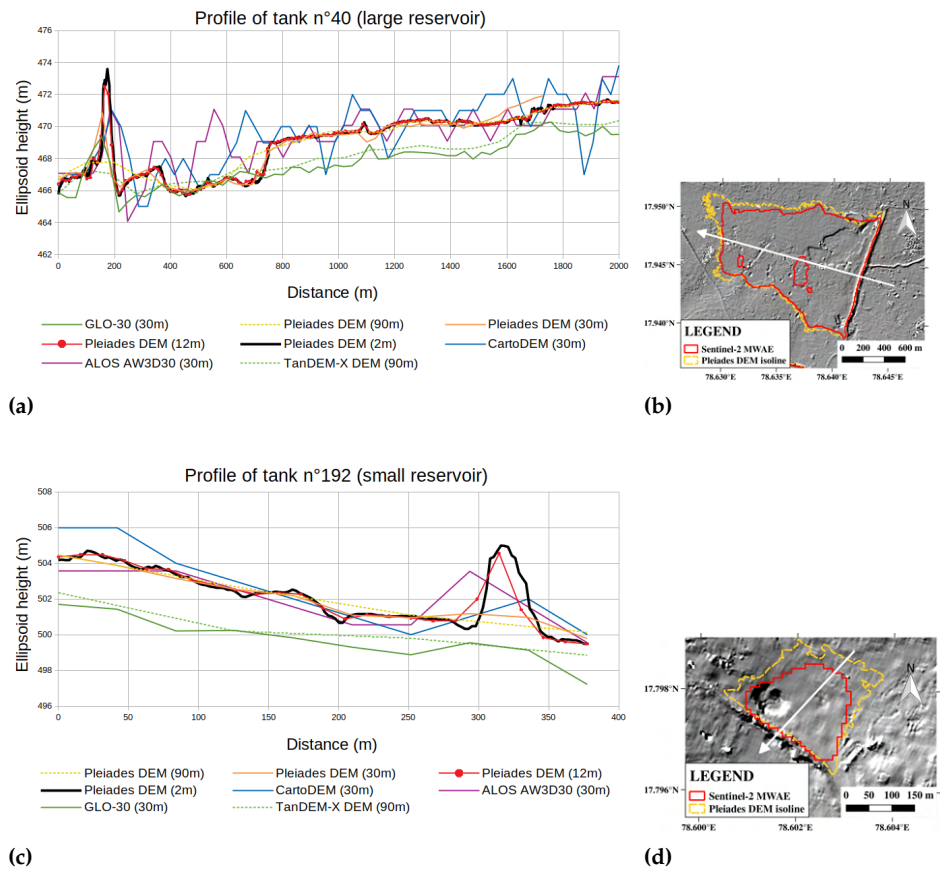


of the total volume of the RHS on individual Pleiades scenes, and 7% on all four Pleiades scenes (as illustrated in Figure 5b).



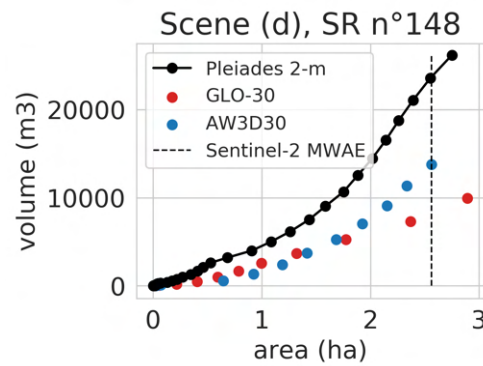
**FIGURE 5.** (a) Distribution (median and quartiles) of Sentinel-2 Maximum Surface Area Extent (MWAE) of the SR per Pleiades scene. (b) Cumulative volumes of the SR of all four Pleiades scenes, in % of the RHS maximum capacity, sorted by decreasing MWAE. SR with an area below 5 ha represents 7% of the RHS maximum capacity (as indicated by the dashed lines).

Apart from the DEM acquisition timing, i.e. the presence or absence of water – anterior to creation or rejuvenation – each source of DEM acquisition have diverse capacity to retrieve the SR geometries. Figure 6 illustrates those differences for two SR that were empty during the acquisition of all DEMs. Coarse resolution smoothes the dyke height and the bottom bathymetry (particularly small depressions at the bottom of SR). This leads to systematic volume underestimation. The dyke of larger reservoirs (as the 114 ha reservoir in Figure 6a) is well detected by Pleiades at 2- and 12-m. The degradation of Pleiades to a 30-m resolution decrease the dyke height by 2, which is globally the same height visible by GLO-30, AW3D30 and CartoDEM. At a 90-m resolution (Pleiades and TanDEM-X) the dyke is nearly flattened. The dyke and the bathymetry of smaller reservoirs (as the 3 ha reservoir in Figure 6c) are almost completely smoothed at a 30-m resolution. The global slopes are coherent between each source of data, with a absolute bias that does not affect geometries, except for CartoDEM and AW3D30 elevation, that have been rounded to the meter before data dissemination. The maximum capacities of SR larger than 5 ha are underestimated by 5, 10 and 14% for Pleiades resampled at 12, 30 and 90 m, against 10, 26 and 39% for SR smaller than 5 ha.

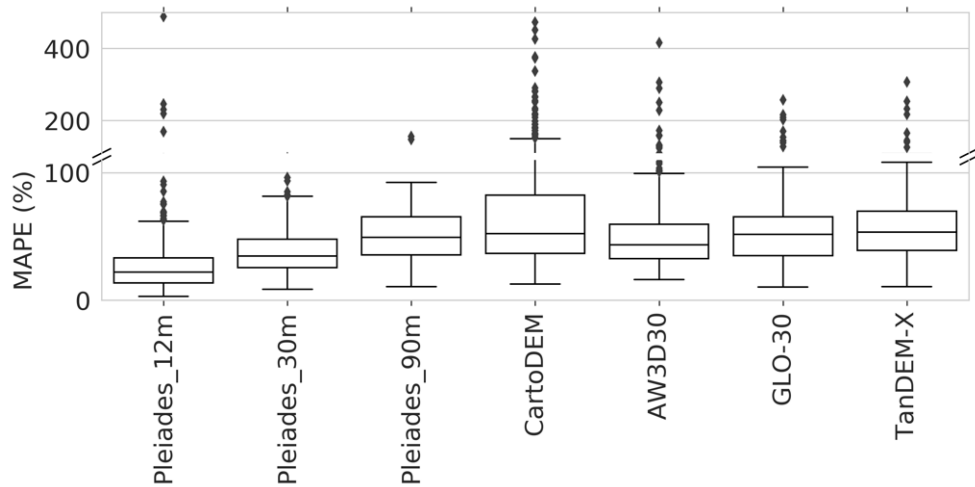


**FIGURE 6.** Profiles of 2 SR for Pleiades at 2-m, 12-m, 30- and 90-m resolution, CartoDEM, AW3D30, GLO-30 and TanDEM-X DEM. Both reservoirs were empty during all DEM acquisitions. (a) and (c) : Profiles of a large (114 ha) and a small (3 ha) reservoir. y values in vertical axis are ellipsoidal heights referenced to the WGS84-G1150 ellipsoid [52] which is the native projection of Pleiades, CartoDEM and TanDEM-X DEM. AW3D30 and GLO-30 are altitudes referenced to the EGM96 geoid [53] and were converted to ellipsoid height for this figure [54]. (b) and (d) : Context of the profile. The map background is Pleiades 2-m DEM hillshade, the red line is Sentinel-2 detected MWAE (Maximum Water Area Extent), the dashed blue line the manually selected Pleiades DEM isoline, and the white arrow is the transect.

We quantified the decrease of SR volume estimates from water area extents for large-scale and resampled DEMs. Figure 7 shows an example for a given SR : the series of volumes of GLO-30 and AW3D30 (red and blue points), computed with a 10 cm height increment, are compared with the corresponding water volumes given by the reference hypsometric relationship of Pleiades 2-m DEM (black curve, computed as explained in section 3.1) for the same area. Volumes of individual SR were generally associated with high uncertainties : the median MAPE of SR volumes (shown in Figure 8) over all SR ranges from 43 to 53% for large-scale DEMs, against 22%, 35% and 49% for Pleiades aggregated at 12, 30 and 90 m.



**FIGURE 7.** V-A pairs of a SR (2.6 ha) for three DEMs. The hypsometric relationship derived from Pleiades 2-m DEM is in black, linearly interpolated between black dots. Red and blue dots are the V-A pairs of GLO-30 and AW3D30.



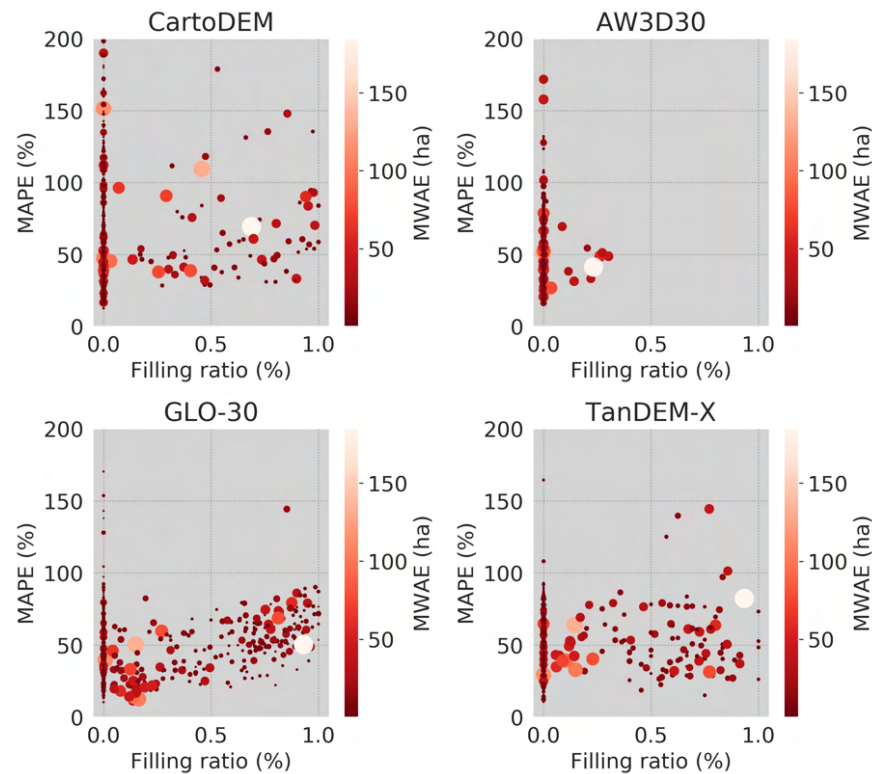
**FIGURE 8.** Distribution (median and quartiles), over all small reservoirs (SR), of the mean absolute percentage error (MAPE) of SR volumes extracted from several DEMs at a 10 cm increment, relative to Pleiades 2-m DEM reference volume.

Apart from resolution and numerical precision, another major requirement for bathymetry retrieval and essential for accurate volume estimation is the absence of water in the SR during DEM acquisition. The filling ratio of the four large-scale DEMs, is reported in Table 2. CartoDEM, GLO-30 and TanDEM-X DEM show high (>33%) filling ratios for scenes (a) and (b). AW3D30 is the only DEM with a low filling ratio on the four scenes (<8%).

We examined the impact of this filling ratio on the estimation of the volumes of SR. Figure 9 shows the MAPE of all SR plotted against their individual filling ratio during DEM acquisition for CartoDEM, AW3D30, GLO-30 and TanDEM-X DEM. The estimation error seems to be correlated to the filling ratio for GLO-30 and AW3D30, slightly less for CartoDEM and TanDEM-X DEM. However, a great variability of errors is visible even for empty SR (filling ratio = 0%), in particular for the smaller ones (little dark red dots, often smaller than 5 ha). This confirms the inability of these four DEMs to correctly measure the volume of smaller SR because of their coarse resolution.

**TABLE 2.** Filling ratio (%) of the observed Rainwater Harvesting System (RHS) at the acquisition date(s) for CartoDEM, AW3D30, GLO-30 and TanDEM-X DEM. It is computed from the water masks provided with the DEMs and Sentinel-2 Maximum Water Area Extent (MWAE).

Scene	CartoDEM	AW3D30	GLO-30	TanDEM-X DEM
a	54	8	78	68
b	33	3	64	50
c	5	0	11	8
d	9	0	14	10



**FIGURE 9.** MAPE of small reservoirs (SR) in relation to their filling ratio at the moment of the DEM acquisition. Each point represents one SR. The color and size of each point correspond to the Sentinel-2 Maximum Water Area Extent (MWAE) of the SR in ha.

#### 4.2. Volume retrieval from Sentinel-2 time series at Pleiades scene scale

Sentinel-2 time series were used to estimate the water volume stored in each tank covered by Pleiades DEM at different hydrological period and filling rate. These volumes, estimated with each DEM derived hypsometric relationship, were compared to the reference, i.e volumes from the hypsometric relationships derived from Pleiades 2-m DEM. The average MAPE and MPE computed on the time series of the volumes of individual SR are reported in Table 3 (the MAPE and MPE were computed on the time series for each SR, then averaged). MAPEs are high for large-scale DEMs, ranging from 52% to 188%. Even the resampling of Pleiades DEM at 12 m induces high error, ranging from 22 to 43%. This shows that large-scale DEMs cannot be used to monitor the volume of individual SR. Note that these MAPEs estimated from Sentinel-2 time series are larger than those estimated from all V-A sample points in section 4.1. Indeed, water area extents detected by Sentinel-2 over time are not equally distributed on the V-A relationship as higher areas rarely occur, and small volumes estimates that have the largest relative uncertainty.

**TABLE 3.** Average MAPE (%) and MPE (%) of volumes computed with Sentinel-2-derived water area extents time series for each SR.

Image	Pleiades 12m		Pleiades 30m		Pleiades 90m		CartoDEM		AW3D30		GLO-30		TanDEM-X	
	MAPE	MPE	MAPE	MPE	MAPE	MPE	MAPE	MPE	MAPE	MPE	MAPE	MPE	MAPE	MPE
a	22.4	19.9	34.4	27.6	44.5	38.9	108.3	-21.9	52.2	7.7	121.9	-1.6	52.5	21.2
b	30.0	23.8	41.2	32.7	53.1	39.5	123.7	-62.0	57.1	0.7	188.2	-83.4	75.1	17.5
c	43.4	30.2	50.4	47.7	67.7	50.5	72.8	-12.8	58.6	-0.2	106.4	-21.7	65.1	50.3
d	25.7	20.8	41.4	33.4	66.7	43.7	96.3	-37.2	81.2	-4.3	111.6	-25.1	64.9	45.7
all	29.8	23.3	41.9	35.5	60.2	43.8	96.5	-30.5	66.0	0.1	121.6	-25.5	63.2	37.8

**TABLE 4.**  $R^2$  (%), MAPE (%) and MPE (%) of the time series of volumes lumped at scene scale with Sentinel-2 areas. The last line presents the metrics computed on the series of V-A points of all scenes.

Image	Pleiades 12m			Pleiades 30m			Pleiades 90m			CartoDEM			AW3D30			GLO-30			TanDEM-X		
	$R^2$	MAPE	MPE	$R^2$	MAPE	MPE	$R^2$	MAPE	MPE	$R^2$	MAPE	MPE	$R^2$	MAPE	MPE	$R^2$	MAPE	MPE	$R^2$	MAPE	MPE
a	99.4	4.4	4.4	98.5	7.0	7.0	97.3	10.2	10.1	8.4	57.1	57.1	95.2	11.9	11.1	-54.2	72.1	72.1	95.9	12.3	7.2
b	99.2	18.0	18.0	98.1	29.1	29.1	94.6	43.0	43.0	77.3	48.4	48.4	97.4	26.1	20.0	70.1	69.1	69.1	83.7	36.9	36.9
c	97.4	28.0	28.0	95.2	37.9	37.9	91.9	46.6	46.6	96.2	18.8	5.4	95.9	20.6	-12.5	98.8	28.5	28.2	90.5	47.9	47.9
d	99.4	10.2	10.2	97.8	29.4	29.4	91.9	44.3	44.3	93.4	15.3	-5.9	80.3	16.3	7.7	85.0	38.4	38.4	71.8	57.2	57.2
all	99.0	16.2	16.2	97.5	28.1	28.1	94.0	38.6	38.6	81.7	29.0	17.4	90.0	18.2	4.6	69.1	45.8	45.7	85.2	42.8	41.9

However, aggregating individual volumes at the Pleiades scene scale improves their correspondence with volumes derived from Pleiades 2 m DEM. Figure 10 shows lumped volume estimates from the eight DEMs. Only a subset of the Sentinel-2 time series is shown for better visibility. High values ( $>100\%$ ) of volume relative uncertainty were found for very small areas ( $<1\%$  of the maximum RHS area). These outliers can dramatically increase MAPE and MPE despite representing a low absolute uncertainty, and were therefore removed from the calculation of the metrics presented in this section (16 points for AW3D30 in scene (c)). The  $R^2$ , MAPE and MPE of scene-scale volumes, computed over the time series, are reported in Table 4.

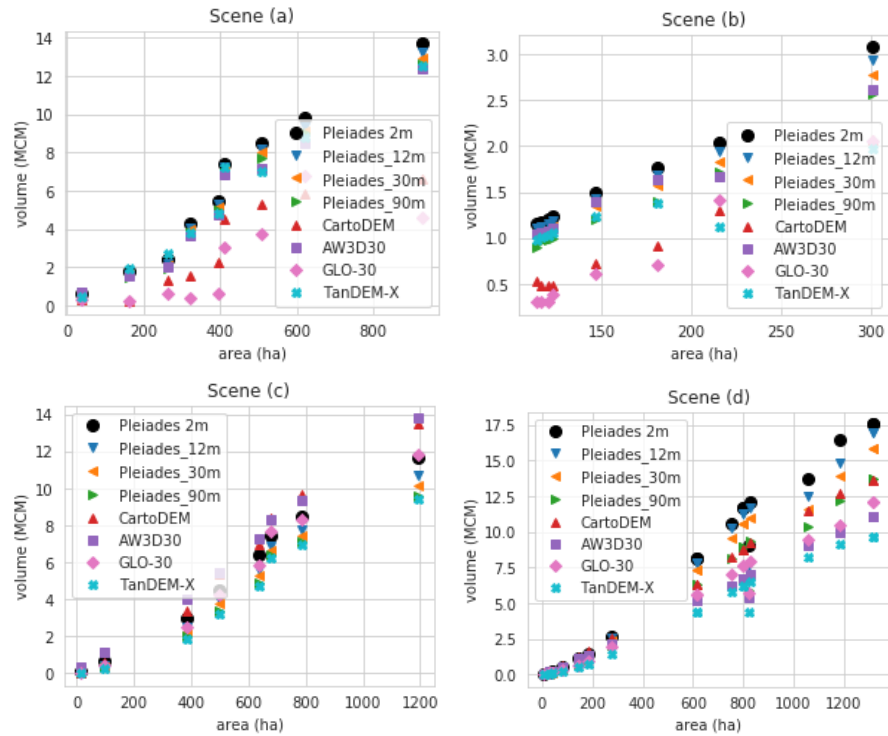


FIGURE 10. Lumped volume and area for all 4 Pleiades scenes. Only a subset of the time series is plotted. Reference values are black dots.

The MAPEs of aggregated values at the Pleiades scene scale (see Table 4) are lower than for individual SR (12% against 72%), which shows that the errors on SR compensate each other and the aggregation of small volumes produce better regional volume estimates than with each individual SR.

On the four areas covered by Pleiades acquisitions, volume estimates are less accurate ( $R^2$  decreases) with the reduction of Pleiades resolution (Table 4). As expected, a lower resolution induces a systematic negative bias ( $MPE > 0$  and  $MAPE \approx MPE$ ) on the four Pleiades scenes, as the the dykes geometries are flattened. Nevertheless, these biases increase drastically above 30m. Indeed, the relative uncertainty reaches 28% on average for a 30 m resolution.

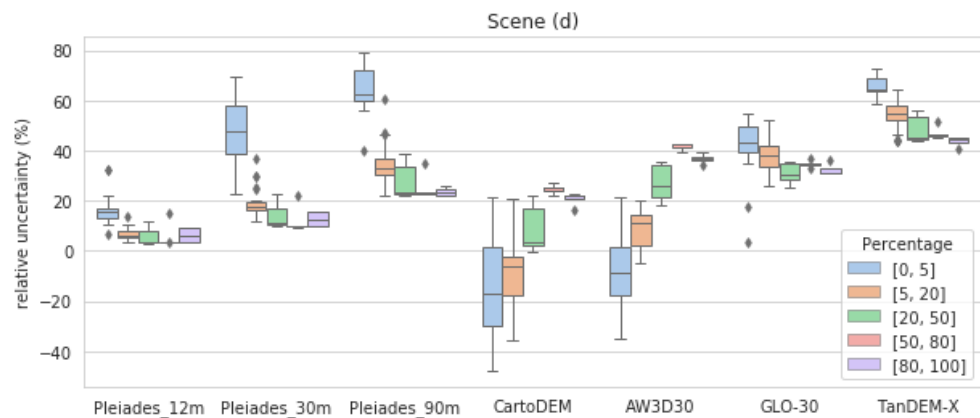
The filling rate of the RHS, i.e. the percentage of the RHS maximum area covered by surface water at the time of acquisition, is high for CartoDEM, GLO-30 and TanDEM-X ( $>33\%$ , for scenes (a) and (b), Table 2), which impacts their performances compare to AW3D30 (filling ratio  $<8\%$ ). GLO-30 has the largest filling ratio (78% and 64% for scenes (a) and (b)) and the least accurate estimates (Table 4) with  $R^2 = -54\%$  and 70% for scenes (a) and (b). Figure 10 shows that those DEM lead to large underestimation of volumes and thus are not suitable for RHS volume monitoring in this region. TanDEM-X DEM performs better

than CartoDEM (higher  $R^2$ ) despite its lower resolution and higher filling ratio. AW3D30 volume estimates are similar to those of the reference Pleiades DEM resampled at 30-m for scenes (a) and (b), i.e. it can be used at the Telangana scale taking the bias presented in this study into account.

On scenes (c) and (d) the filling ratio is lower than 11% for all DEM. CartoDEM, AW3D30 and GLO-30 estimates are relatively good ( $R^2$  from 80 to 99%). The coarser TanDEM-X DEM produces the least accurate volume estimates. GLO-30 produces the best estimates for image (c) ( $R^2=98.8\%$ ) among all DEMs including Pleiades resampled at 30-m, and CartoDEM performs better than the three other large-scale DEMs and Pleiades 30-m for image (d) ( $R^2=96.5\%$ ), despite the fact that AW3D30 has the lower filling ratio for these two scenes.

For almost all scenes and DEMs, volumes are generally estimated with a negative bias, except for CartoDEM and AW3D30 in scene (c) and (d), for which the MPE can even be negative.

Volume estimates have different accuracies for low and high filling rates. Relative uncertainties reaches their maximum (80% and -50%) for smallest water area extent (<5% of the total RHS surface area) and decreases for higher filling rates (Figure 11). Contrary to resampled Pleiades DEM, GLO-30 and TanDEM-X DEM, for which increasing volumes are estimated with a decreasing negative bias, CartoDEM and AW3D30 overestimate low volumes and underestimate large one.



**FIGURE 11.** Relative uncertainty of volume estimates at scene scale for scene (d). The V-A points of the time series of all four scenes are gathered by their proportion of surface water area relative to the maximum area of the RHS of scene (d).

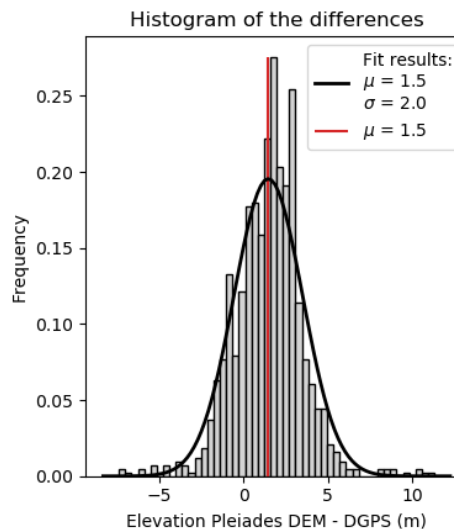
## 5. Discussion

This section discusses and compares the ability of each selected available global DEM to retrieve the water volumes of SR using surface water detection time series (here Sentinel-2 images). The main factors impacting the retrieval accuracy, i.e. surface water extent accuracy, DEM accuracy, DEM spatial and numerical resolution, filling ratio and SR size are explored regarding the thematic objectives related to the monitoring of RHS storage at SR level or regional level.

### 5.1. Pleiades DEM accuracy

In this article, we quantified the difference between water volumes estimated from coarse DEM and Pleiades DEM at 2 m. The accuracy of the reference Pleiades DEM elevation was estimated with 1,146 DGPS measurements on four different tanks. The histogram of the differences between the DGPS elevation and the corresponding Pleiades DEM elevation is shown in figure 12. A normal distribution was fit to the data with an average of 1.5

m and a standard deviation of 2.0 m, which can be a measure of the punctual vertical error. Note that this is an estimation of the pixel level accuracy (i.e. at 2 m resolution). The uncertainty is expected to be reduced by spatial aggregation, i.e. when computing the volume of tanks [55]. The RMSE and the absolute vertical accuracy are estimated at 2.5 m and 1.6 m, respectively. In comparison, the RMSE of AW3D30 is 5 m and ranges between 3.4 and 4.72 for CartoDEM, and the absolute vertical accuracy is 2 m for TanDEM-X DEM and ranges from 4.7 to 7.3 m for CartoDEM. The error of Pleiades DEM elevation is lowest than the other DEM despite the 10 year gap between DGPS and Pleiades acquisition, although it should be noted that the accuracies computed for Pleiades DEM and coarser DEM are not directly comparable.



**FIGURE 12.** Comparison of Pleiades DEM and differential-GPS (DGPS) measurements. (a) Histogram of the differences between Pleiades DEM and DGPS measurements elevation. A normal distribution is fit to the data with a mean  $\mu$  (vertical red line) and standard deviation  $\sigma$  of 1.5 m and 2.0 m, respectively.

### 5.2. Surface water extent accuracy

Sentinel-2 surface water areas are often underestimated because of (i) the different resolution and (ii) the well known imperfections of water detection methods from multispectral data, especially in seasonal shallow temporary lakes, on the edge or caused by vegetation [56]. The comparison of Sentinel-2-derived water maps with higher resolution images (as explained in section 3.2.2) showed an average negative bias on volume estimates : the average relative error (MAPE) was estimated at 17% on water areas and 31% on volumes, which is in accordance with the 15% negative bias found in Perin et al. [57]. This bias could be reduced by detecting water on optical images with a higher spatial resolution than Sentinel-2, such as Venus (5.3 m resolution, 2 days revisit time) [58]. However, the use of optical images only allows the retrieval of a partial time series because of missing images during cloudy periods. Radar observations such as those from Radarsat-2, TanDEM-X, or Sentinel-1 can provide dense time series but underestimate surface water coverage. Previous studies quantified that using Sentinel-1 data rather than optical data could underestimate reservoirs water areas by 15% [25] and water volumes by 27% [20].

However, in the particular case of this study, the accuracy of water maps is not the focus. Water maps are used to simulate several filling rates of the RHS for the specific purpose of quantifying the sensitivity of volume retrieval to DEM features at scene scale. We



chose to use water maps produced from optical images, yet we could also have generated random surface water area extents for each SR.

### 5.3. Hypsometry relationship

The construction of individual V-A relationship relies on a vertical subsampling of each DEM. This can be challenging for (i) coarse DEMs for which the number of pixels by SR is low, (ii) DEMs rounded to the meter, (iii) DEMs for which the depth of SR is lowered by the presence of water at the time of acquisition. The bilinear resampling of DEMs at a finer resolution (2-m in this study) allows a better vertical sampling by solving issues (i) and (ii). The density of V-A points remains nonetheless inferior for large-scale DEMs than for Pleiades 2-m DEM (as illustrated in Figure 7), producing a “sparser” hypsometric curve and reducing the accuracy of volume estimation from Sentinel-2 images. Existing methods applied for medium to large dams have been used to extrapolate water volume below the water surface detected by satellite data based DEMs [27]. We have chosen to not use them as they add more source of differences between the use of selected DEMs. This explains a part of the systematic underestimation of SR volumes presented in this study.

### 5.4. Heterogeneous negative biases of volume restitutions from freely available DEMs

We have quantified the impact of (i) the spatial resolution of the DEM, (ii) the SR filling rate during its acquisition and (iii) its numerical precision (float or integer format) on volume estimates. Volumes retrieved from DEMs with low resolution and high filling ratio are generally underestimated (up to 72% at scene scale with a filling ratio of 78% for GLO-30), while integer DEMs induce unpredictable uncertainties. Indeed, the discrepancies with Pleiades 2-m DEM are enhanced by the rounding of elevation values, which can either under or overestimate volumes of individual SR depending if the rounding is upward or downward. The systematic negative bias due to coarse resolution can be compensated by upward rounding errors, which eventually and incidentally improve volume estimates at scene scale. This explains why volume estimates from CartoDEM and AW3D30 at scene scale are occasionally better than Pleiades 30-m DEM, which is supposedly more accurate due to its acquisition period and its higher native resolution.

The long period between DEMs acquisition dates can also impact volume estimates. SR bathymetries can evolve quickly, especially on the influence of public policies for SR rejuvenation that started in 2015 to increase the water total storage (Kakatiya program). It should be relevant to monitor those changes at least once in a decade with new VHR DEM acquisitions during dry months to update SR bathymetries and monitor new ones. Note that we chose not to include in this study the DEM from the Shuttle Radar Topography Mission (SRTM) [59] based on 2000 acquisitions, that we estimated too ancient and obsolete regarding the RHS observed with Pleiades acquisitions (16 and 19 years) given those recent changes that occurred in the study area.

### 5.5. SR delineation

As presented in section 3.2.3, the hypsometry retrieval process is highly dependent on the agreement between SR limits and the DEM isolines. The use of optical satellite data such as Sentinel-2-derived MWAE (Maximum Water Area Extent) for SR delineation is a fast semi-automatic process. Nevertheless, the discrepancy between the surface water area extents and pleiades DEM isolines imposes to manually delineate SR with Pleiades DEM isoline, a time-consuming task that can not be reproduced at large scales. Note that this task was fully automated in Arendt et al. [28] and [50] by using a Blue Spot detection algorithm, which identifies natural depressions subject to flooding directly from a DEM. In particular, Arendt et al. [28] used this method in the Cuvelai Basin, a semi-arid area flooded part of the year, to determine a high resolution theoretical MWAE of natural depressions that were measured using TanDEM-X DEM commercial product at 12-m resolution and estimated their potential water storage. A combination of the Blue Spot detection algorithm

and surface water maps derived from Sentinel-2 could be a lead for the automatic retrieval of active SR limits.

### 5.6. Monitoring the RHS of Telangana : what DEM and what uncertainty?

This study shows that VHR DEM resampled at 12 or 30 m enables to estimate water volumes from surface water area with an error (i.e. MAPE computed with the native 2 m resolution as a reference) less than 16 and 28% at the scale of the study area, with higher errors for lowest volumes and smallest water bodies. This error ranges from 18 to 45% for the large-scale DEMs. On the other hand, the quantification of the 15 main largest and highly filled SR of the RHS would enable to quantify 40% of this potential water resource (figure 5b). This can be particular relevant as the volume of these largest reservoirs can represent a higher portion of the RHS volume during the dry season, as the largest reservoirs are generally the last ones to empty. At regional scale, this amount of water is small compare to the groundwater and large dams resources, but remain important for local farmers in upstream areas disconnected from the main large rivers and irrigation water.

## 6. Conclusions and perspectives

The small reservoirs (SR) composing the Rainwater Harvesting System (RHS) are a crucial water resource for irrigation and livestock in upstream rural areas in South India. We used the bathymetry from a 2-m Digital Elevation Model (DEM) over a 1,800 km<sup>2</sup> area to derive volume-area relationships for more than 500 SR. Times series of Sentinel-2-derived surface water area over several years were used to simulate several filling rates of the RHS and compute volume estimates at larger scales. We used these retrievals as a reference to quantify the uncertainty of RHS volume retrieval from 3 large-scale DEMs at 30-m resolution (AW3D30, GLO-30, CartoDEM) and one DEM at 90-m resolution (TanDEM-X DEM).

Volume estimates at SR scale were associated with high relative uncertainties, ranging from 47% to 78% for large-scale DEMs against 17%, 29% and 46% for Pleiades DEM aggregated at 12, 30 and 90 m. Volume estimates lumped at larger scale have better accuracies because of the compensation of the uncertainties of individual SR. We found that at scene scale, the Pleiades DEM at 12-m resolution could estimate the RHS volume with a R<sup>2</sup> from 97.4% to 99.4% and a MAPE (Mean Absolute Percentage Error) from 4% to 28%, but with higher uncertainties in dry conditions (low RHS filling rates). The Pleiades DEM aggregated at 30-m was able to estimate the RHS volume at scene scale with a R<sup>2</sup> from 95.2 to 98.5% and a MAPE from 7 to 38%. However, volumes are estimated with a quasi-systematic negative bias when the resolution decreases due to (i) the underestimation of the volume of very small reservoirs (<5ha) that are hardly visible by larger DEM and (ii) the smoothing of the SR geometry. The use of a 90-m resolution DEM degrades volume estimates even more, and we recommend using a resolution no coarser than 30 meters and to focus on the largest SR composing the major part of the RHS volume in this region.

Results indicate that the acquisition period of DEM is crucial, i.e. the filling ratio of the RHS at the time of acquisition. AW3D30 had the lowest RHS filling ratio of all DEM (<6% on all 4 scenes) and hence showed the best results on average over the 4 scenes (R<sup>2</sup> from 80.3 to 97.4% and MAPE from 8 to 26%). On the other hand, CartoDEM and GLO-30 were able to produce better results at scene scale for a low filling ratio (<10%). In absence of VHR DEMs, the best strategy is to use a 30-m global DEM acquired recently, or a combination of several DEM with a low filling ratio (<10%) covering the study area, to approximate the volume of the RHS at regional scale from time series of surface water maps. Yet an average relative uncertainty up to 26% can be expected. However, at very local or SR scale, or for low water stocks, high-resolution DEMs are essential to estimate and monitor surface water storage with precision in near-real time in combination with high resolution optical observations. Unlike the next hydrological SWOT mission that will not be able to monitor such small water bodies, upcoming very high resolution stereoscopy satellite missions

such as CO3D [31] will provide large-scale accurate DEMs with lower costs. It is seen to be essential in the next years to build a near-real time early warning hydrological drought systems in India.

This study was performed on the RHS located in an area of 1,813 km<sup>2</sup> representative of the Telanagana RHS. The estimated maximum storage capacity reaches 57.6 million cubic meters (32 mm), which is representative of the maximum capacity estimated at the scale of the Telangana by Pascal *et al.* [36] (30 mm). The main limitation of the extension of this study at regional scale is the manual SR delineation process. This could be improved by the use of automatic methods such as the Blue Spot detection algorithm Arendt *et al.* [28], Nielsen *et al.* [50] which could delineate a high number of SR in a shorter amount of time, allowing to extend the RHS monitoring at larger scale. This study opens up possibilities for a seasonal monitoring of RHS storage volumes that can be analysed in terms of water inflows (rainfall, surface runoff) and outflows (evaporation, infiltration or agricultural uses) between two measurements dates at regional scale, while these studies have been done only for individual SR so far [6–8]. Those flow components are highly important to quantify in order to properly understand the impact of the RHS on the water cycle (including groundwater additional recharge) in these semi-arid areas, and modelling approaches used to understand those flows could benefit from the very high resolution satellite data such as those used in this study [60].

**Author Contributions:** Conceptualization, Claire Pascal, Sylvain Ferrant, Santiago Peña-Luque, Simon Gascoin and Olivier Merlin; Data curation, Claire Pascal, Sylvain Ferrant and Adrien Selles; Formal analysis, Sylvain Ferrant; Funding acquisition, Sylvain Ferrant; Investigation, Claire Pascal, Sylvain Ferrant, Simon Gascoin and Olivier Merlin; Methodology, Claire Pascal, Sylvain Ferrant, Santiago Peña-Luque, Simon Gascoin and Olivier Merlin; Project administration, Sylvain Ferrant; Software, Claire Pascal, Sylvain Ferrant and Santiago Peña-Luque; Supervision, Sylvain Ferrant and Olivier Merlin; Validation, Claire Pascal, Sylvain Ferrant, Adrien Selles and Jean-Christophe Maréchal; Visualization, Claire Pascal; Writing – original draft, Claire Pascal and Sylvain Ferrant; Writing – review & editing, Claire Pascal, Sylvain Ferrant, Santiago Peña-Luque, Adrien Selles, Jean-Christophe Maréchal, Simon Gascoin and Olivier Merlin. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the Programme National de Télédétection Spatiale (PNTS, grant number PNTS-2020-18).

**Data Availability Statement:** Not applicable.

**Acknowledgments:** This work has been supported by the Programme National de Télédétection Spatiale. Pleiades images acquisition was supported by the CNES and were acquired using the DINAMIS program (Dispositif Institutionnel National d'Approvisionnement Mutualisé en Imagerie Spatiale). Sentinel-2 images were processed by Theia and downloaded on <https://theia.cnes.fr>. This project relies heavily on open-source softwares. The programming was done in OTB 7.0.0, Python 3.7 and associated libraries including numpy, pandas, gdal and matplotlib. This research was realized as part of a PhD thesis funded by a french national research ministry doctoral fellowship.

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

## Abbreviations

The following abbreviations are used in this manuscript :

DEM	Digital Elevation Model
MAPE	Mean Absolute Percentage Error
MPE	Mean Percentage Error
MWAE	Maximum Water Area Extent
RHS	Rainwater Harvesting System
SR	Small Reservoir

## References

1. Wada, Y.; Beek, L.P.H.v.; Viviroli, D.; Dürr, H.H.; Weingartner, R.; Bierkens, M.F.P. Global monthly water stress : 2. Water demand and severity of water stress. *Water Resources Research* **2011**, *47*. <https://doi.org/https://doi.org/10.1029/2010WR009792>. 631-633
2. Habets, F.; Molénat, J.; Carluier, N.; Douez, O.; Leenhardt, D. The cumulative impacts of small reservoirs on hydrology : A review. *Science of The Total Environment* **2018**, *643*, 850–867. <https://doi.org/10.1016/j.scitotenv.2018.06.188>. 634-635
3. Krol, M.S.; de Vries, M.J.; van Oel, P.R.; de Araújo, J.C. Sustainability of Small Reservoirs and Large Scale Water Availability Under Current Conditions and Climate Change. *Water Resources Management* **2011**, *25*, 3017–3026. <https://doi.org/10.1007/s11269-011-9787-0>. 636-638
4. Sakthivadivel, R. The groundwater recharge movement in India. In *The Agricultural Groundwater Revolution : Opportunities and Threats to Development. Comprehensive Assessment of Water Management in Agriculture Series.*; Giordano, M.; Villholth, K.G., Eds.; CABI Publishing : Wallingford, United Kingdom, 2007. 639-640
5. Glendenning, C.J.; van Ogtrop, F.F.; Mishra, A.K.; Vervoort, R.W. Balancing watershed and local scale impacts of rain water harvesting in India—A review. *Agricultural Water Management* **2012**, *107*, 1–13. <https://doi.org/10.1016/j.agwat.2012.01.011>. 642-643
6. Boisson, A.; Baisset, M.; Alazard, M.; Perrin, J.; Villesseche, D.; Dewandel, B.; Kloppmann, W.; Chandra, S.; Picot-Colbeaux, G.; Sarah, S.; et al. Comparison of surface and groundwater balance approaches in the evaluation of managed aquifer recharge structures : Case of a percolation tank in a crystalline aquifer in India. *Journal of Hydrology* **2014**, *519*, 1620–1633. <https://doi.org/10.1016/j.jhydrol.2014.09.022>. 644-647
7. Massuel, S.; Perrin, J.; Mascré, C.; Mohamed, W.; Boisson, A.; Ahmed, S. Managed aquifer recharge in South India : What to expect from small percolation tanks in hard rock? *Journal of Hydrology* **2014**, *512*, 157–167. <https://doi.org/10.1016/j.jhydrol.2014.02.062>. 648-650
8. Perrin, J.; Ferrant, S.; Massuel, S.; Dewandel, B.; Maréchal, J.; Aulong, S.; Ahmed, S. Assessing water availability in a semi-arid watershed of southern India using a semi-distributed model. *Journal of Hydrology* **2012**, *460-461*, 143–155. <https://doi.org/10.1016/j.jhydrol.2012.07.002>. 651-653
9. Baup, F.; Frappart, F.; Maubant, J. Combining high-resolution satellite images and altimetry to estimate the volume of small lakes. *Hydrology and Earth System Sciences* **2014**, *18*, 2007–2020. <https://doi.org/10.5194/hess-18-2007-2014>. 654-655
10. Crétaux, J.F.; Arsen, A.; Calmant, S.; Kouraev, A.; Vuglinski, V.; Bergé-Nguyen, M.; Gennero, M.C.; Nino, F.; Abarca Del Rio, R.; Cazenave, A.; et al. SOLS : A lake database to monitor in the Near Real Time water level and storage variations from remote sensing data. *Advances in Space Research* **2011**, *47*, 1497–1507. <https://doi.org/10.1016/j.asr.2011.01.004>. 656-658
11. Duan, Z.; Bastiaanssen, W.G.M. Estimating water volume variations in lakes and reservoirs from four operational satellite altimetry databases and satellite imagery data. *Remote Sensing of Environment* **2013**, *134*, 403–416. <https://doi.org/10.1016/j.rse.2013.03.010>. 659-661
12. Frappart, F.; Minh, K.D.; L’Hermitte, J.; Cazenave, A.; Ramillien, G.; Le Toan, T.; Mognard-Campbell, N. Water volume change in the lower Mekong from satellite altimetry and imagery data. *Geophysical Journal International* **2006**, *167*, 570–584. <https://doi.org/10.1111/j.1365-246X.2006.03184.x>. 662-664
13. Tachikawa, T.; Hato, M.; Kaku, M.; Iwasaki, A. Characteristics of ASTER GDEM version 2. In *Proceedings of the 2011 IEEE International Geoscience and Remote Sensing Symposium*, 2011, pp. 3657–3660. ISSN : 2153-7003, <https://doi.org/10.1109/IGARSS.2011.6050017>. 665-667
14. Krieger, G.; Zink, M.; Bachmann, M.; Bräutigam, B.; Schulze, D.; Martone, M.; Rizzoli, P.; Steinbrecher, U.; Walter Antony, J.; De Zan, F.; et al. TanDEM-X : A radar interferometer with two formation-flying satellites. *Acta Astronautica* **2013**, *89*, 83–98. <https://doi.org/10.1016/j.actaastro.2013.03.008>. 668-670
15. Tadono, T.; Ishida, H.; Oda, F.; Naito, S.; Minakawa, K.; Iwamoto, H. Precise Global DEM Generation by ALOS PRISM. In *Proceedings of the ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*. Copernicus GmbH, 2014, Vol. II-4, pp. 71–76. ISSN : 2194-9042, <https://doi.org/10.5194/isprsannals-II-4-71-2014>. 671-673
16. Bonnema, M.; Hossain, F. Inferring reservoir operating patterns across the Mekong Basin using only space observations. *Water Resources Research* **2017**, *53*, 3791–3810. <https://doi.org/https://doi.org/10.1002/2016WR019978>. 674-675
17. Li, Y.; Gao, H.; Jasinski, M.F.; Zhang, S.; Stoll, J.D. Deriving High-Resolution Reservoir Bathymetry From ICESat-2 Prototype Photon-Counting Lidar and Landsat Imagery. *IEEE Transactions on Geoscience and Remote Sensing* **2019**, *57*, 7883–7893. <https://doi.org/10.1109/TGRS.2019.2917012>. 676-678
18. Avisse, N.; Tilmant, A.; Müller, M.F.; Zhang, H. Monitoring small reservoirs’ storage with satellite remote sensing in inaccessible areas. *Hydrology and Earth System Sciences* **2017**, *21*, 6445–6459. <https://doi.org/10.5194/hess-21-6445-2017>. 679-680
19. Bhagwat, T.; Klein, I.; Huth, J.; Leinenkugel, P. Volumetric Analysis of Reservoirs in Drought-Prone Areas Using Remote Sensing Products. *Remote Sensing* **2019**, *11*, 1974. <https://doi.org/10.3390/rs11171974>. 681-682
20. Vanthof, V.; Kelly, R. Water storage estimation in ungauged small reservoirs with the TanDEM-X DEM and multi-source satellite observations. *Remote Sensing of Environment* **2019**, *235*, 111437. <https://doi.org/10.1016/j.rse.2019.111437>. 683-684
21. Pekel, J.F.; Cottam, A.; Gorelick, N.; Belward, A.S. High-resolution mapping of global surface water and its long-term changes. *Nature* **2016**, *540*, 418–422. Number : 7633 Publisher : Nature Publishing Group, <https://doi.org/10.1038/nature20584>. 685-686
22. Pickens, A.H.; Hansen, M.C.; Hancher, M.; Stehman, S.V.; Tyukavina, A.; Potapov, P.; Marroquin, B.; Sherani, Z. Mapping and sampling to characterize global inland water dynamics from 1999 to 2018 with full Landsat time-series. *Remote Sensing of Environment* **2020**, *243*, 111792. <https://doi.org/10.1016/j.rse.2020.111792>. 687-689

23. Cordeiro, M.C.R.; Martinez, J.M.; Peña-Luque, S. Automatic water detection from multidimensional hierarchical clustering for Sentinel-2 images and a comparison with Level 2A processors. *Remote Sensing of Environment* **2021**, *253*, 112209. <https://doi.org/10.1016/j.rse.2020.112209>. 690  
691
24. Martinis, S.; Plank, S.; Ćwik, K. The Use of Sentinel-1 Time-Series Data to Improve Flood Monitoring in Arid Areas. *Remote Sensing* **2018**, *10*, 583. Number : 4 Publisher : Multidisciplinary Digital Publishing Institute, <https://doi.org/10.3390/rs10040583>. 692  
693  
694
25. Peña-Luque, S.; Ferrant, S.; Cordeiro, M.C.R.; Ledauphin, T.; Maxant, J.; Martinez, J.M. Sentinel-1&2 Multitemporal Water Surface Detection Accuracies, Evaluated at Regional and Reservoirs Level. *Remote Sensing* **2021**, *13*, 3279. <https://doi.org/10.3390/rs13163279>. 695  
696  
697
26. Magome, J.; Ishidaira, H.; Takeuchi, K. Method for satellite monitoring of water storage in reservoirs for efficient regional water management. In Proceedings of the Water Resources Systems-Hydrological Risk, Management and Development; Bloschl, G.; Franks, S.; Kumagai, M.; Musiak, K.; Rosbjerg, D., Eds.; Int Assoc Hydrological Sciences : Wallingford, 2003; pp. 303–310. ISSN : 0144-7815 Issue : 281 WOS :000189482100037. 698  
699  
700  
701
27. Liu, K.; Song, C.; Wang, J.; Ke, L.; Zhu, Y.; Zhu, J.; Ma, R.; Luo, Z. Remote Sensing-Based Modeling of the Bathymetry and Water Storage for Channel-Type Reservoirs Worldwide. *Water Resources Research* **2020**, *56*, e2020WR027147. [\\_eprint : https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2020WR027147](https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2020WR027147), <https://doi.org/10.1029/2020WR027147>. 702  
703  
704
28. Arendt, R.; Reinhardt-Imjela, C.; Schulte, A.; Faulstich, L.; Ullmann, T.; Beck, L.; Martinis, S.; Johannes, P.; Lengright, J. Natural Pans as an Important Surface Water Resource in the Cuvelai Basin—Metrics for Storage Volume Calculations and Identification of Potential Augmentation Sites. *Water* **2021**, *13*, 177. <https://doi.org/10.3390/w13020177>. 705  
706  
707
29. Gleyzes, M.A.; Perret, L.; Kubik, P. PLEIADES SYSTEM ARCHITECTURE AND MAIN PERFORMANCES. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* **2012**, XXXIX-B1, 537–542. <https://doi.org/10.5194/isprsarchives-XXXIX-B1-537-2012>. 708  
709  
710
30. Domeneghetti, A.; Schumann, G.J.P.; Frasson, R.P.M.; Wei, R.; Pavelsky, T.M.; Castellarin, A.; Brath, A.; Durand, M.T. Characterizing water surface elevation under different flow conditions for the upcoming SWOT mission. *Journal of Hydrology* **2018**, *561*, 848–861. <https://doi.org/10.1016/j.jhydrol.2018.04.046>. 711  
712  
713
31. Lebègue, L.; Cazala-Hourcade, E.; Languille, F.; Artigues, S.; Melet, O. CO3D, A WORLDWIDE ONE ONE-METER ACCURACY DEM FOR 2025. In Proceedings of the ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. Copernicus GmbH, 2020, Vol. XLIII-B1-2020, pp. 299–304. ISSN : 1682-1750, <https://doi.org/https://doi.org/10.5194/isprs-archives-XLIII-B1-2020-299-2020>. 714  
715  
716  
717
32. Ferrant, S.; Selles, A.; Le Page, M.; Herrault, P.A.; Pelletier, C.; Al-Bitar, A.; Mermoz, S.; Gascoin, S.; Bouvet, A.; Saqalli, M.; et al. Detection of Irrigated Crops from Sentinel-1 and Sentinel-2 Data to Estimate Seasonal Groundwater Use in South India. *Remote Sensing* **2017**, *9*, 1119. <https://doi.org/10.3390/rs9111119>. 718  
719  
720
33. Ferrant, S.; Selles, A.; Le Page, M.; AlBitar, A.; Mermoz, S.; Gascoin, S.; Bouvet, A.; Ahmed, S.; Kerr, Y. SENTINEL-1&2 FOR NEAR REAL TIME CROPPING PATTERN MONITORING IN DROUGHT PRONE AREAS. APPLICATION TO IRRIGATION WATER NEEDS IN TELANGANA, SOUTH-INDIA. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* **2019**, XLII-3/W6, 285–292. <https://doi.org/10.5194/isprs-archives-XLII-3-W6-285-2019>. 721  
722  
723  
724
34. Maréchal, J.C.; Dewandel, B.; Ahmed, S.; Galeazzi, L.; Zaidi, F.K. Combined estimation of specific yield and natural recharge in a semi-arid groundwater basin with irrigated agriculture. *Journal of Hydrology* **2006**, *329*, 281–293. <https://doi.org/10.1016/j.jhydrol.2006.02.022>. 725  
726  
727
35. Lehner, B.; Liermann, C.R.; Revenga, C.; Vörösmarty, C.; Fekete, B.; Crouzet, P.; Döll, P.; Endejan, M.; Frenken, K.; Magome, J.; et al. High-resolution mapping of the world’s reservoirs and dams for sustainable river-flow management. *Frontiers in Ecology and the Environment* **2011**, *9*, 494–502. <https://doi.org/https://doi.org/10.1890/100125>. 728  
729  
730
36. Pascal, C.; Ferrant, S.; Selles, A.; Maréchal, J.C.; Gascoin, S.; Merlin, O. High-Resolution Mapping of Rainwater Harvesting System Capacity from Satellite Derived Products in South India. In Proceedings of the 2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS, 2021, pp. 7011–7014. ISSN : 2153-7003, <https://doi.org/10.1109/IGARSS47720.2021.9553131>. 731  
732  
733
37. Beyer, R.; Alexandrov, O.; ScottMcMichael. NeoGeographyToolkit/StereoPipeline : ASP 2.6.2, 2019. <https://doi.org/10.5281/ZENODO.3247734>. 734  
735
38. Marti, R.; Gascoin, S.; Berthier, E.; de Pinel, M.; Houet, T.; Laffly, D. Mapping snow depth in open alpine terrain from stereo satellite imagery. *The Cryosphere* **2016**, *10*, 1361–1380. Publisher : Copernicus GmbH, <https://doi.org/https://doi.org/10.5194/tc-10-1361-2016>. 736  
737  
738
39. Shean, D.E.; Alexandrov, O.; Moratto, Z.M.; Smith, B.E.; Joughin, I.R.; Porter, C.; Morin, P. An automated, open-source pipeline for mass production of digital elevation models (DEMs) from very-high-resolution commercial stereo satellite imagery. *ISPRS Journal of Photogrammetry and Remote Sensing* **2016**, *116*, 101–117. <https://doi.org/10.1016/j.isprsjprs.2016.03.012>. 739  
740  
741
40. Muralikrishnan, S.; Narender, B.; Reddy, S.; Pillai, A. Evaluation of Indian National DEM from Cartosat-1 Data, 2011. Indian Space Research Organisation National Remote Sensing Center. 742  
743
41. Wessel, B. TanDEM-X DEM Product Specification, 2016. Public Document TD-GS-PS-0021. Earth Observation Center : Weßling, Germany. 744  
745
42. Takaku, J.; Tadono, T.; Tsutsui, K. Generation of High Resolution Global DSM from ALOS PRISM. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* **2014**, XL-4, 243–248. <https://doi.org/10.5194/isprsarchives-XL-4-243-2014>. 746  
747  
748

43. Hagolle, O.; Huc, M.; Desjardins, C.; Auer, S.; Richter, R. MAJA Algorithm Theoretical Basis Document 2017. Publisher : Zenodo, <https://doi.org/10.5281/zenodo.1209633>. 749  
750
44. Breiman, L. Random Forests. *Machine Learning* 2001, 45, 5–32. <https://doi.org/10.1023/A:1010933404324>. 751
45. Tulbure, M.G.; Broich, M.; Stehman, S.V.; Kommareddy, A. Surface water extent dynamics from three decades of seasonally continuous Landsat time series at subcontinental scale in a semi-arid region. *Remote Sensing of Environment* 2016, 178, 142–157. <https://doi.org/10.1016/j.rse.2016.02.034>. 752  
753  
754
46. Bangira, T.; Alfieri, S.M.; Menenti, M.; van Niekerk, A. Comparing Thresholding with Machine Learning Classifiers for Mapping Complex Water. *Remote Sensing* 2019, 11, 1351. Number : 11 Publisher : Multidisciplinary Digital Publishing Institute, <https://doi.org/10.3390/rs11111351>. 755  
756  
757
47. Nobre, A.; Cuartas, L.; Hodnett, M.; Rennó, C.; Rodrigues, G.; Silveira, A.; Waterloo, M.; Saleska, S. Height Above the Nearest Drainage – a hydrologically relevant new terrain model. *Journal of Hydrology* 2011, 404, 13–29. <https://doi.org/10.1016/j.jhydrol.2011.03.051>. 758  
759  
760
48. McFeeters, S.K. The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features. *International Journal of Remote Sensing* 1996, 17, 1425–1432. Publisher : Taylor & Francis, <https://doi.org/10.1080/01431169608948714>. 761  
762
49. Xu, H. Modification of normalised difference water index (NDWI) to enhance open water features in remotely sensed imagery. *International Journal of Remote Sensing* 2006, 27, 3025–3033. Publisher : Taylor & Francis \_eprint : <https://doi.org/10.1080/01431160600589179>, <https://doi.org/10.1080/01431160600589179>. 763  
764  
765
50. Nielsen, N.H.; Larsen, M.R.A.; Rasmussen, S.F. Development of a screening method to assess flood risk on Danish national roads and highway systems. *Water Science and Technology* 2011, 63, 2957–2966. <https://doi.org/10.2166/wst.2011.157>. 766  
767
51. Rodrigues, L.N.; Sano, E.E.; Steenhuis, T.S.; Passo, D.P. Estimation of Small Reservoir Storage Capacities with Remote Sensing in the Brazilian Savannah Region. *Water Resources Management* 2012, 26, 873–882. <https://doi.org/10.1007/s11269-011-9941-8>. 768  
769
52. Gruber, A.; Wessel, B.; Martone, M.; Roth, A. The TanDEM-X DEM Mosaicking : Fusion of Multiple Acquisitions Using InSAR Quality Parameters. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 2016, 9, 1047–1057. Conference Name : IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, <https://doi.org/10.1109/JSTARS.2015.2421879>. 770  
771  
772  
773
53. Lemoine, F.G.K. The Development of the Joint NASA GSFC and the National Imagery and Mapping Agency (NIMA) Geopotential Model EGM96. Technical report, 1998. 774  
775
54. Grohmann, C.H. Evaluation of TanDEM-X DEMs on selected Brazilian sites : Comparison with SRTM, ASTER GDEM and ALOS AW3D30. *Remote Sensing of Environment* 2018, 212, 121–133. <https://doi.org/10.1016/j.rse.2018.04.043>. 776  
777
55. Deschamps-Berger, C.; Gascoin, S.; Berthier, E.; Deems, J.; Gutmann, E.; Dehecq, A.; Shean, D.; Dumont, M. Snow depth mapping from stereo satellite imagery in mountainous terrain : evaluation using airborne laser-scanning data. *The Cryosphere* 2020, 14, 2925–2940. Publisher : Copernicus GmbH, <https://doi.org/10.5194/tc-14-2925-2020>. 778  
779  
780
56. Doña, C.; Chang, N.B.; Caselles, V.; Sánchez, J.M.; Pérez-Planells, L.; Bisquert, M.D.M.; García-Santos, V.; Imen, S.; Camacho, A. Monitoring Hydrological Patterns of Temporary Lakes Using Remote Sensing and Machine Learning Models : Case Study of La Mancha Húmeda Biosphere Reserve in Central Spain. *Remote Sensing* 2016, 8, 618. Number : 8 Publisher : Multidisciplinary Digital Publishing Institute, <https://doi.org/10.3390/rs8080618>. 781  
782  
783  
784
57. Perin, V.; Tulbure, M.G.; Gaines, M.D.; Reba, M.L.; Yaeger, M.A. A multi-sensor satellite imagery approach to monitor on-farm reservoirs. *Remote Sensing of Environment* 2021, p. 112796. <https://doi.org/10.1016/j.rse.2021.112796>. 785  
786
58. Ferrier, P.; Crebassol, P.; Dedieu, G.; Hagolle, O.; Meygret, A.; Tinto, F.; Yaniv, Y.; Herscovitz, J. VEN $\mu$ S (Vegetation and environment monitoring on a new micro satellite). In Proceedings of the 2010 IEEE International Geoscience and Remote Sensing Symposium, 2010, pp. 3736–3739. ISSN : 2153-7003, <https://doi.org/10.1109/IGARSS.2010.5652087>. 787  
788  
789
59. Farr, T.G.; Rosen, P.A.; Caro, E.; Crippen, R.; Duren, R.; Hensley, S.; Kobrick, M.; Paller, M.; Rodriguez, E.; Roth, L.; et al. The Shuttle Radar Topography Mission. *Reviews of Geophysics* 2007, 45. <https://doi.org/10.1029/2005RG000183>. 790  
791
60. Boisson, A.; Villesseche, D.; Selles, A.; Alazard, M.; Chandra, S.; Ferrant, S.; Maréchal, J.C. Long Term Monitoring of Rainwater Harvesting Tanks : Is Multi-Years Management Possible in Crystalline South Indian Aquifers? *Hydrological Processes* 2022, n/a, e14759. <https://doi.org/10.1002/hyp.14759>. 792  
793  
794

## 5.4 Discussion et perspectives

Cette section discute des incertitudes liées à l'estimation du volume à partir des MNT Pléiades et des cartes d'eau Sentinel-2, et propose des pistes pour l'extension de ce type de méthodes à large échelle et de nouvelles applications.

### 5.4.1 Précision du MNT

L'évaluation de la précision des MNT Pléiades nécessite l'acquisition de mesures terrain d'altitude avec une date d'acquisition proche. Seules les données de Perrin et al. (2012) sont disponibles dans le cadre de ces travaux pour la validation des MNT Pléiades. Elles consistent en 1146 mesures acquises en 2010 par GPS différentiel (DGPS) sur les bathymétries de quatre tanks localisés sur l'emprise Pléiades (d) (voir description des données en section 2.2.2). L'erreur de l'altitude Pléiades par rapport aux mesures DGPS est caractérisée par trois métriques :

- L'écart-type de la distribution des différences entre mesure DGPS et altitude Pléiades,
- Le RMSE entre mesures DGPS et altitudes Pléiades (erreur fournie avec les MNT CartoDEM (Muralikrishnan et al., 2011) et ALOS (Takaku et al., 2014) à 30 m),
- L'erreur verticale absolue (incertitude de l'altitude Pléiades par rapport à la mesure DGPS) exprimée par un intervalle de confiance à 90% (erreur fournie avec le MNT TanDEM-X à 12 m et CartoDEM).

Sur les 1146 points, les trois erreurs sont respectivement estimées à 2 m, 2.5 m et l'erreur verticale absolue est inférieure à 1.6 m. Ces erreurs sont inférieures à celles d'autres MNT (bien que non directement comparables) : RMSE de 5 m pour ALOS à 30 m, RMSE entre 3.4 et 4.72 pour CartoDEM à 30 m, erreur verticale absolue de 2 m pour TanDEM-X (Wessel, 2016) et comprise entre 4.7 et 7.3 m pour CartoDEM. Cette erreur, ponctuelle, diminue avec l'agrégation spatiale, notamment lors du calcul des volumes des tanks (Deschamps-Berger et al., 2020).

Pendant, l'erreur pourrait être encore plus faible si elle était comparée à partir de mesures DGPS synchrones avec les acquisitions Pléiades. En effet, la présence de bassins profonds sur les tanks 5 et 15 sur le MNT Pléiades coïncidant avec des altitudes DGPS plus grandes (points oranges et rouges sur la figure 5.8) suggèrent qu'ils ont été creusés dans les 9 ans séparant les deux acquisitions, vraisemblablement dans le cadre de la mission Kakatiya.

### 5.4.2 Précision des surfaces en eau détectées sur Sentinel-2

Les incertitudes des cartes d'eau Sentinel-2 sont évaluées avec des images optiques très haute résolution sur 16 tanks et différentes dates pour un total de 29 points (section 2.3.4.3). Les surfaces et les volumes dérivés à partir des relations hypsométriques sont généralement sous-estimés par les cartes Sentinel-2, avec une erreur relative moyenne de +17.5 % sur les surfaces et +30.7% sur les volumes, du fait de la résolution des images et des imperfections des algorithmes de classification (voir la distribution des erreurs en figure 5.9). Ainsi, la capacité du RHS, estimée à 23.5 mm, pourrait représenter une fraction plus importante de l'eau de surface. Ce biais pourrait être réduit en détectant l'eau sur des images optiques à une résolution plus haute, comme Venüs (résolutions spatiale et temporelle de 5.3 m et deux jours) (Ferrier et al., 2010) ou Planet (revisite journalière et résolution de 3 m) (Cooley et al., 2017), ce qui pourrait se révéler essentiel dans les futures études du RHS.

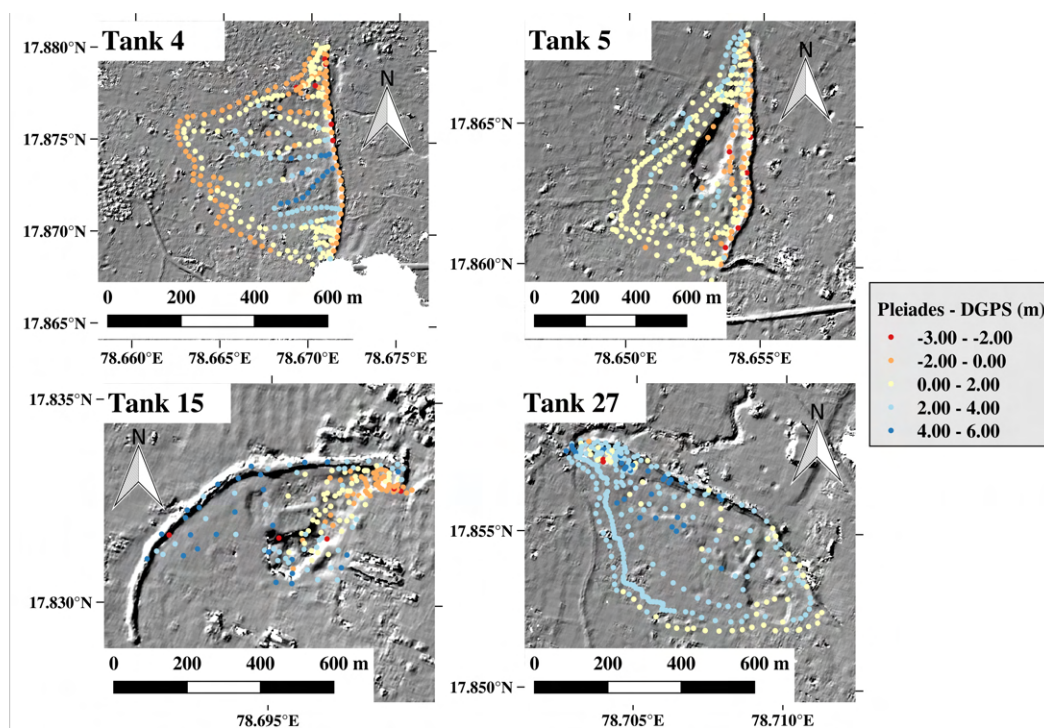


FIGURE 5.8 – Localisation et bathymétries des quatre tanks de Perrin et al. (2012). Les mesures DGPS sont indiquées par des points dont la couleur correspond à la différence entre l'altitude du MNT Pléiades et la mesure DGPS en mètres.

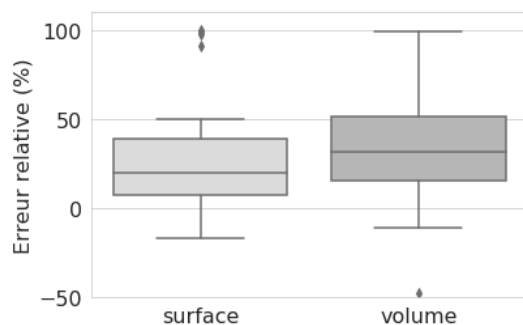


FIGURE 5.9 – Distribution des erreurs des surfaces en eau Sentinel-2 et volumes dérivés par rapport aux images optiques très haute résolution.

### 5.4.3 Délimitation des tanks

Les travaux présentés dans cette thèse se concentrent sur le RHS d'une zone de 1813 km<sup>2</sup> couverte par quatre emprises Pléiades, avec une capacité estimée à 58.3 MCM (32 mm) (table 5.1). L'extension de cette étude à plus large échelle est limitée par le processus manuel de délimitation des tanks. En effet, le calcul de la courbe hypsométrique dépend fortement de la correspondance entre les limites du tank et les isolignes du MNT. L'utilisation de données satellitaires pour délimiter les tanks comme le MWAE dérivé de Sentinel-2 est un processus semi-automatique rapide, mais induit dans le cadre de notre méthode une sur-estimation des volumes les plus élevés en raison des imprécisions des méthodes de détection d'eau et de la présence d'une digue. C'est pourquoi les tanks ont été délimités par une isoligne du MNT Pléiades sélectionnée manuellement, une tâche



longue qui ne peut être reproduite à plus large échelle. Cette tâche est automatisée dans Arendt et al. (2021) et Nielsen et al. (2011) avec l’algorithme de détection Blue Spot qui identifie les dépressions naturelles à partir d’un MNT. En particulier, cette méthode a été utilisée dans le bassin de Cuvelai, une zone semi-aride localisée entre l’Angola et la Namibie inondée une partie du temps, pour identifier un MWAE et une capacité volumétrique maximale théorique des dépressions naturelles à partir du MNT TanDEM-X à 12 m de résolution. Une combinaison de l’algorithme de détection Blue Spot et de cartes en eau dérivées des données optiques Sentinel-2 pourrait permettre de délimiter les tanks actifs pour étendre ce type d’étude à plus large échelle.

#### 5.4.4 Futures applications

Les MNT haute résolution sont essentiels pour estimer et suivre le stock d’eau de surface avec précision en quasi temps réel en combinaison avec des observations optiques haute résolution. La saison d’acquisition (qui doit se faire lorsque les tanks sont vides) et la revisite (qui permet de mettre à jour les bathymétries changeantes) sont essentielles pour le suivi du RHS. À ce jour, ces MNT comme Pléiades ne sont disponibles que sur de faibles emprises, mais les MNT précis à large échelle fournis par les futures missions de stéréoscopie comme CO3D (Lebègue et al., 2020) seront déterminants pour le développement de systèmes d’alerte de sécheresse en quasi temps réel.

Ce type d’étude ouvre de nouvelles possibilités pour un suivi saisonnier du volume du RHS qui peut être interprété en terme de flux entrants (pluie, ruissellement) et sortants (évaporation et infiltration) entre deux dates d’acquisition à une échelle régionale, ce qui n’a été fait jusqu’à présent que pour des tanks individuels (Boisson et al., 2014; Massuel et al., 2014; Perrin et al., 2012). Ces flux sont déterminants pour comprendre l’impact du RHS sur le cycle de l’eau (dont l’eau souterraine) dans ces zones semi-arides.

---

## 5.5 Ce qu'il faut retenir

- La capacité du RHS du Telangana a été estimée à  $24 \pm 9$  mm (contre 113 mm pour la capacité cumulée des principaux lacs de barrage) à partir de modèles numériques de terrain (MNT) Pléiades et de cartes de surface en eau dérivées de Sentinel-2 et de Landsat.
- Les MNT très haute résolution à fort taux de revisite sont essentiels pour quantifier le volume des tanks du RHS. Quantifier un volume avec un MNT à une résolution de 30 m au lieu de 2 m induit au minimum une erreur de 47% sur les tanks individuels et de 20% pour des volumes cumulés sur une zone de  $20 \times 20$  km<sup>2</sup>. Cette erreur plus faible s'explique en partie par le fait que le volume d'eau contenu dans les grands tanks, mieux estimé que dans les petits tanks, constitue une majeure partie des volumes cumulés.
- La validation du MNT Pléiades acquis en 2019 par des données DGPS acquises 10 ans avant indiquent une erreur ponctuelle d'environ 2 m (écart-type de la distribution des différences entre mesures DGPS et altitude Pléiades), qui diminue avec l'agrégation spatiale. Les volumes estimés sur certains tanks avec les mesures terrain et le MNT Pléiades montrent parfois de grandes différences, témoignant des changements de bathymétrie des tanks suite à leur réhabilitation récente dans le cadre du programme Kakatiya.
- Le contrôle de la date d'acquisition est essentiel pour avoir accès à la bathymétrie des tanks vides.
- L'arrivée des prochaines missions topographiques à très haute résolution comme CO3D combinées à des séries temporelles d'images optiques ouvre le champ pour l'étude des systèmes de récolte des eaux de pluie (RHS - *Rainwater Harvesting System*), de leur suivi pour le développement de systèmes d'alerte de sécheresse et de leur impact hydrologique jusqu'à présent mal compris.



# Chapitre 6

## Conclusions et perspectives

L'objectif général de la thèse est de cartographier les ressources en eau (à la surface, dans la zone vadose et dans la zone saturée) à des résolutions multiples et de comprendre leur dynamique dans les régions semi-arides fortement irriguées avec une approche par télédétection multi-capteurs. Ces travaux se sont concentrés sur l'État indien du Telangana (114 800 km<sup>2</sup>), caractérisé par des aquifères hétérogènes à faible capacité de stockage exploités pour l'irrigation de manière déstructurée. Les recherches se sont orientées selon trois axes : (i) l'étude qualitative des séries temporelles satellitaires et *in situ* avec une attention particulière portée au potentiel des données gravimétrique GRACE et d'humidité du sol en surface SMOS et SMAP pour le suivi des ressources en eau souterraine et de l'irrigation, (ii) la cartographie des ressources en eau souterraine par désagrégation des données GRACE, et (iii) la cartographie des volumes d'eau stockée en surface par le système de récupération des eaux de pluie (RHS - *Rainwater Harvesting System*) à partir de données satellitaires haute résolution.

### 6.1 Potentiel des données GRACE et SMOS/SMAP pour le suivi multi-échelle des ressources en eau

Les données GRACE fournissent depuis 2002 des anomalies de stock d'eau total (TWS - *Total Water Storage*) à des résolutions spatiales d'au moins 300 km. La succession d'une saison humide où les précipitations reconstituent les ressources en eau et d'une saison sèche où elles sont vidées par les prélèvements de l'aquifère d'origine anthropique induisent des fluctuations annuelles de TWS très forte à l'échelle du Telangana pouvant aller jusqu'à 60 cm. Contrairement aux aquifères sédimentaires profonds du nord de l'Inde constituant des réservoirs massifs et subissant une baisse de 2 à 3 cm/an (Asoka et al., 2017; Rodell et al., 2018; Tiwari et al., 2009), les aquifères cristallins du Telangana à faible capacité de stockage se remplissent avec les pluies de la mousson et se vident presque entièrement chaque année avec l'intensité et la fréquence des pompages. Ils sont cependant soumis à l'alternance de périodes climatiques plus ou moins humides de quelques années en raison des phénomènes El Niño et La Niña, impactant les études de tendances qui mènent à des conclusions différentes selon la période étudiée. Ainsi, les tendances positives identifiées sur le sud de l'Inde dans la littérature (+1-2 cm/an calculée entre 2002 et 2013 (Asoka et al., 2017) et +0.7 cm/an entre 2002 et 2018 (Rodell et al., 2018)) sont la conséquence de la grande sécheresse ayant précédé la mission GRACE, tandis qu'aucune pente significative n'émerge du signal entre 2002 et 2021.

Le TWS de GRACE est une donnée unique en hydrologie, puisqu'il permet d'estimer les va-

riations de stock d'eau souterraine (GWS) par satellite en soustrayant du TWS les réservoirs hydrologiques de surface et des couches peu profondes du sol (Breña-Naranjo et al., 2014; Cao and Roy, 2020; Frappart et al., 2019; Frappart and Ramillien, 2018; Papa et al., 2015; Rodell et al., 2018; Rzepecka and Birylo, 2020; Tiwari et al., 2009; Zhang et al., 2020). Les variations de GWS dérivées des données GRACE à l'échelle du Telangana présentent une bonne correspondance avec celles dérivées des données *in situ* ( $R^2 = 0.70$ ), démontrant le potentiel de ces méthodes pour le suivi du GWS à la résolution de GRACE sur les régions dépourvues de réseau dense de puits d'observation.

Mais dans ce contexte où les rares ressources en eau sont réparties de manière très hétérogène dans les aquifères et en surface, une échelle plus fine est nécessaire pour rendre compte de la complexité des processus hydrologiques. À l'échelle de 25 km (résolution de grille des données SMOS), les données SMOS et le NDVI de MODIS présentent des dynamiques différentes selon l'occupation du sol et le degré d'anthropisation. En particulier, les régions très irriguées sont caractérisées par un cycle annuel bimodal de l'humidité de surface dérivée des données SMOS et SMAP, ainsi que du NDVI de MODIS, dont la temporalité des deux pics correspondent respectivement à la fin de la mousson et au début de la deuxième saison culturale. Le pic en saison sèche détecté à l'échelle du Telangana et à l'échelle de 25 km en l'absence de précipitations sur des zones très irriguées est une preuve visible par satellite de l'intensité de l'irrigation, constat appuyé par (i) sa présence uniquement sur les régions agricoles fortement irriguées, (ii) son absence des séries temporelles d'humidité du sol modélisées sans prise en compte de l'irrigation, (iii) la cohérence temporelle des pics en saison sèche d'humidité du sol SMOS et de NDVI, et (iv) la forte corrélation entre l'amplitude du pic d'humidité SMOS et la superficie de riz cultivé ( $R = 0.81$ ).

Les données d'humidité du sol de SMOS possèdent donc un fort potentiel pour l'estimation de l'eau utilisée pour l'irrigation, particulièrement en saison sèche. Cependant les techniques actuelles pour une estimation quantitative, basée sur l'inversion du bilan hydrique du sol (Brocca et al., 2018) à partir des données satellitaires d'humidité du sol ne parviennent pas à séparer les contributions de la pluie et de l'irrigation. Ceci est particulièrement dommageable dans cette région où les cultures sont irriguées même pendant les épisodes de pluie. Une autre méthode d'estimation en cours de développement, basée sur le modèle PrISM, pourrait donner d'autres résultats (Pellarin et al., 2020). Ce type d'estimation pourrait aussi bénéficier des missions futures proposant des estimations d'humidité du sol de surface à plus haute résolution comme SMOS-HR, actuellement en phase 0 au Centre National d'Études Spatiales (CNES) (Rodríguez-Fernández et al., 2019).

## 6.2 Désagrégation des données GRACE pour le suivi de l'eau souterraine à haute résolution

Le TWS de GRACE possède un grand potentiel pour l'estimation du GWS, cependant sa faible résolution (300-400 km) ne permet pas d'obtenir des cartes à une échelle pertinente pour la gestion des territoires et de l'irrigation. Cette dernière décennie, de nombreuses études ont cherché à obtenir des cartes mensuelles de GWS à une résolution supérieure par désagrégation des données GRACE. Cette désagrégation passe par leur assimilation dans des modèles hydrologiques (Giotto et al., 2016; Houborg et al., 2012; Nie et al., 2019; Schumacher et al., 2018; Tian et al., 2017; Zaitchik et al., 2008) ou avec des approches empiriques basées sur le caractère supposément multi-échelle de la relation entre le TWS et des variables auxiliaires disponibles à plus haute résolution spatiale et supposées explicatives de la variabilité de TWS (Ali et al., 2021; Jyolsna et al., 2021; Karunakalage et al., 2021; Sahour et al., 2020; Seyoum and Milewski, 2017; Vishwakarma et al., 2021; Zhang

et al., 2021a,b). Les cartes de GWS désagrégé sont validées à partir de mesures piézométriques. Cependant, les méthodes de validation présentes dans la littérature reposent essentiellement sur une appréciation arbitraire des métriques ou des pentes calculées entre les séries temporelles des données désagrégées et des données de validation, sans inclure l’aspect spatial et sans référence à laquelle comparer les métriques obtenues.

Dans les travaux présentés dans cette thèse, une nouvelle méthode de validation a été développée permettant (i) de valider de manière appropriée les produits désagrégés à la fois dans le temps et dans l’espace, et (ii) de comparer les performances des cartes de GWS désagrégé et de la donnée non désagrégée, en s’appuyant sur le calcul d’un gain. Cette méthode de validation a été utilisée pour évaluer la qualité des cartes de GWS désagrégé à 50 km à partir du TWS à l’échelle du Telangana (100 000 km<sup>2</sup>) avec deux méthodes de désagrégation empirique existantes “statiques” (modélisation du TWS avec une fonction dont les paramètres sont constants dans le temps) et une autre méthode empirique “dynamique” développée dans le cadre de ces travaux (modélisation de la dynamique temporelle du TWS à haute résolution).

Les méthodes statiques calibrées à partir des données de NDVI, d’humidité du sol et des précipitations fournissent des cartes généralement plus informatives que la basse résolution (TWS sans désagrégation) sur les aspects temporels et spatiaux. Le double aspect de la validation souligne les hétérogénéités des performances de l’algorithme à la fois dans l’espace et, pour la première fois pour les données désagrégées de GRACE, dans le temps. Ainsi, la désagrégation présente des performances plus faible sur les zones urbaines, et éprouve des difficultés pour représenter la distribution spatiale du GWS dans les périodes où il est particulièrement faible (à la fin de la saison sèche et au début de la mousson). En revanche, les méthodes dynamiques développées pour la première fois dans ces travaux ne permettent pas de produire des cartes de GWS meilleures que la basse résolution. Le développement de ces méthodes dynamiques est encouragé, du fait de leur potentiel pour représenter les tendances interannuelles à haute résolution.

Les techniques de désagrégation souffrent encore de plusieurs obstacles. D’abord, les performances de la méthode sont potentiellement limitées par le caractère exogène des données satellitaires utilisées pour sa calibration et *in situ* pour sa validation. A basse résolution, le  $R^2$  entre le GWS *in situ* et le GWS obtenu par déconvolution du signal GRACE n’excède pas 0.70, et de meilleures performances peuvent difficilement être attendues pour les cartes de GWS désagrégé. Aussi, de manière conceptuelle, une méthode dynamique basée sur le caractère multi-échelle d’une relation empirique liant le TWS et d’autres variables auxiliaires calibrées à basse résolution serait insuffisante pour désagréger le TWS dans un milieu où les variables utilisées pour la désagrégation présentent des dynamiques temporelles variées aux interprétations physiques différentes entre la basse et la haute résolution. C’est typiquement le cas d’une région irriguée comme le Telangana où la variété des cycles annuels des séries temporelles d’humidité du sol de SMOS et de NDVI de MODIS sur différents pixels de 25 km (cycle annuel bimodal ou unimodal selon si la région est fortement irriguée ou non) n’est pas représentée par les variables agrégées à l’échelle du Telangana.

La mission GRACE-FO lancée en 2018 permet la continuité des acquisitions de GRACE depuis 2002. Sur le long terme, des missions gravimétriques futures (MARVEL (Lemoine et al., 2020; Lemoine and Manda, 2020), MAGIC (Massotti et al., 2021)) prévoient d’améliorer significativement la précision et la qualité des estimations gravimétriques en proposant de nouvelles configurations pour les constellations satellitaires. Cependant, la résolution spatiale, bien que légèrement améliorée, demeurera faible (plusieurs centaines de km). Aussi, les méthodes de désagrégation resteront nécessaires pour le suivi des ressources en eau souterraine à partir de ces futures données gravimétriques. Dans les futurs efforts pour développer des méthodes de désagrégation des données gravimétriques, une méthode de validation exhaustive comme celle développée dans ces travaux

sera très utile pour mieux interpréter les failles des modèles et identifier les pistes d’amélioration. En particulier, cela donne la possibilité de développer et d’évaluer des modèles différents selon les périodes de l’année afin de mieux représenter l’impact des processus hydrologiques dominants (infiltration, évapotranspiration, ruissellement).

### 6.3 Suivi du volume d’eau de surface avec des données très haute résolution

Un troisième aspect de cette thèse concerne le suivi du volume du RHS, réseau de milliers de petites retenues d’eau dans les régions amont pour entre autres favoriser la recharge de l’aquifère. Cependant, leur volume est peu suivi du fait de leur grand nombre et de leur petite taille (typiquement quelques hectares). Seules des études à l’échelle de tanks individuels ont été réalisées (Boisson et al., 2014; Brauns et al., 2022; Massuel et al., 2014; Perrin et al., 2012), et le véritable impact du RHS à l’échelle régionale est méconnu.

Les méthodes actuelles d’estimation de volume par les données satellitaires reposent sur l’association de mesures altimétriques ou de modèles numériques de terrain (MNT), avec des cartes d’eau produites à partir d’images optiques ou radar. Ces estimations se font généralement sur de grands lacs isolés, et sont difficilement applicables sur les tanks du RHS de par la faible résolution des MNT disponibles gratuitement (minimum 30 m) et l’absence de contrôle sur la date d’acquisition qui ne garantit pas l’accès à leur bathymétrie.

En cela, la cartographie de la bathymétrie de ces tanks par MNT très haute résolution à fort taux de revisite est précieuse pour l’étude du RHS. Le potentiel de tels MNT pour le suivi du volume du RHS a été évalué avec quatre MNT Pléiades de  $20 \times 20 \text{ km}^2$  acquis en juin (fin de la saison sèche où les tanks sont vides).

D’abord, la capacité du RHS à l’échelle du Telangana a été évaluée partir d’une relation entre l’étendue d’eau maximale et le volume maximal du RHS calibrée avec les MNT Pléiades et des cartes d’eau dérivées d’images optiques Sentinel-2. L’extension de cette relation à l’échelle du Telangana avec des cartes d’eau dérivées des images Landsat a permis d’estimer la capacité du RHS à  $24 \pm 9 \text{ mm}$ . Il s’agit d’un réservoir non négligeable comparé à la capacité cumulée de 113 mm des principaux lacs de barrage, important à l’échelle locale pour l’irrigation des cultures dans les zones amont. Cependant, la capacité du RHS reste faible par rapport aux 600 mm de fluctuation annuelle de l’ensemble des ressources en eau, suggérant à l’instar de Brauns et al. (2022) que l’impact du RHS sur la recharge pourrait être minime par rapport aux précipitations ou au flux de retour de l’irrigation à l’échelle d’un territoire donné, sans toutefois éliminer la possibilité d’un effet dans un périmètre proche.

Ensuite, l’utilisation combinée des MNT Pléiades et de cartes d’eau dérivées de Sentinel-2 permet le suivi volumétrique du RHS en quasi-temps réel et notamment de leur décharge pendant la saison sèche. L’apport de la très haute résolution a été quantifié par rapport à l’utilisation d’autres MNT globaux ou nationaux à 30 m ou 90 m. L’estimation du volume des tanks individuels par ces MNT est associée à des incertitudes relatives importantes par rapport aux MNT Pléiades, variant de 47% à 78%. Les volumes agrégés à l’échelle d’une emprise Pléiades de  $20 \times 20 \text{ km}^2$  ont de meilleures précisions, avec une incertitude relative moyenne de 20%, du fait de la compensation des erreurs individuelles des tanks.

La mission CO3D (Lebègue et al., 2020) pour la cartographie de la topographie de la Terre à très haute résolution est donc essentielle pour le suivi volumétrique du RHS. En particulier, l’analyse des variations volumétriques en terme de flux entrants (pluie, ruissellement) et sortants (évaporation

et infiltration) à l'échelle régionale, à ce jour réalisée uniquement sur un petit nombre de tanks (Boisson et al., 2014; Massuel et al., 2014; Perrin et al., 2012), est essentielle pour comprendre l'impact du RHS sur le cycle de l'eau dans les régions semi-arides.

## 6.4 Sur la synergie des données satellitaires multicateurs

Chacun des trois aspects de cette thèse combine plusieurs données multi-sources (données satellitaires, produits de classifications, sorties de modèles ou données *in situ*), afin de les comparer et in fine tenter de comprendre les dynamiques hydrologiques, estimer le stock d'eau souterraine ou la quantité d'eau utilisée pour l'irrigation (chapitre 3), estimer la dynamique du GWS avec l'échelle des usages et de la gestion des territoires par la désagrégation des données GRACE (chapitre 4) ou estimer le volume d'eau de surface (chapitre 5).

L'utilisation de données exogènes permet de mieux contraindre le bilan hydrologique à toutes les échelles, et ainsi de comprendre l'interaction entre les processus hydrologiques naturels et anthropiques, au niveau du cycle annuel des précipitations, des évolutions à moyen ou long terme, l'intensité de l'exploitation de l'aquifère, l'irrigation, ou l'occupation du sol. Cependant, les incertitudes natives des données, la diversité des capteurs, de leur configuration et de leur résolution spatiale et temporelle, les incertitudes des algorithmes de classification, ainsi que l'impact du rééchantillonnage à différentes résolutions rend difficile l'établissement des relations quantitatives entre les données. Ces facteurs peuvent en particulier expliquer la correspondance limitée entre les mesures piézométriques et le GWS issu des données satellitaires à la résolution de GRACE ou désagrégé (variété des capteurs et rééchantillonnages), les difficultés pour l'estimation de l'irrigation avec les données d'humidité de SMOS et de pluie (incertitudes des données), ou la correspondance entre les MNT Pléiades et des cartes d'eau issues de Sentinel-2 (différences de résolution et incertitude de la classification). De plus, les travaux présentés dans les chapitres 3 et 4 se sont focalisés sur des données agrégées à l'échelle mensuelle pour s'adapter à la résolution temporelle limitante de GRACE et des données piézométriques, alors qu'une analyse plus fine pourrait être menée si des données avec une résolution temporelle plus élevée étaient disponibles. Des analyses journalières pourraient potentiellement inclure les données GRACE, dont des solutions journalières locales peuvent être dérivées (Eicker et al., 2020; Gouweleeuw et al., 2018; Kurtenbach et al., 2012, 2009). Aussi, les futures missions satellitaires visant l'amélioration des résolutions spatiale et temporelle et de la précision des mesures (MARVEL (Lemoine et al., 2020; Lemoine and Manda, 2020) et MAGIC (Massotti et al., 2021) pour les mesures gravimétriques, SWOT pour l'altimétrie des surfaces en eau (Domeneghetti et al., 2018), CO3D pour la topographie (Lebègue et al., 2020) ou SMOS-HR pour l'humidité du sol (Rodríguez-Fernández et al., 2019)) pourraient lever certains verrous de connaissance hydrologique à large échelle.





# Acronymes

**CCI** : *Climate Change Initiative*.

**CEFIRES** : Centre Franco-Indien de Recherche sur les Eaux Souterraines.

**CWC** : *Central Water Commission*.

**CWS** : *Canopy Water Storage* (stock d'eau dans la canopée).

**CWSA** : *Canopy Water Storage* (anomalie du stock d'eau dans la canopée).

**DEM** : *Digital Elevation Model* (modèle numérique de terrain).

**GLDAS** : *Global Land Data Assimilation System*.

**GLEAM** : *Global Land Evaporation Amsterdam Model*.

**GRACE** : *Gravity Recovery and Climate Experiment*.

**GSW** : *Global Surface Water*.

**GWL** : *GroundWater Level* (niveau d'eau mesuré dans les puits d'observation).

**GWLK** : *GroundWater Level Kharif* (niveau d'eau minimal durant la période humide).

**GWLR** : *GroundWater Level Rabi* (niveau d'eau minimal durant la période humide).

**GWS** : *GroundWater Storage* (stock d'eau souterraine).

**GWSA** : *GroundWater Storage Anomaly* (anomalie du stock d'eau souterraine).

**HR** : Haute Résolution.

**IDW** : *Inverse Distance Weighting* (interpolation par pondération inverse à la distance).

**IFCGR** : *Indo-French Center for Groundwater Research*.

**LR** : *Low Resolution*, basse résolution.

**MAR** : *Managed Aquifer Recharge*.

**MCM** : *Million Cubic Meter*, (million de mètres cubes).

**ML** : *MultiLinear model* (modèle linéaire multiple).

**MNT** : Modèle Numérique de Terrain.

**MWAE** : *Maximum Water Area Extent*.

**PET** : *Potential EvapoTranspiration* (évapotranspiration potentielle).

**R** : Coefficient de corrélation de Pearson.

**R<sup>2</sup>** : Coefficient de détermination.

**RF** : *Random Forest* (forêt aléatoire).

**RHS** : *Rainwater Harvesting System*.

**RMSE** : *Root Mean Squared Error* (erreur quadratique moyenne).

**RZSM** : *Root Zone Soil Moisture* (humidité en zone racinaire).

**SH** : *Spherical Harmonics* (harmoniques sphériques).

**SM** : *Soil Moisture* (humidité du sol).

**SMA** : *Soil Moisture Anomaly* (anomalie d'humidité du sol).

**SMOS** : *Soil Moisture and Ocean Salinity*.

**SMAP** : *Soil Moisture Active Passive*.

**SSM** : *Surface Soil Moisture* (humidité de surface du sol).

**SR** : *Small Reservoir*.

**SW** : *Surface Water Storage* (stock d'eau de surface).

**SWA** : *Surface Water Storage Anomaly* (anomalie du stock d'eau de surface).

**Sy** : *Specific Yield* (rendement spécifique ou porosité efficace).

**TIN** : *Triangulated Irregular Network* (interpolation par triangulation irrégulière).

**TWS** : *Total Water Storage* (stock d'eau total).

**WRIS** : *India Water Resources Information System*.

# Bibliographie

- Prateek Aggarwal, Anjali Viswamohanan, Danwant Narayanaswamy, and Shruti Sharma. Unpacking India's Electricity Subsidies : Reporting, transparency, and efficacy. December 2020.
- A. Al Bitar, Yann Kerr, O. Merlin, F. Cabot, and Jean-Pierre Wigneron. Global drought index from smos soil moisture. In 2013 IEEE International Geoscience and Remote Sensing Symposium IGARSS, 2013.
- Ahmad Al Bitar, Arnaud Mialon, Yann H. Kerr, François Cabot, Philippe Richaume, Elsa Jacqueline, Arnaud Quesney, Ali Mahmoodi, Stéphane Tarot, Marie Parrens, Amen Al-Yaari, Thierry Pellarin, Nemesio Rodriguez-Fernandez, and Jean-Pierre Wigneron. The global SMOS Level 3 daily soil moisture and brightness temperature maps. Earth System Science Data, 9(1) : 293–315, June 2017. ISSN 1866-3508. doi : 10.5194/essd-9-293-2017.
- Akhtar Ali. Asia's Large-Scale Irrigation Systems : Challenges and Options. April 2012.
- Shoaib Ali, Dong Liu, Qiang Fu, Muhammad Jehanzeb Masud Cheema, Quoc Bao Pham, Md Mazur Rahman, Thanh Duc Dang, and Duong Tran Anh. Improving the Resolution of GRACE Data for Spatio-Temporal Groundwater Storage Assessment. Remote Sensing, 13(17) :3513, January 2021. doi : 10.3390/rs13173513.
- Robert Arendt, Christian Reinhardt-Imjela, Achim Schulte, Leona Faulstich, Tobias Ullmann, Lorenz Beck, Sandro Martinis, Petrina Johannes, and Joachim Lengricht. Natural Pans as an Important Surface Water Resource in the Cuvelai Basin—Metrics for Storage Volume Calculations and Identification of Potential Augmentation Sites. Water, 13(2) :177, January 2021. ISSN 2073-4441. doi : 10.3390/w13020177.
- Akarsh Asoka, Tom Gleeson, Yoshihide Wada, and Vimal Mishra. Relative contribution of monsoon precipitation and pumping to changes in groundwater storage in India. Nature Geoscience, 10(2) :109–117, February 2017. ISSN 1752-0908. doi : 10.1038/ngeo2869.
- Nicolas Avisse, Amaury Tilmant, Marc François Müller, and Hua Zhang. Monitoring small reservoirs' storage with satellite remote sensing in inaccessible areas. Hydrology and Earth System Sciences, 21(12) :6445–6459, December 2017. ISSN 1607-7938. doi : 10.5194/hess-21-6445-2017.
- Reena Badiani and Katrina Jessoe. The Impact of Electricity Subsidies on Groundwater Extraction and Agricultural Production. page 41, 2013.
- Tsitsi Bangira, Silvia Maria Alfieri, Massimo Menenti, and Adriaan van Niekerk. Comparing Thresholding with Machine Learning Classifiers for Mapping Complex Water. Remote Sensing, 11(11) :1351, January 2019. ISSN 2072-4292. doi : 10.3390/rs11111351.

- F. Baup, F. Frappart, and J. Maubant. Combining high-resolution satellite images and altimetry to estimate the volume of small lakes. Hydrology and Earth System Sciences, 18(5) :2007–2020, May 2014. ISSN 1607-7938. doi : 10.5194/hess-18-2007-2014.
- J Benveniste, M Roca, G Levrini, P Vincent, S Baker, O Zanife, C Zelli, and O Bombaci. The Radar Altimetry Mission : RA-2, MWR, DORIS and LRR. page 11, 2001.
- Srinivas Bettadpur. Level-2 Gravity Field Product User Handbook. 4 :21, 2018.
- Ross Beyer, Oleg Alexandrov, and ScottMcMichael. NeoGeographyToolkit/StereoPipeline : ASP 2.6.2. Zenodo, June 2019.
- Tejas Bhagwat, Igor Klein, Juliane Huth, and Patrick Leinenkugel. Volumetric Analysis of Reservoirs in Drought-Prone Areas Using Remote Sensing Products. Remote Sensing, 11(17) :1974, August 2019. ISSN 2072-4292. doi : 10.3390/rs11171974.
- Soumendra N. Bhanja and Abhijit Mukherjee. In situ and satellite-based estimates of usable groundwater storage across India : Implications for drinking water supply and food security. Advances in Water Resources, 126 :15–23, April 2019. ISSN 0309-1708. doi : 10.1016/j.advwatres.2019.02.001.
- Soumendra N. Bhanja, Abhijit Mukherjee, Dipankar Saha, Isabella Velicogna, and James S. Famiglietti. Validation of GRACE based groundwater storage anomaly using in-situ groundwater level measurements in India. Journal of Hydrology, 543 :729–738, December 2016. ISSN 0022-1694. doi : 10.1016/j.jhydrol.2016.10.042.
- Soumendra N. Bhanja, Abhijit Mukherjee, Matthew Rodell, Yoshihide Wada, Siddhartha Chattopadhyay, Isabella Velicogna, Kishore Pangaluru, and James S. Famiglietti. Groundwater rejuvenation in parts of India influenced by water-policy change implementation. Scientific Reports, 7(1) :1–7, August 2017. ISSN 2045-2322. doi : 10.1038/s41598-017-07058-2.
- Wanjuan Bie, Teng Fei, Xinyu Liu, Huizeng Liu, and Guofeng Wu. Small water bodies mapped from Sentinel-2 MSI (MultiSpectral Imager) imagery with higher accuracy. International Journal of Remote Sensing, 41(20) :7912–7930, October 2020. ISSN 0143-1161. doi : 10.1080/01431161.2020.1766150.
- Filsa Bioresita, Anne Puissant, André Stumpf, and Jean-Philippe Malet. A Method for Automatic and Rapid Mapping of Water Surfaces from Sentinel-1 Imagery. Remote Sensing, 10(2) :217, February 2018. ISSN 2072-4292. doi : 10.3390/rs10020217.
- A. Boisson, M. Baisset, M. Alazard, J. Perrin, D. Villesseche, B. Dewandel, W. Kloppmann, S. Chandra, G. Picot-Colbeaux, S. Sarah, S. Ahmed, and J. C. Maréchal. Comparison of surface and groundwater balance approaches in the evaluation of managed aquifer recharge structures : Case of a percolation tank in a crystalline aquifer in India. Journal of Hydrology, 519 :1620–1633, November 2014. ISSN 0022-1694. doi : 10.1016/j.jhydrol.2014.09.022.
- A. Boisson, D. Villesseche, M. Baisset, J. Perrin, M. Viossanges, W. Kloppmann, S. Chandra, B. Dewandel, G. Picot-Colbeaux, R. Rangarajan, J. C. Maréchal, and S. Ahmed. Questioning the impact and sustainability of percolation tanks as aquifer recharge structures in semi-arid crystalline context. Environmental Earth Sciences, 73(12) :7711–7721, June 2015. ISSN 1866-6299. doi : 10.1007/s12665-014-3229-2.

- Matthew Bonnema and Faisal Hossain. Inferring reservoir operating patterns across the Mekong Basin using only space observations. Water Resources Research, 53(5) :3791–3810, 2017. ISSN 1944-7973. doi : 10.1002/2016WR019978.
- Bentje Brauns, Somsubhra Chattopadhyay, Dan J. Lapworth, Sian E. Loveless, Alan M. MacDonald, Andrew A. McKenzie, Muddu Sekhar, Siva Naga Venkat Nara, and Veena Srinivasan. Assessing the role of groundwater recharge from tanks in crystalline bedrock aquifers in Karnataka, India, using hydrochemical tracers. Journal of Hydrology X, 15 :100121, May 2022. ISSN 2589-9155. doi : 10.1016/j.hydroa.2022.100121.
- Leo Breiman. Random Forests. Machine Learning, 45(1) :5–32, October 2001. ISSN 1573-0565. doi : 10.1023/A:1010933404324.
- Jose Agustin Breña-Naranjo, Anthony D. Kendall, and David W. Hyndman. Improved methods for satellite-based groundwater storage estimates : A decade of monitoring the high plains aquifer from space and ground observations. Geophysical Research Letters, 41(17) :6167–6173, 2014. ISSN 1944-8007. doi : 10.1002/2014GL061213.
- L. Brocca, F. Melone, T. Moramarco, W. Wagner, V. Naeimi, Z. Bartalis, and S. Hasenauer. Improving runoff prediction through the assimilation of the ASCAT soil moisture product. Hydrology and Earth System Sciences, 14(10) :1881–1893, October 2010. ISSN 1027-5606. doi : 10.5194/hess-14-1881-2010.
- Luca Brocca, Luca Ciabatta, Christian Massari, Tommaso Moramarco, Sebastian Hahn, Stefan Hasenauer, Richard Kidd, Wouter Dorigo, Wolfgang Wagner, and Vincenzo Levizzani. Soil as a natural rain gauge : Estimating global rainfall from satellite soil moisture data. Journal of Geophysical Research : Atmospheres, 119(9) :5128–5141, 2014. ISSN 2169-8996. doi : 10.1002/2014JD021489.
- Luca Brocca, Thierry Pellarin, Wade T. Crow, Luca Ciabatta, Christian Massari, Dongryeol Ryu, Chun-Hsu Su, Christoph Rüdiger, and Yann Kerr. Rainfall estimation by inverting SMOS soil moisture estimates : A comparison of different methods over Australia. Journal of Geophysical Research : Atmospheres, 121(20) :12,062–12,079, 2016. ISSN 2169-8996. doi : 10.1002/2016JD025382.
- Luca Brocca, Angelica Tarpanelli, Paolo Filippucci, Wouter Dorigo, Felix Zaussinger, Alexander Gruber, and Diego Fernández-Prieto. How much water is used for irrigation? A new approach exploiting coarse resolution satellite soil moisture products. International Journal of Applied Earth Observation and Geoinformation, 73 :752–766, December 2018. ISSN 0303-2434. doi : 10.1016/j.jag.2018.08.023.
- Mary J. Brodzik, Brendan Billingsley, Terry Haran, Bruce Raup, and Matthew H. Savoie. EASE-Grid 2.0 : Incremental but Significant Improvements for Earth-Gridded Data Sets. ISPRS International Journal of Geo-Information, 1(1) :32–45, June 2012. ISSN 2220-9964. doi : 10.3390/ijgi1010032.
- Isabel Caballero, Javier Ruiz, and Gabriel Navarro. Sentinel-2 Satellites Provide Near-Real Time Evaluation of Catastrophic Floods in the West Mediterranean. Water, 11(12) :2499, December 2019. doi : 10.3390/w11122499.

- Yanping Cao and Shouraseni Sen Roy. Spatial patterns of seasonal level trends of groundwater in India during 2002–2016. Weather, 75(4) :123–128, 2020. ISSN 1477-8696. doi : 10.1002/wea.3370.
- Don P. Chambers, John Wahr, Mark E. Tamisiea, and R. Steven Nerem. Ocean mass from GRACE and glacial isostatic adjustment. Journal of Geophysical Research : Solid Earth, 115(B11), 2010. ISSN 2156-2202. doi : 10.1029/2010JB007530.
- J. L. Chen, C. R. Wilson, B. D. Tapley, and S. Grand. GRACE detects coseismic and postseismic deformation from the Sumatra-Andaman earthquake. Geophysical Research Letters, 34(13), 2007. ISSN 1944-8007. doi : 10.1029/2007GL030356.
- Jianli Chen, Jin Li, Zizhan Zhang, and Shengnan Ni. Long-term groundwater variations in Northwest India from satellite gravity measurements. Global and Planetary Change, 116 :130–138, May 2014. ISSN 0921-8181. doi : 10.1016/j.gloplacha.2014.02.007.
- Li Chen, Qisheng He, Kun Liu, Jinyang Li, and Chenlin Jing. Downscaling of GRACE-Derived Groundwater Storage Based on the Random Forest Model. Remote Sensing, 11(24) :2979, December 2019. ISSN 2072-4292. doi : 10.3390/rs11242979.
- S. Chidambaram, G. Senthil Kumar, M. V. Prasanna, A. John Peter, AL. Ramanathan, and K. Srinivasamoorthy. A study on the hydrogeology and hydrogeochemistry of groundwater from different depths in a coastal aquifer : Annamalai Nagar, Tamilnadu, India. Environmental Geology, 57 (1) :59–73, March 2009. ISSN 1866-6299. doi : 10.1007/s00254-008-1282-4.
- V. M. Chowdary, R. Vinu Chandran, N. Neeti, R. V. Bothale, Y. K. Srivastava, P. Ingle, D. Ramakrishnan, D. Dutta, A. Jeyaram, J. R. Sharma, and Ravindra Singh. Assessment of surface and sub-surface waterlogged areas in irrigation command areas of Bihar state using remote sensing and GIS. Agricultural Water Management, 95(7) :754–766, July 2008. ISSN 0378-3774. doi : 10.1016/j.agwat.2008.02.009.
- Robert B. Cleveland, William S. Cleveland, Jean E. McRae, and Irma Terpenning. STL : A Seasonal-Trend Decomposition Procedure Based on Loess. Journal of Official Statistics, 6(1) : 3–73, 1990.
- Sarah W. Cooley, Laurence C. Smith, Leon Stepan, and Joseph Mascaro. Tracking Dynamic Northern Surface Water Changes with High-Frequency Planet CubeSat Imagery. Remote Sensing, 9(12) :1306, December 2017. ISSN 2072-4292. doi : 10.3390/rs9121306.
- Savannah S. Cooley and Felix W. Landerer. Gravity Recovery and Climate Experiment Follow-on (GRACE-FO) Level-3 Data Product User Handbook. [https://podaac-tools.jpl.nasa.gov/drive/files/allData/gracefo/docs/GRACE-FO\\_L3\\_Handbook\\_JPL-D-103133\\_20190327.pdf](https://podaac-tools.jpl.nasa.gov/drive/files/allData/gracefo/docs/GRACE-FO_L3_Handbook_JPL-D-103133_20190327.pdf), 2019.
- Maurício C. R. Cordeiro, Jean-Michel Martinez, and Santiago Peña-Luque. Automatic water detection from multidimensional hierarchical clustering for Sentinel-2 images and a comparison with Level 2A processors. Remote Sensing of Environment, 253 :112209, February 2021. ISSN 0034-4257. doi : 10.1016/j.rse.2020.112209.
- Corrado Corradini. Soil moisture in the development of hydrological processes and its determination at different spatial scales. Journal of Hydrology, 516 :1–5, August 2014. ISSN 0022-1694. doi : 10.1016/j.jhydrol.2014.02.051.

- J. F. Crétaux, A. Arsen, S. Calmant, A. Kouraev, V. Vuglinski, M. Bergé-Nguyen, M. C. Gennero, F. Nino, R. Abarca Del Rio, A. Cazenave, and P. Maisongrande. 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(9) :1497–1507, May 2011. ISSN 0273-1177. doi : 10.1016/j.asr.2011.01.004.
- W. T. Crow, M. J. van den Berg, G. J. Huffman, and T. Pellarin. Correcting rainfall using satellite-based surface soil moisture retrievals : The Soil Moisture Analysis Rainfall Tool (SMART). Water Resources Research, 47(8), 2011. ISSN 1944-7973. doi : 10.1029/2011WR010576.
- Wade T. Crow, George J. Huffman, Rajat Bindlish, and Thomas J. Jackson. Improving Satellite-Based Rainfall Accumulation Estimates Using Spaceborne Surface Soil Moisture Retrievals. Journal of Hydrometeorology, 10(1) :199–212, February 2009. ISSN 1525-7541, 1525-755X. doi : 10.1175/2008JHM986.1.
- Narendra Narayan Das, Dara Entekhabi, Eni G. Njoku, Jiancheng J. C. Shi, Joel T. Johnson, and Andreas Colliander. Tests of the SMAP Combined Radar and Radiometer Algorithm Using Airborne Field Campaign Observations and Simulated Data. IEEE Transactions on Geoscience and Remote Sensing, 52(4) :2018–2028, April 2014. ISSN 1558-0644. doi : 10.1109/TGRS.2013.2257605.
- O. De Viron, I. Panet, V. Mikhailov, M. Van Camp, and M. Diamant. Retrieving earthquake signature in grace gravity solutions. Geophysical Journal International, 174(1) :14–20, July 2008. ISSN 0956-540X. doi : 10.1111/j.1365-246X.2008.03807.x.
- B Delaunay. Sur la sphère vide. A la mémoire de Georges Voronoï. page 9, 1934.
- S. B. Deolankar. The Deccan Basalts of Maharashtra, India —Their Potential as Aquifers. Groundwater, 18(5) :434–437, 1980. ISSN 1745-6584. doi : 10.1111/j.1745-6584.1980.tb03416.x.
- César Deschamps-Berger, Simon Gascoin, Etienne Berthier, Jeffrey Deems, Ethan Gutmann, Amaury Dehecq, David Shean, and Marie Dumont. Snow depth mapping from stereo satellite imagery in mountainous terrain : Evaluation using airborne laser-scanning data. The Cryosphere, 14(9) :2925–2940, September 2020. ISSN 1994-0416. doi : 10.5194/tc-14-2925-2020.
- B. Dewandel, P. Lachassagne, R. Wyns, J. C. Maréchal, and N. S. Krishnamurthy. A generalized 3-D geological and hydrogeological conceptual model of granite aquifers controlled by single or multiphase weathering. Journal of Hydrology, 330(1) :260–284, October 2006. ISSN 0022-1694. doi : 10.1016/j.jhydrol.2006.03.026.
- B. Dewandel, J. Perrin, S. Ahmed, S. Aulong, Z. Hrkal, P. Lachassagne, M. Samad, and S. Massuel. Development of a tool for managing groundwater resources in semi-arid hard rock regions : Application to a rural watershed in South India. Hydrological Processes, 24(19) :2784–2797, September 2010. ISSN 08856087. doi : 10.1002/hyp.7696.
- B. Dewandel, J.C. Maréchal, O. Bour, B. Ladouche, S. Ahmed, S. Chandra, and H. Pauwels. Upscaling and regionalizing hydraulic conductivity and effective porosity at watershed scale in deeply weathered crystalline aquifers. Journal of Hydrology, 416–417 :83–97, January 2012. ISSN 00221694. doi : 10.1016/j.jhydrol.2011.11.038.



- Benoît Dewandel, Patrick Lachassagne, Faisal K. Zaidi, and Subash Chandra. A conceptual hydrodynamic model of a geological discontinuity in hard rock aquifers : Example of a quartz reef in granitic terrain in South India. Journal of Hydrology, 405(3) :474–487, August 2011. ISSN 0022-1694. doi : 10.1016/j.jhydrol.2011.05.050.
- Benoît Dewandel, Yvan Caballero, Jérôme Perrin, Alexandre Boisson, Fabrice Dazin, Sylvain Ferant, Subash Chandra, and Jean-Christophe Maréchal. A methodology for regionalizing 3-D effective porosity at watershed scale in crystalline aquifers. Hydrological Processes, 31(12) : 2277–2295, 2017a. ISSN 1099-1085. doi : 10.1002/hyp.11187.
- Benoît Dewandel, Julie Jeanpert, Bernard Ladouche, Jean-Lambert Join, and Jean-Christophe Maréchal. Inferring the heterogeneity, transmissivity and hydraulic conductivity of crystalline aquifers from a detailed water-table map. Journal of Hydrology, 550 :118–129, July 2017b. ISSN 00221694. doi : 10.1016/j.jhydrol.2017.03.075.
- Kamel Didan, Armando Barreto Munoz, and Alfredo Huete. MODIS Vegetation Index User’s Guide (MOD13 Series). page 35, 2015.
- A. Domeneghetti, G. J. P. Schumann, R. P. M. Frasson, R. Wei, T. M. Pavelsky, A. Castellarin, A. Brath, and M. T. Durand. Characterizing water surface elevation under different flow conditions for the upcoming SWOT mission. Journal of Hydrology, 561 :848–861, June 2018. ISSN 0022-1694. doi : 10.1016/j.jhydrol.2018.04.046.
- Carolina Doña, Ni-Bin Chang, Vicente Caselles, Juan Manuel Sánchez, Lluís Pérez-Planells, Maria Del Mar Bisquert, Vicente García-Santos, Sanaz Imen, and Antonio Camacho. Monitoring Hydrological Patterns of Temporary Lakes Using Remote Sensing and Machine Learning Models : Case Study of La Mancha Húmeda Biosphere Reserve in Central Spain. Remote Sensing, 8(8) : 618, August 2016. doi : 10.3390/rs8080618.
- Zheng Duan and W. G. M. Bastiaanssen. Estimating water volume variations in lakes and reservoirs from four operational satellite altimetry databases and satellite imagery data. Remote Sensing of Environment, 134 :403–416, July 2013. ISSN 0034-4257. doi : 10.1016/j.rse.2013.03.010.
- John L Dwyer. Landsat Collection 1 Level 1 Product Definition. page 32, 2019.
- Annette Eicker, Laura Jensen, Viviana Wöhnke, Henryk Dobsław, Andreas Kvas, Torsten Mayer-Gürr, and Robert Dill. Daily GRACE satellite data evaluate short-term hydro-meteorological fluxes from global atmospheric reanalyses. Scientific Reports, 10(1) :4504, March 2020. ISSN 2045-2322. doi : 10.1038/s41598-020-61166-0.
- Markus Enenkel, Caroline Steiner, Thomas Mistelbauer, Wouter Dorigo, Wolfgang Wagner, Linda See, Clement Atzberger, Stefan Schneider, and Edith Roggenhofer. A Combined Satellite-Derived Drought Indicator to Support Humanitarian Aid Organizations. Remote Sensing, 8(4) :340, April 2016. ISSN 2072-4292. doi : 10.3390/rs8040340.
- Teresa L. Evans, Maycira Costa, Kevin Telmer, and Thiago S. F. Silva. Using ALOS/PALSAR and RADARSAT-2 to Map Land Cover and Seasonal Inundation in the Brazilian Pantanal. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 3(4) :560–575, December 2010. ISSN 2151-1535. doi : 10.1109/JSTARS.2010.2089042.

- Bin Fang, Venkat Lakshmi, Rajat Bindlish, Thomas J. Jackson, Michael Cosh, and Jeffrey Basara. Passive Microwave Soil Moisture Downscaling Using Vegetation Index and Skin Surface Temperature. Vadose Zone Journal, 12(3), August 2013. ISSN 1539-1663. doi : 10.2136/vzj2013.05.0089.
- Bin Fang, Venkataraman Lakshmi, Rajat Bindlish, and Thomas J. Jackson. Downscaling of SMAP Soil Moisture Using Land Surface Temperature and Vegetation Data. Vadose Zone Journal, 17(1) :170198, 2018. ISSN 15391663. doi : 10.2136/vzj2017.11.0198.
- Quanlong Feng, Jianhua Gong, Jiantao Liu, and Yi Li. Flood Mapping Based on Multiple End-member Spectral Mixture Analysis and Random Forest Classifier—The Case of Yuyao, China. Remote Sensing, 7(9) :12539–12562, September 2015. ISSN 2072-4292. doi : 10.3390/rs70912539.
- S. Ferrant, A. Selles, M. Le Page, A. AlBitar, S. Mermoz, S. Gascoin, A. Bouvet, S. Ahmed, and Y. Kerr. SENTINEL-1&2 FOR NEAR REAL TIME CROPPING PATTERN MONITORING IN DROUGHT PRONE AREAS. APPLICATION TO IRRIGATION WATER NEEDS IN TELANGANA, SOUTH-INDIA. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XLII-3/W6 :285–292, July 2019. ISSN 2194-9034. doi : 10.5194/isprs-archives-XLII-3-W6-285-2019.
- Sylvain Ferrant, Yvan Caballero, Jérôme Perrin, Simon Gascoin, Benoit Dewandel, Stéphanie Aulong, Fabrice Dazin, Shakeel Ahmed, and Jean-Christophe Maréchal. Projected impacts of climate change on farmers’ extraction of groundwater from crystalline aquifers in South India. Scientific Reports, 4(1) :1–10, January 2014. ISSN 2045-2322. doi : 10.1038/srep03697.
- Sylvain Ferrant, Adrien Selles, Michel Le Page, Pierre-Alexis Herrault, Charlotte Pelletier, Ahmad Al-Bitar, Stéphane Mermoz, Simon Gascoin, Alexandre Bouvet, Mehdi Saqalli, Benoit Dewandel, Yvan Caballero, Shakeel Ahmed, Jean-Christophe Maréchal, and Yann Kerr. Detection of Irrigated Crops from Sentinel-1 and Sentinel-2 Data to Estimate Seasonal Groundwater Use in South India. Remote Sensing, 9(11) :1119, November 2017. doi : 10.3390/rs9111119.
- Pierric Ferrier, Philippe Crebassol, Gérard Dedieu, Olivier Hagolle, Aimé Meygret, Francesc Tinto, Yoram Yaniv, and Jacob Herscovitz. VEN $\mu$ S (Vegetation and environment monitoring on a new micro satellite). In 2010 IEEE International Geoscience and Remote Sensing Symposium, pages 3736–3739, July 2010. doi : 10.1109/IGARSS.2010.5652087.
- Jin Kathrine Fosli, A. Amarender Reddy, and Radhika Rani. The Policy of Free Electricity to Agriculture Sector : Implications and Perspectives of the Stakeholders in India. Journal of Development Policy and Practice, 6(2) :252–269, July 2021. ISSN 2455-1333. doi : 10.1177/24551333211025203.
- F. Frappart, K. Do Minh, J. L’Hermitte, A. Cazenave, G. Ramillien, T. Le Toan, and N. Mognard-Campbell. Water volume change in the lower Mekong from satellite altimetry and imagery data. Geophysical Journal International, 167(2) :570–584, November 2006. ISSN 0956-540X. doi : 10.1111/j.1365-246X.2006.03184.x.
- F. Frappart, F. Papa, A. Güntner, J. Tomasella, J. Pfeffer, G. Ramillien, T. Emilio, J. Schiatti, L. Seoane, J. da Silva Carvalho, D. Medeiros Moreira, M. P. Bonnet, and F. Seyler. The spatio-temporal variability of groundwater storage in the Amazon River Basin. Advances in Water Resources, 124 :41–52, February 2019. ISSN 0309-1708. doi : 10.1016/j.advwatres.2018.12.005.

- Frédéric Frappart and Guillaume Ramillien. Monitoring Groundwater Storage Changes Using the Gravity Recovery and Climate Experiment (GRACE) Satellite Mission : A Review. Remote Sensing, 10(6) :829, June 2018. doi : 10.3390/rs10060829.
- Lee-Lueng Fu, Edward J. Christensen, Charles A. Yamarone Jr., Michel Lefebvre, Yves Ménard, Michel Dorrer, and Philippe Escudier. TOPEX/POSEIDON mission overview. Journal of Geophysical Research : Oceans, 99(C12) :24369–24381, 1994. ISSN 2156-2202. doi : 10.1029/94JC01761.
- Alexandra Gemitzi and Venkat Lakshmi. Downscaling GRACE data to estimate groundwater use at the aquifer scale. page 5, 2017.
- Alexandra Gemitzi and Venkat Lakshmi. Estimating Groundwater Abstractions at the Aquifer Scale Using GRACE Observations. Geosciences, 8(11) :419, November 2018. doi : 10.3390/geosciences8110419.
- Nureldin A. A. Gido, Hadi Amin, Mohammad Bagherbandi, and Faramarz Nilfouroushan. Satellite Monitoring of Mass Changes and Ground Subsidence in Sudan’s Oil Fields Using GRACE and Sentinel-1 Data. Remote Sensing, 12(11) :1792, January 2020. doi : 10.3390/rs12111792.
- Richard O. Gilbert. 6.5 Sen’s Nonparametric Estimator of Slope. In Statistical Methods for Environmental Pollution Monitoring. John Wiley & Sons, February 1987. ISBN 978-0-471-28878-7.
- Mark Giordano and Karen G. Villholth. The Agricultural Groundwater Revolution : Opportunities and Threats to Development. CABI, 2007. ISBN 978-1-84593-172-8.
- Manuela Giroto, Gabriëlle J. M. De Lannoy, Rolf H. Reichle, and Matthew Rodell. Assimilation of gridded terrestrial water storage observations from GRACE into a land surface model. Water Resources Research, 52(5) :4164–4183, 2016. ISSN 1944-7973. doi : 10.1002/2015WR018417.
- C. J. Glendenning, F. F. van Ogtrop, A. K. Mishra, and R. W. Vervoort. Balancing watershed and local scale impacts of rain water harvesting in India—A review. Agricultural Water Management, 107 :1–13, May 2012. ISSN 0378-3774. doi : 10.1016/j.agwat.2012.01.011.
- M. A. Gleyzes, L. Perret, and P. Kubik. PLEIADES SYSTEM ARCHITECTURE AND MAIN PERFORMANCES. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XXXIX-B1 :537–542, July 2012. ISSN 2194-9034. doi : 10.5194/isprsarchives-XXXIX-B1-537-2012.
- Ben T. Gouweleeuw, Andreas Kvas, Christian Gruber, Animesh K. Gain, Thorsten Mayer-Gürr, Frank Flechtner, and Andreas Güntner. Daily GRACE gravity field solutions track major flood events in the Ganges–Brahmaputra Delta. Hydrology and Earth System Sciences, 22(5) :2867–2880, May 2018. ISSN 1027-5606. doi : 10.5194/hess-22-2867-2018.
- N. Guihéneuf, A. Boisson, O. Bour, B. Dewandel, J. Perrin, A. Dausse, M. Viossanges, S. Chandra, S. Ahmed, and J. C. Maréchal. Groundwater flows in weathered crystalline rocks : Impact of piezometric variations and depth-dependent fracture connectivity. Journal of Hydrology, 511 : 320–334, April 2014. ISSN 0022-1694. doi : 10.1016/j.jhydrol.2014.01.061.

- Florence Habets, Jérôme Molénat, Nadia Carluer, Olivier Douez, and Delphine Leenhardt. The cumulative impacts of small reservoirs on hydrology : A review. Science of The Total Environment, 643 :850–867, December 2018. ISSN 0048-9697. doi : 10.1016/j.scitotenv.2018.06.188.
- Olivier Hagolle, Mireille Huc, Camille Desjardins, Stefan Auer, and Rudolf Richter. MAJA Algorithm Theoretical Basis Document. December 2017. doi : 10.5281/zenodo.1209633.
- Khaled H. Hamed and A. Ramachandra Rao. A modified Mann-Kendall trend test for autocorrelated data. Journal of Hydrology, 204(1) :182–196, January 1998. ISSN 0022-1694. doi : 10.1016/S0022-1694(97)00125-X.
- Andy Hardy, Georgina Ettritch, Dónall E. Cross, Pete Bunting, Francis Liywalii, Jacob Sakala, Andrew Silumesii, Douglas Singini, Mark Smith, Tom Willis, and Chris J. Thomas. Automatic Detection of Open and Vegetated Water Bodies Using Sentinel 1 to Map African Malaria Vector Mosquito Breeding Habitats. Remote Sensing, 11(5) :593, January 2019. ISSN 2072-4292. doi : 10.3390/rs11050593.
- Tejasvi Hora, Veena Srinivasan, and Nandita B. Basu. The Groundwater Recovery Paradox in South India. Geophysical Research Letters, 46(16) :9602–9611, 2019. ISSN 1944-8007. doi : 10.1029/2019GL083525.
- Rasmus Houborg, Matthew Rodell, Bailing Li, Rolf Reichle, and Benjamin F. Zaitchik. Drought indicators based on model-assimilated Gravity Recovery and Climate Experiment (GRACE) terrestrial water storage observations. Water Resources Research, 48(7), 2012. ISSN 1944-7973. doi : 10.1029/2011WR011291.
- Suraj Jena, Rabindra Kumar Panda, Meenu Ramadas, Binayak P. Mohanty, Alok Kumar Samantaray, and Susanta Kishore Pattanaik. Characterization of groundwater variability using hydrological, geological, and climatic factors in data-scarce tropical savanna region of India. Journal of Hydrology : Regional Studies, 37 :100887, October 2021. ISSN 22145818. doi : 10.1016/j.ejrh.2021.100887.
- P. J. Jyolsna, B. V. N. P. Kambhammettu, and Saisrinivas Gorugantula. Application of random forest and multi-linear regression methods in downscaling GRACE derived groundwater storage changes. Hydrological Sciences Journal, 66(5) :874–887, April 2021. ISSN 0262-6667. doi : 10.1080/02626667.2021.1896719.
- Anuradha Karunakalage, Tandriila Sarkar, Suresh Kannaujiya, Prakash Chauhan, Pranshu Pranjali, Ajay Kumar Taloor, and Senthil Kumar. The appraisal of groundwater storage dwindling effect, by applying high resolution downscaling GRACE data in and around Mehsana district, Gujarat, India. Groundwater for Sustainable Development, 13 :100559, May 2021. ISSN 2352801X. doi : 10.1016/j.gsd.2021.100559.
- Jongyoun Kim and Terri S. Hogue. Improving Spatial Soil Moisture Representation Through Integration of AMSR-E and MODIS Products. IEEE Transactions on Geoscience and Remote Sensing, 50(2) :446–460, February 2012. ISSN 1558-0644. doi : 10.1109/TGRS.2011.2161318.
- Byoung Chul Ko, Hyeong Hun Kim, and Jae Yeal Nam. Classification of Potential Water Bodies Using Landsat 8 OLI and a Combination of Two Boosted Random Forest Classifiers. Sensors, 15(6) :13763–13777, June 2015. ISSN 1424-8220. doi : 10.3390/s150613763.

- Randal D. Koster, Sarith P. P. Mahanama, Ben Livneh, Dennis P. Lettenmaier, and Rolf H. Reichle. Skill in streamflow forecasts derived from large-scale estimates of soil moisture and snow. Nature Geoscience, 3(9) :613–616, September 2010. ISSN 1752-0908. doi : 10.1038/ngeo944.
- Gerhard Krieger, Manfred Zink, Markus Bachmann, Benjamin Bräutigam, Daniel Schulze, Michele Martone, Paola Rizzoli, Ulrich Steinbrecher, John Walter Antony, Francesco De Zan, Irena Hajsek, Kostas Papathanassiou, Florian Kugler, Marc Rodriguez Cassola, Marwan Younis, Stefan Baumgartner, Paco López-Dekker, Pau Prats, and Alberto Moreira. TanDEM-X : A radar interferometer with two formation-flying satellites. Acta Astronautica, 89 :83–98, August 2013. ISSN 0094-5765. doi : 10.1016/j.actaastro.2013.03.008.
- Maarten S. Krol, Marjella J. de Vries, Pieter R. van Oel, and José Carlos de Araújo. Sustainability of Small Reservoirs and Large Scale Water Availability Under Current Conditions and Climate Change. Water Resources Management, 25(12) :3017–3026, September 2011. ISSN 1573-1650. doi : 10.1007/s11269-011-9787-0.
- E. Kurtenbach, A. Eicker, T. Mayer-Gürr, M. Holschneider, M. Hayn, M. Fuhrmann, and J. Kusche. Improved daily GRACE gravity field solutions using a Kalman smoother. Journal of Geodynamics, 59–60 :39–48, September 2012. ISSN 0264-3707. doi : 10.1016/j.jog.2012.02.006.
- Enrico Kurtenbach, Torsten Mayer-Gürr, and Annette Eicker. Deriving daily snapshots of the Earth’s gravity field from GRACE L1B data using Kalman filtering. Geophysical Research Letters, 36(17), 2009. ISSN 1944-8007. doi : 10.1029/2009GL039564.
- Juliette Lambin, Rosemary Morrow, Lee-Lueng Fu, Josh K. Willis, Hans Bonekamp, John Lillibridge, Jacqueline Perbos, Gérard Zaouche, Parag Vaze, Walid Bannoura, François Parisot, Eric Thouvenot, Sophie Coutin-Faye, Eric Lindstrom, and Mike Mignogno. The OSTM/Jason-2 Mission. Marine Geodesy, 33(sup1) :4–25, August 2010. ISSN 0149-0419. doi : 10.1080/01490419.2010.491030.
- F. W. Landerer and S. C. Swenson. Accuracy of scaled GRACE terrestrial water storage estimates. Water Resources Research, 48(4), 2012. ISSN 1944-7973. doi : 10.1029/2011WR011453.
- L. Lebègue, E. Cazala-Hourcade, F. Languille, S. Artigues, and O. Melet. CO3D, A WORLD-WIDE ONE ONE-METER ACCURACY DEM FOR 2025. In ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, volume XLIII-B1-2020, pages 299–304. Copernicus GmbH, August 2020. doi : 10.5194/isprs-archives-XLIII-B1-2020-299-2020.
- Bernhard Lehner, Catherine Reidy Liermann, Carmen Revenga, Charles Vörösmarty, Balazs Fekete, Philippe Crouzet, Petra Döll, Marcel Endejan, Karen Frenken, Jun Magome, Christer Nilsson, James C. Robertson, Raimund Rödel, Nikolai Sindorf, and Dominik Wisser. High-resolution mapping of the world’s reservoirs and dams for sustainable river-flow management. Frontiers in Ecology and the Environment, 9(9) :494–502, 2011. ISSN 1540-9309. doi : 10.1890/100125.
- J. M. Lemoine, B. Meyssignac, M. Manda, E. Samain, S. Bourgogne, A. Blazquez, G. Balmino, L. Louise, and J. Michaud. MARVEL Mission Proposal : The Latest Update. 2020 :G020–08, December 2020.
- Jean-Michel Lemoine and Mioara Manda. The MARVEL gravity and reference frame mission proposal. page 13359, May 2020.

- Yao Li, Huilin Gao, Michael F. Jasinski, Shuai Zhang, and Jeremy D. Stoll. Deriving High-Resolution Reservoir Bathymetry From ICESat-2 Prototype Photon-Counting Lidar and Landsat Imagery. IEEE Transactions on Geoscience and Remote Sensing, 57(10) :7883–7893, October 2019. ISSN 0196-2892, 1558-0644. doi : 10.1109/TGRS.2019.2917012.
- Kai Liu, Chunqiao Song, Jida Wang, Linghong Ke, Yunqiang Zhu, Jingying Zhu, Ronghua Ma, and Zhu Luo. Remote Sensing-Based Modeling of the Bathymetry and Water Storage for Channel-Type Reservoirs Worldwide. Water Resources Research, 56(11) :e2020WR027147, 2020. ISSN 1944-7973. doi : 10.1029/2020WR027147.
- Meng Liu, Ronglin Tang, Zhao-Liang Li, Yunjun Yao, and Guangjian Yan. Global Land Surface Evapotranspiration Estimation From Meteorological and Satellite Data Using the Support Vector Machine and Semiempirical Algorithm. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 11(2) :513–521, February 2018. ISSN 2151-1535. doi : 10.1109/JSTARS.2017.2788462.
- Di Long, Bridget R. Scanlon, Laurent Longuevergne, Alexander Y. Sun, D. Nelun Fernando, and Himanshu Save. GRACE satellite monitoring of large depletion in water storage in response to the 2011 drought in Texas. Geophysical Research Letters, 40(13) :3395–3401, 2013. ISSN 1944-8007. doi : 10.1002/grl.50655.
- Di Long, Laurent Longuevergne, and Bridget R. Scanlon. Global analysis of approaches for deriving total water storage changes from GRACE satellites. Water Resources Research, 51(4) :2574–2594, 2015. ISSN 1944-7973. doi : 10.1002/2014WR016853.
- Laurent Longuevergne, Bridget R. Scanlon, and Clark R. Wilson. GRACE Hydrological estimates for small basins : Evaluating processing approaches on the High Plains Aquifer, USA. Water Resources Research, 46(11), 2010. ISSN 1944-7973. doi : 10.1029/2009WR008564.
- B. D. Loomis, S. B. Luthcke, and T. J. Sabaka. Regularization and error characterization of GRACE mascons. Journal of Geodesy, 93(9) :1381–1398, September 2019. ISSN 1432-1394. doi : 10.1007/s00190-019-01252-y.
- Kyle MacRitchie. README Document for the Tropical Rainfall Measurement Mission (TRMM) Version 7, 2017.
- J. Magome, H. Ishidaira, and K. Takeuchi. Method for satellite monitoring of water storage in reservoirs for efficient regional water management. In G. Bloschl, S. Franks, M. Kumagai, K. Musiake, and D. Rosbjerg, editors, Water Resources Systems-Hydrological Risk, Management and Development, number 281, pages 303–310, Wallingford, 2003. Int Assoc Hydrological Sciences. ISBN 978-1-901502-32-9.
- Sylvain Mangiarotti, Flavie Le Jean, Mireille Huc, and Christophe Letellier. Global modeling of aggregated and associated chaotic dynamics. Chaos, Solitons & Fractals, 83 :82–96, February 2016. ISSN 09600779. doi : 10.1016/j.chaos.2015.11.031.
- J. C. Maréchal, B. Dewandel, and K. Subrahmanyam. Use of hydraulic tests at different scales to characterize fracture network properties in the weathered-fractured layer of a hard rock aquifer. Water Resources Research, 40(11), 2004. ISSN 1944-7973. doi : 10.1029/2004WR003137.

- J.C. Maréchal, B. Dewandel, S. Ahmed, L. Galeazzi, and F.K. Zaidi. Combined estimation of specific yield and natural recharge in a semi-arid groundwater basin with irrigated agriculture. Journal of Hydrology, 329(1-2) :281–293, September 2006. ISSN 00221694. doi : 10.1016/j.jhydrol.2006.02.022.
- Jean-Christophe Maréchal, Adrien Selles, Benoit Dewandel, Alexandre Boisson, Jérôme Perrin, and Shakeel Ahmed. An Observatory of Groundwater in Crystalline Rock Aquifers Exposed to a Changing Environment : Hyderabad, India. Vadose Zone Journal, 17(1) :180076, 2018. ISSN 1539-1663. doi : 10.2136/vzj2018.04.0076.
- Brecht Martens, Diego G. Miralles, Hans Lievens, Robin van der Schalie, Richard A. M. de Jeu, Diego Fernández-Prieto, Hylke E. Beck, Wouter A. Dorigo, and Niko E. C. Verhoest. GLEAM v3 : Satellite-based land evaporation and root-zone soil moisture. Geoscientific Model Development, 10(5) :1903–1925, May 2017. ISSN 1991-959X. doi : 10.5194/gmd-10-1903-2017.
- Luca Massotti, Christian Siemes, Günther March, Roger Haagmans, and Pierluigi Silvestrin. Next Generation Gravity Mission Elements of the Mass Change and Geoscience International Constellation : From Orbit Selection to Instrument and Mission Design. Remote Sensing, 13(19) :3935, January 2021. ISSN 2072-4292. doi : 10.3390/rs13193935.
- S. Massuel, J. Perrin, C. Mascré, W. Mohamed, A. Boisson, and S. Ahmed. Managed aquifer recharge in South India : What to expect from small percolation tanks in hard rock? Journal of Hydrology, 512 :157–167, May 2014. ISSN 0022-1694. doi : 10.1016/j.jhydrol.2014.02.062.
- Sylvain Massuel, Jérôme Perrin, Cédric Mascré, Mohamed Wajid, and Benoit Dewandel. Percolation efficiency of an artificial tank in semi-arid hard rock aquifer context, South India. page 9, 2008.
- S. K. McFEETERS. The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features. International Journal of Remote Sensing, 17(7) :1425–1432, May 1996. ISSN 0143-1161. doi : 10.1080/01431169608948714.
- Olivier Merlin, Frédéric Jacob, Jean-Pierre Wigneron, Jeffrey Walker, and Ghani Chehbouni. Multidimensional Disaggregation of Land Surface Temperature Using High-Resolution Red, Near-Infrared, Shortwave-Infrared, and Microwave-L Bands. IEEE Transactions on Geoscience and Remote Sensing, 50(5) :1864–1880, May 2012. ISSN 1558-0644. doi : 10.1109/TGRS.2011.2169802.
- Olivier Merlin, Maria José Escorihuela, Miquel Aran Mayoral, Olivier Hagolle, Ahmad Al Bitar, and Yann Kerr. Self-calibrated evaporation-based disaggregation of SMOS soil moisture : An evaluation study at 3 km and 100 m resolution in Catalunya, Spain. Remote Sensing of Environment, 130 :25–38, March 2013. ISSN 00344257. doi : 10.1016/j.rse.2012.11.008.
- Olivier Merlin, Yoann Malbêteau, Youness Notfi, Stefan Bacon, Salah Er-Raki Saïd Khabba, and Lionel Jarlan. Performance Metrics for Soil Moisture Downscaling Methods : Application to DISPATCH Data in Central Morocco. Remote Sensing, 7(4) :3783–3807, April 2015. doi : 10.3390/rs70403783.
- Dominic Merriott. Factors associated with the farmer suicide crisis in India. Journal of Epidemiology and Global Health, 6(4) :217–227, April 2016. ISSN 2210-6014. doi : 10.1016/j.jegh.2016.03.003.

- Milewski, Thomas, Seyoum, and Rasmussen. Spatial Downscaling of GRACE TWSA Data to Identify Spatiotemporal Groundwater Level Trends in the Upper Floridan Aquifer, Georgia, USA. Remote Sensing, 11(23) :2756, November 2019. ISSN 2072-4292. doi : 10.3390/rs11232756.
- D. G. Miralles, T. R. H. Holmes, R. a. M. De Jeu, J. H. Gash, A. G. C. A. Meesters, and A. J. Dolman. Global land-surface evaporation estimated from satellite-based observations. Hydrology and Earth System Sciences, 15(2) :453–469, February 2011. ISSN 1027-5606. doi : 10.5194/hess-15-453-2011.
- Michelle E. Miro and James S. Famiglietti. Downscaling GRACE Remote Sensing Datasets to High-Resolution Groundwater Storage Change Maps of California’s Central Valley. Remote Sensing, 10(1) :143, January 2018. doi : 10.3390/rs10010143.
- J. L. Monteith. Evaporation and environment. Symposia of the Society for Experimental Biology, 19 :205–234, 1965.
- Qiaozhen Mu, Faith Ann Heinsch, Maosheng Zhao, and Steven W. Running. Development of a global evapotranspiration algorithm based on MODIS and global meteorology data. Remote Sensing of Environment, 111(4) :519–536, December 2007. ISSN 0034-4257. doi : 10.1016/j.rse.2007.04.015.
- Qiaozhen Mu, Maosheng Zhao, and Steven W. Running. Improvements to a MODIS global terrestrial evapotranspiration algorithm. Remote sensing of environment, 115(8) :1781–1800, 2011.
- S. Muralikrishnan, B. Narender, Shashivardhan Reddy, and Abhijit Pillai. Evaluation of Indian National DEM from Cartosat-1 Data, 2011.
- Madeleine Nicolas. Impact de l’hétérogénéité Sur La Recharge Naturelle et Artificielle Des Aquifères Cristallins Altérés et Fracturés. PhD thesis, UNIVERSITÉ DE RENNES 1, 2019.
- Madeleine Nicolas, Olivier Bour, Adrien Selles, Benoit Dewandel, Vincent Bailly-Comte, Subash Chandra, Shakeel Ahmed, and Jean-Christophe Maréchal. Managed Aquifer Recharge in fractured crystalline rock aquifers : Impact of horizontal preferential flow on recharge dynamics. Journal of Hydrology, 573 :717–732, June 2019. ISSN 0022-1694. doi : 10.1016/j.jhydrol.2019.04.003.
- Wanshu Nie, Benjamin F. Zaitchik, Matthew Rodell, Sujay V. Kumar, Kristi R. Arsenault, Bailing Li, and Augusto Getirana. Assimilating GRACE Into a Land Surface Model in the Presence of an Irrigation-Induced Groundwater Trend. Water Resources Research, 55(12) :11274–11294, 2019. ISSN 1944-7973. doi : 10.1029/2019WR025363.
- N. H. Nielsen, M. R. A. Larsen, and S. F. Rasmussen. Development of a screening method to assess flood risk on Danish national roads and highway systems. Water Science and Technology, 63(12) :2957–2966, June 2011. ISSN 0273-1223, 1996-9732. doi : 10.2166/wst.2011.157.
- Shaowei Ning, Hiroshi Ishidaira, and Jie Wang. Statistical Downscaling of Grace-Derived Terrestrial Water Storage Using Satellite and Gldas Products. Journal of Japan Society of Civil Engineers, Series B1 (Hydraulic Engineering), 70(4) :I\_133–I\_138, 2014. doi : 10.2208/jscejhe.70.I\_133.



- Stella Ofori-Ampofo, Charlotte Pelletier, and Stefan Lang. Crop Type Mapping from Optical and Radar Time Series Using Attention-Based Deep Learning. *Remote Sensing*, 13(22) :4668, January 2021. doi : 10.3390/rs13224668.
- J. Overgaard, D. Rosbjerg, and M. B. Butts. Land-surface modelling in hydrological perspective – a review. *Biogeosciences*, 3(2) :229–241, May 2006. ISSN 1726-4170. doi : 10.5194/bg-3-229-2006.
- Yann Pageot, Frédéric Baup, Jordi Inglada, Nicolas Baghdadi, and Valérie Demarez. Detection of Irrigated and Rainfed Crops in Temperate Areas Using Sentinel-1 and Sentinel-2 Time Series. *Remote Sensing*, 12(18) :3044, January 2020. ISSN 2072-4292. doi : 10.3390/rs12183044.
- K Palanisami, Ruth Meinzen-Dick, and Mark Giordano. Climate Change and Water Supplies : Options for Sustaining Tank Irrigation Potential in India. *Economic and Political Weekly*, 45 (26/27) :183–190, 2010. ISSN 0012-9976.
- Dileep K. Panda and John Wahr. Spatiotemporal evolution of water storage changes in India from the updated GRACE-derived gravity records. *Water Resources Research*, 52(1) :135–149, 2016. ISSN 1944-7973. doi : 10.1002/2015WR017797.
- Isabelle Panet, Valentin Mikhailov, Michel Diament, Fred Pollitz, Geoffrey King, Olivier De Viron, Matthias Holschneider, Richard Biancale, and Jean-Michel Lemoine. Coseismic and post-seismic signatures of the Sumatra 2004 December and 2005 March earthquakes in GRACE satellite gravity. *Geophysical Journal International*, 171(1) :177–190, October 2007. ISSN 0956-540X. doi : 10.1111/j.1365-246X.2007.03525.x.
- Fabrice Papa, Frédéric Frappart, Yoann Malbeteau, Mohammad Shamsudduha, Venugopal Vuruputur, Muddu Sekhar, Guillaume Ramillien, Catherine Prigent, Filipe Aires, Rajesh Kumar Pandey, Sujit Bala, and Stephane Calmant. Satellite-derived surface and sub-surface water storage in the Ganges–Brahmaputra River Basin. *Journal of Hydrology : Regional Studies*, 4 : 15–35, September 2015. ISSN 2214-5818. doi : 10.1016/j.ejrh.2015.03.004.
- C. Pascal, S. Ferrant, A. Selles, J.-C. Maréchal, S. Gascoïn, and O. Merlin. High-Resolution Mapping of Rainwater Harvesting System Capacity from Satellite Derived Products in South India. In *2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS*, pages 7011–7014, July 2021. doi : 10.1109/IGARSS47720.2021.9553131.
- Santiago Peña-Luque, Sylvain Ferrant, Mauricio C. R. Cordeiro, Thomas Ledauphin, Jerome Maxant, and Jean-Michel Martinez. Sentinel-1&2 Multitemporal Water Surface Detection Accuracies, Evaluated at Regional and Reservoirs Level. *Remote Sensing*, 13(16) :3279, August 2021. ISSN 2072-4292. doi : 10.3390/rs13163279.
- Jean-François Pekel, Andrew Cottam, Noel Gorelick, and Alan S. Belward. High-resolution mapping of global surface water and its long-term changes. *Nature*, 540(7633) :418–422, December 2016. ISSN 1476-4687. doi : 10.1038/nature20584.
- Thierry Pellarin, Samuel Louvet, Claire Gruhier, Guillaume Quantin, and Cedric Legout. A simple and effective method for correcting soil moisture and precipitation estimates using AMSR-E measurements. *Remote Sensing of Environment*, 136 :28–36, September 2013. ISSN 00344257. doi : 10.1016/j.rse.2013.04.011.

- Thierry Pellarin, Carlos Román-Cascón, Christian Baron, Rajat Bindlish, Luca Brocca, Pierre Camberlin, Diego Fernández-Prieto, Yann H. Kerr, Christian Massari, Jeremy Panthou, Benoit Perrimond, Nathalie Philippon, and Guillaume Quantin. The Precipitation Inferred from Soil Moisture (PrISM) Near Real-Time Rainfall Product : Evaluation and Comparison. Remote Sensing, 12(3) :481, January 2020. doi : 10.3390/rs12030481.
- Charlotte Pelletier, Silvia Valero, Jordi Inglada, Nicolas Champion, Claire Marais Sicre, and Gérard Dedieu. Effect of Training Class Label Noise on Classification Performances for Land Cover Mapping with Satellite Image Time Series. Remote Sensing, 9(2) :173, February 2017. ISSN 2072-4292. doi : 10.3390/rs9020173.
- Liqing Peng. Variability of Atmospheric Evaporative Demand. 2019.
- Liqing Peng, Zhenzhong Zeng, Zhongwang Wei, Anping Chen, Eric F. Wood, and Justin Sheffield. Determinants of the ratio of actual to potential evapotranspiration. Global Change Biology, 25 (4) :1326–1343, 2019. ISSN 1365-2486. doi : 10.1111/gcb.14577.
- Howard Latimer Penman and Bernard Augustus Keen. Natural evaporation from open water, bare soil and grass. Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences, 193(1032) :120–145, April 1948. doi : 10.1098/rspa.1948.0037.
- J. Perrin, S. Ferrant, S. Massuel, B. Dewandel, J.C. Maréchal, S. Aulong, and S. Ahmed. Assessing water availability in a semi-arid watershed of southern India using a semi-distributed model. Journal of Hydrology, 460–461 :143–155, August 2012. ISSN 00221694. doi : 10.1016/j.jhydrol.2012.07.002.
- Jerome Perrin, Shakeel Ahmed, and Daniel Hunkeler. The effects of geological heterogeneities and piezometric fluctuations on groundwater flow and chemistry in a hard-rock aquifer, southern India. Hydrogeology Journal, 19(6) :1189, May 2011. ISSN 1435-0157. doi : 10.1007/s10040-011-0745-y.
- Gautam Pingle. Irrigation in Telangana : The Rise and Fall of Tanks. Economic & Political Weekly, page 9, 2011.
- A. Poulouvalis, M. Anadranistakis, A. Liakatas, S. Alexandris, and P. Kerkides. Semi-empirical approach for estimating actual evapotranspiration in Greece. Agricultural Water Management, 51(2) :143–152, October 2001. ISSN 0378-3774. doi : 10.1016/S0378-3774(01)00121-4.
- C. H. B. Priestley and R. J. Taylor. On the Assessment of Surface Heat Flux and Evaporation Using Large-Scale Parameters. Monthly Weather Review, 100(2) :81–92, February 1972. ISSN 1520-0493, 0027-0644. doi : 10.1175/1520-0493(1972)100<0081:OTAOSH>2.3.CO;2.
- G. L. Ramillien, L. Seoane, F. Frappart, R. Biancale, S. Gratton, X. Vasseur, and S. Bourgoigne. Constrained Regional Recovery of Continental Water Mass Time-variations from GRACE-based Geopotential Anomalies over South America. Surveys in Geophysics, 33(5) :887–905, September 2012. ISSN 1573-0956. doi : 10.1007/s10712-012-9177-z.
- P R Reddy. An over view of Irrigation Tanks Rehabilitation in semi arid hard rock terrain. page 7, October 2015.

- Rolf H. Reichle, Gabrielle J. M. De Lannoy, Qing Liu, Joseph V. Ardizzone, Andreas Colliander, Austin Conaty, Wade Crow, Thomas J. Jackson, Lucas A. Jones, John S. Kimball, Randal D. Koster, Sarith P. Mahanama, Edmond B. Smith, Aaron Berg, Simone Bircher, David Bosch, Todd G. Caldwell, Michael Cosh, Ángel González-Zamora, Chandra D. Holifield Collins, Karsten H. Jensen, Stan Livingston, Ernesto Lopez-Baeza, José Martínez-Fernández, Heather McNairn, Mahta Moghaddam, Anna Pacheco, Thierry Pellarin, John Prueger, Tracy Rowlandson, Mark Seyfried, Patrick Starks, Zhongbo Su, Marc Thibeault, Rogier van der Velde, Jeffrey Walker, Xiaoling Wu, and Yijian Zeng. Assessment of the SMAP Level-4 Surface and Root-Zone Soil Moisture Product Using In Situ Measurements. *Journal of Hydrometeorology*, 18(10) : 2621–2645, July 2017. ISSN 1525-755X. doi : 10.1175/JHM-D-17-0063.1.
- Sarah Ringerud, Christian D. Kummerow, and Christa D. Peters-Lidard. A Semi-Empirical Model for Computing Land Surface Emissivity in the Microwave Region. *IEEE Transactions on Geoscience and Remote Sensing*, 53(4) :1935–1946, April 2015. ISSN 1558-0644. doi : 10.1109/TGRS.2014.2351232.
- Riccardo E. M. Riva, Brian C. Gunter, Timothy J. Urban, Bert L. A. Vermeersen, Roderik C. Lindenbergh, Michiel M. Helsen, Jonathan L. Bamber, Roderik S. W. van de Wal, Michiel R. van den Broeke, and Bob E. Schutz. Glacial Isostatic Adjustment over Antarctica from combined ICESat and GRACE satellite data. *Earth and Planetary Science Letters*, 288(3) :516–523, November 2009. ISSN 0012-821X. doi : 10.1016/j.epsl.2009.10.013.
- M. Rodell, P. R. Houser, U. Jambor, J. Gottschalck, K. Mitchell, C.-J. Meng, K. Arsenault, B. Cosgrove, J. Radakovich, M. Bosilovich, J. K. Entin, J. P. Walker, D. Lohmann, and D. Toll. The Global Land Data Assimilation System. *Bulletin of the American Meteorological Society*, 85(3) :381–394, March 2004. ISSN 0003-0007, 1520-0477. doi : 10.1175/BAMS-85-3-381.
- M. Rodell, B. F. Chao, A. Y. Au, J. S. Kimball, and K. C. McDonald. Global Biomass Variation and Its Geodynamic Effects : 1982–98. *Earth Interactions*, 9(2) :1–19, April 2005. ISSN 1087-3562. doi : 10.1175/EI126.1.
- M. Rodell, J. S. Famiglietti, D. N. Wiese, J. T. Reager, H. K. Beaudoin, F. W. Landerer, and M.-H. Lo. Emerging trends in global freshwater availability. *Nature*, 557(7707) :651–659, May 2018. ISSN 1476-4687. doi : 10.1038/s41586-018-0123-1.
- Matthew Rodell, Jianli Chen, Hiroko Kato, James S. Famiglietti, Joe Nigro, and Clark R. Wilson. Estimating groundwater storage changes in the Mississippi River basin (USA) using GRACE. *Hydrogeology Journal*, 15(1) :159–166, February 2007. ISSN 1431-2174. doi : 10.1007/s10040-006-0103-7.
- Matthew Rodell, Isabella Velicogna, and James S. Famiglietti. Satellite-based estimates of groundwater depletion in India. *Nature*, 460(7258) :999–1002, August 2009. ISSN 1476-4687. doi : 10.1038/nature08238.
- N. Rodríguez-Fernández, A. Mialon, O. Merlin, C. Suere, F. Cabot, A. Khazâal, J. Costeraste, B. Palacin, R. Rodríguez-Suquet, T. Tournier, T. Decoopman, É Anterrieu, M. Colom, J. Morel, Y. Kerr, B. Rougé, J. Boutin, G. Picard, T. Pellarin, M. Escorihuela, A. Al Bitar, and P. Richaume. SMOS-HR : A High Resolution L-Band Passive Radiometer for Earth Science and Applications. *undefined*, 2019.

- Ignacio Rodríguez-Iturbe and Amilcare Porporato. Ecohydrology of Water-Controlled Ecosystems : Soil Moisture and Plant Dynamics. Cambridge University Press, February 2007. ISBN 978-1-139-44195-7.
- Steven W. Running, Qiaozhen Mu, Maosheng Zhao, and Alvaro Moreno. User's Guide MODIS Global Terrestrial Evapotranspiration (ET) Product (MOD16A2/A3 and Year-end Gap-filled MOD16A2GF/A3GF) NASA Earth Observing System MODIS Land Algorithm (For Collection 6). [https://lpdaac.usgs.gov/documents/494/MOD16\\_User\\_Guide\\_V6.pdf](https://lpdaac.usgs.gov/documents/494/MOD16_User_Guide_V6.pdf), 2019.
- Zofia Rzepecka and Monika Birylo. Groundwater Storage Changes Derived from GRACE and GLDAS on Smaller River Basins—A Case Study in Poland. Geosciences, 10(4) :124, March 2020. ISSN 2076-3263. doi : 10.3390/geosciences10040124.
- Sabah Sabaghy, Jeffrey P. Walker, Luigi J. Renzullo, Ruzbeh Akbar, Steven Chan, Julian Chaubell, Narendra Das, R. Scott Dunbar, Dara Entekhabi, Anouk Gevaert, Thomas J. Jackson, Alexander Loew, Olivier Merlin, Mahta Moghaddam, Jian Peng, Jinzheng Peng, Jeffrey Piepmeyer, Christoph Rüdiger, Vivien Stefan, Xiaoling Wu, Nan Ye, and Simon Yueh. Comprehensive analysis of alternative downscaled soil moisture products. Remote Sensing of Environment, 239 : 111586, March 2020. ISSN 0034-4257. doi : 10.1016/j.rse.2019.111586.
- Hossein Sahour, Mohamed Sultan, Mehdi Vazifedan, Karem Abdelmohsen, Sita Karki, John A. Yellich, Esayas Gebremichael, Fahad Alshehri, and Tamer M. Elbayoumi. Statistical Applications to Downscale GRACE-Derived Terrestrial Water Storage Data and to Fill Temporal Gaps. Remote Sensing, 12(3) :533, February 2020. ISSN 2072-4292. doi : 10.3390/rs12030533.
- Nilda Sánchez, Ángel González-Zamora, María Piles, and José Martínez-Fernández. A New Soil Moisture Agricultural Drought Index (SMADI) Integrating MODIS and SMOS Products : A Case of Study over the Iberian Peninsula. Remote Sensing, 8(4) :287, April 2016. ISSN 2072-4292. doi : 10.3390/rs8040287.
- Himanshu Save, Srinivas Bettadpur, and Byron D. Tapley. High-resolution CSR GRACE RL05 mascons. Journal of Geophysical Research : Solid Earth, 121(10) :7547–7569, 2016. ISSN 2169-9356. doi : 10.1002/2016JB013007.
- Bridget R. Scanlon, Zizhan Zhang, Himanshu Save, David N. Wiese, Felix W. Landerer, Di Long, Laurent Longuevergne, and Jianli Chen. Global evaluation of new GRACE mascon products for hydrologic applications. Water Resources Research, 52(12) :9412–9429, 2016. ISSN 1944-7973. doi : 10.1002/2016WR019494.
- R. Schmidt, F. Flechtner, U. Meyer, K.-H. Neumayer, Ch. Dahle, R. König, and J. Kusche. Hydrological Signals Observed by the GRACE Satellites. Surveys in Geophysics, 29(4) :319–334, October 2008. ISSN 1573-0956. doi : 10.1007/s10712-008-9033-3.
- M. Schumacher, E. Forootan, A. I. J. M. van Dijk, H. Müller Schmied, R. S. Crosbie, J. Kusche, and P. Döll. Improving drought simulations within the Murray-Darling Basin by combined calibration/assimilation of GRACE data into the WaterGAP Global Hydrology Model. Remote Sensing of Environment, 204 :212–228, January 2018. ISSN 0034-4257. doi : 10.1016/j.rse.2017.10.029.
- Adrien Selles and Jean-Christophe Maréchal. Irrigation : L'Inde en quête d'une « révolution bleue ». <http://theconversation.com/irrigation-linde-en-quete-dune-revolution-bleue-180732>, 2022.

- Sonia I. Seneviratne, Thierry Corti, Edouard L. Davin, Martin Hirschi, Eric B. Jaeger, Irene Lehner, Boris Orlowsky, and Adriaan J. Teuling. Investigating soil moisture–climate interactions in a changing climate : A review. Earth-Science Reviews, 99(3) :125–161, May 2010. ISSN 0012-8252. doi : 10.1016/j.earscirev.2010.02.004.
- K.-W. Seo, C. R. Wilson, J. S. Famiglietti, J. L. Chen, and M. Rodell. Terrestrial water mass load changes from Gravity Recovery and Climate Experiment (GRACE). Water Resources Research, 42(5), 2006. ISSN 1944-7973. doi : 10.1029/2005WR004255.
- Wondwosen Seyoum, Dongjae Kwon, and Adam Milewski. Downscaling GRACE TWSA Data into High-Resolution Groundwater Level Anomaly Using Machine Learning-Based Models in a Glacial Aquifer System. Remote Sensing, 11(7) :824, April 2019. ISSN 2072-4292. doi : 10.3390/rs11070824.
- Wondwosen M. Seyoum and Adam M. Milewski. Improved methods for estimating local terrestrial water dynamics from GRACE in the Northern High Plains. Advances in Water Resources, 110 : 279–290, December 2017. ISSN 0309-1708. doi : 10.1016/j.advwatres.2017.10.021.
- Abhishek Sharma. Rethinking Tanks : Opportunities for Revitalizing Irrigation Tanks : Empirical Findings from Ananthapur District, Andhra Pradesh, India. International Water Management Institute, Colombo, Sri Lanka, 2003. ISBN 978-92-9090-526-4.
- Justin Sheffield and Eric F. Wood. Drought : Past Problems and Future Scenarios. Routledge, London, March 2011. ISBN 978-1-84977-525-0. doi : 10.4324/9781849775250.
- Prepared Suhung Shen, Justin Sheffield, Liqing Peng, and Ashley Heath. RM-OBS/PU Potential Evapotranspiration and Supporting Forcing L4 3-Hourly 0.25x0.25 degree V002. page 16, 2021.
- S. Siebert, J. Burke, J. M. Faures, K. Frenken, J. Hoogeveen, P. Döll, and F. T. Portmann. Groundwater use for irrigation – a global inventory. Hydrology and Earth System Sciences, 14 (10) :1863–1880, October 2010. ISSN 1027-5606. doi : 10.5194/hess-14-1863-2010.
- Kunwar P. Singh, Amrita Malik, Vinod K. Singh, Dinesh Mohan, and Sarita Sinha. Chemometric analysis of groundwater quality data of alluvial aquifer of Gangetic plain, North India. Analytica Chimica Acta, 550(1) :82–91, September 2005. ISSN 0003-2670. doi : 10.1016/j.aca.2005.06.056.
- Leelambar Singh and Subbarayan Saravanan. Satellite-derived GRACE groundwater storage variation in complex aquifer system in India. Sustainable Water Resources Management, 6(3) :43, May 2020. ISSN 2363-5045. doi : 10.1007/s40899-020-00399-3.
- V. Srinivasan, S. Thompson, K. Madhyastha, G. Penny, K. Jeremiah, and S. Lele. Why is the Arkavathy River drying ? A multiple-hypothesis approach in a data-scarce region. Hydrology and Earth System Sciences, 19(4) :1905–1917, April 2015. ISSN 1027-5606. doi : 10.5194/hess-19-1905-2015.
- Alexander Y. Sun. Predicting groundwater level changes using GRACE data. Water Resources Research, 49(9) :5900–5912, 2013. ISSN 1944-7973. doi : 10.1002/wrcr.20421.
- Alexander Y. Sun, Ronald Green, Matthew Rodell, and Sean Swenson. Inferring aquifer storage parameters using satellite and in situ measurements : Estimation under uncertainty. Geophysical Research Letters, 37(10), 2010. ISSN 1944-8007. doi : 10.1029/2010GL043231.

- Alexander Y. Sun, Ronald Green, Sean Swenson, and Matthew Rodell. Toward calibration of regional groundwater models using GRACE data. Journal of Hydrology, 422–423 :1–9, February 2012. ISSN 00221694. doi : 10.1016/j.jhydrol.2011.10.025.
- Alexander Y. Sun, Bridget R. Scanlon, Zizhan Zhang, David Walling, Soumendhra N. Bhanja, Abhijit Mukherjee, and Zhi Zhong. Combining Physically-Based Modeling and Deep Learning for Fusing GRACE Satellite Data : Can We Learn from Mismatch ? Water Resources Research, 55(2) :1179–1195, February 2019. ISSN 0043-1397, 1944-7973. doi : 10.1029/2018WR023333.
- Tetsushi Tachikawa, Masami Hato, Manabu Kaku, and Akira Iwasaki. Characteristics of ASTER GDEM version 2. In 2011 IEEE International Geoscience and Remote Sensing Symposium, pages 3657–3660, July 2011. doi : 10.1109/IGARSS.2011.6050017.
- T. Tadono, H. Ishida, F. Oda, S. Naito, K. Minakawa, and H. Iwamoto. Precise Global DEM Generation by ALOS PRISM. In ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, volume II-4, pages 71–76. Copernicus GmbH, April 2014. doi : 10.5194/isprsannals-II-4-71-2014.
- J. Takaku, T. Tadono, and K. Tsutsui. Generation of High Resolution Global DSM from ALOS PRISM. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XL-4 :243–248, April 2014. ISSN 2194-9034. doi : 10.5194/isprsarchives-XL-4-243-2014.
- B. D. Tapley, S. Bettadpur, M. Watkins, and C. Reigber. The gravity recovery and climate experiment : Mission overview and early results. Geophysical Research Letters, 31(9), 2004a. ISSN 1944-8007. doi : 10.1029/2004GL019920.
- Byron D. Tapley, Srinivas Bettadpur, John C. Ries, Paul F. Thompson, and Michael M. Watkins. GRACE Measurements of Mass Variability in the Earth System. Science, 305(5683) :503–505, July 2004b. ISSN 0036-8075, 1095-9203. doi : 10.1126/science.1099192.
- Siyuan Tian, Paul Tregoning, Luigi J. Renzullo, Albert I. J. M. van Dijk, Jeffrey P. Walker, Valentijn R. N. Pauwels, and Sébastien Allgeyer. Improved water balance component estimates through joint assimilation of GRACE water storage and SMOS soil moisture retrievals. Water Resources Research, 53(3) :1820–1840, 2017. ISSN 1944-7973. doi : 10.1002/2016WR019641.
- V. M. Tiwari, J. Wahr, and S. Swenson. Dwindling groundwater resources in northern India, from satellite gravity observations. Geophysical Research Letters, 36(18), 2009. ISSN 1944-8007. doi : 10.1029/2009GL039401.
- Mirela G. Tulbure, Mark Broich, Stephen V. Stehman, and Anil Kommareddy. Surface water extent dynamics from three decades of seasonally continuous Landsat time series at subcontinental scale in a semi-arid region. Remote Sensing of Environment, 178 :142–157, June 2016. ISSN 00344257. doi : 10.1016/j.rse.2016.02.034.
- Victoria Vanthof and Richard Kelly. Water storage estimation in ungauged small reservoirs with the TanDEM-X DEM and multi-source satellite observations. Remote Sensing of Environment, 235 :111437, December 2019. ISSN 00344257. doi : 10.1016/j.rse.2019.111437.

- Kaushlendra Verma and Yashwant B. Katpatal. Groundwater Monitoring Using GRACE and GLDAS Data after Downscaling Within Basaltic Aquifer System. Groundwater, 58(1) :143–151, 2020. ISSN 1745-6584. doi : 10.1111/gwat.12929.
- Bramha Dutt Vishwakarma, Jinwei Zhang, and Nico Sneeuw. Downscaling GRACE total water storage change using partial least squares regression. Scientific Data, 8(1) :95, March 2021. ISSN 2052-4463. doi : 10.1038/s41597-021-00862-6.
- Naresh Krishna Vissa, P. C. Anandh, Mama Manjali Behera, and Sameeksha Mishra. ENSO-induced groundwater changes in India derived from GRACE and GLDAS. Journal of Earth System Science, 128(5) :115, May 2019. ISSN 0973-774X. doi : 10.1007/s12040-019-1148-z.
- Yoshihide Wada, L. P. H. van Beek, Daniel Viviroli, Hans H. Dürr, Rolf Weingartner, and Marc F. P. Bierkens. Global monthly water stress : 2. Water demand and severity of water stress. Water Resources Research, 47(7), 2011. ISSN 1944-7973. doi : 10.1029/2010WR009792.
- Yoshihide Wada, L. P. H. van Beek, and Marc F. P. Bierkens. Nonsustainable groundwater sustaining irrigation : A global assessment. Water Resources Research, 48(6), 2012. ISSN 1944-7973. doi : 10.1029/2011WR010562.
- Wolfgang Wagner, Günter Blöschl, Paolo Pampaloni, Jean-Christophe Calvet, Bizzarro Bizzarri, Jean-Pierre Wigneron, and Yann Kerr. Operational readiness of microwave remote sensing of soil moisture for hydrologic applications. Hydrology Research, 38(1) :1–20, February 2007. ISSN 0029-1277. doi : 10.2166/nh.2007.029.
- John Wahr, Mery Molenaar, and Frank Bryan. Time variability of the Earth’s gravity field : Hydrological and oceanic effects and their possible detection using GRACE. Journal of Geophysical Research : Solid Earth, 103(B12) :30205–30229, 1998. ISSN 2156-2202. doi : 10.1029/98JB02844.
- John Wahr, Sean Swenson, Victor Zlotnicki, and Isabella Velicogna. Time-variable gravity from GRACE : First results : TIME-VARIABLE GRAVITY FROM GRACE. Geophysical Research Letters, 31(11) :n/a–n/a, June 2004. ISSN 00948276. doi : 10.1029/2004GL019779.
- I. A. Walter, R. G. Allen, R. Elliott, M. E. Jensen, D. Itenfisu, B. Mecham, T. A. Howell, R. Snyder, P. Brown, S. Echings, T. Spofford, M. Hattendorf, R. H. Cuenca, J. L. Wright, and D. Martin. ASCE’s Standardized Reference Evapotranspiration Equation. pages 1–11, April 2012. doi : 10.1061/40499(2000)126.
- Lingyue Wang, Meiling Liu, Xiangnan Liu, Ling Wu, Peng Wan, and Chuanyu Wu. Pretrained convolutional neural network for classifying rice-cropping systems based on spatial and spectral trajectories of Sentinel-2 time series. Journal of Applied Remote Sensing, 14(1) :014506, January 2020. ISSN 1931-3195, 1931-3195. doi : 10.1117/1.JRS.14.014506.
- Michael M. Watkins, David N. Wiese, Dah-Ning Yuan, Carmen Boening, and Felix W. Landerer. Improved methods for observing Earth’s time variable mass distribution with GRACE using spherical cap mascons. Journal of Geophysical Research : Solid Earth, 120(4) :2648–2671, 2015. ISSN 2169-9356. doi : 10.1002/2014JB011547.
- Birgit Wessel. TanDEM-X DEM Product Specification. 2016.

- David N. Wiese, Felix W. Landerer, and Michael M. Watkins. Quantifying and reducing leakage errors in the JPL RL05M GRACE mascon solution. Water Resources Research, 52(9) :7490–7502, 2016. ISSN 1944-7973. doi : 10.1002/2016WR019344.
- Thomas Wintgens, Anders Nättorp, Lakshmanan Elango, and Shyam R Asolekar. Natural Water Treatment Systems for Safe and Sustainable Water Supply in the Indian Context. page 20, 2016.
- Pat J.-F. Yeh, S. C. Swenson, J. S. Famiglietti, and M. Rodell. Remote sensing of groundwater storage changes in Illinois using the Gravity Recovery and Climate Experiment (GRACE). Water Resources Research, 42(12), 2006. ISSN 1944-7973. doi : 10.1029/2006WR005374.
- Wenjie Yin, Litang Hu, Menglin Zhang, Jingrui Wang, and Shin-Chan Han. Statistical Downscaling of GRACE-Derived Groundwater Storage Using ET Data in the North China Plain. Journal of Geophysical Research : Atmospheres, 123(11) :5973–5987, 2018. ISSN 2169-8996. doi : 10.1029/2017JD027468.
- Wenjie Yin, Tongqing Li, Wei Zheng, Litang Hu, Shin-Chan Han, Natthachet Tangdamrongsub, Michal Šprlák, and Zhiyong Huang. Improving regional groundwater storage estimates from GRACE and global hydrological models over Tasmania, Australia. Hydrogeology Journal, 28 (5) :1809–1825, August 2020. ISSN 1431-2174, 1435-0157. doi : 10.1007/s10040-020-02157-3.
- Benjamin F. Zaitchik, Matthew Rodell, and Rolf H. Reichle. Assimilation of GRACE Terrestrial Water Storage Data into a Land Surface Model : Results for the Mississippi River Basin. Journal of Hydrometeorology, 9(3) :535–548, June 2008. ISSN 1525-755X. doi : 10.1175/2007JHM951.1.
- Gangqiang Zhang, Wei Zheng, Wenjie Yin, and Weiwei Lei. Improving the Resolution and Accuracy of Groundwater Level Anomalies Using the Machine Learning-Based Fusion Model in the North China Plain. Sensors, 21(1) :46, January 2021a. doi : 10.3390/s21010046.
- Jianxin Zhang, Kai Liu, and Ming Wang. Seasonal and Interannual Variations in China’s Groundwater Based on GRACE Data and Multisource Hydrological Models. Remote Sensing, 12(5) : 845, March 2020. ISSN 2072-4292. doi : 10.3390/rs12050845.
- Jianxin Zhang, Kai Liu, and Ming Wang. Downscaling Groundwater Storage Data in China to a 1-km Resolution Using Machine Learning Methods. Remote Sensing, 13(3) :523, February 2021b. ISSN 2072-4292. doi : 10.3390/rs13030523.
- Wei Zhao and Ainong Li. A Downscaling Method for Improving the Spatial Resolution of AMSR-E Derived Soil Moisture Product Based on MSG-SEVIRI Data. Remote Sensing, 5(12) :6790–6811, December 2013. doi : 10.3390/rs5126790.
- Detang Zhong, Shusen Wang, and Junhua Li. A Self-Calibration Variance-Component Model for Spatial Downscaling of GRACE Observations Using Land Surface Model Outputs. Water Resources Research, 57(1) :e2020WR028944, 2021. ISSN 1944-7973. doi : 10.1029/2020WR028944.
- Jingping Zuo, Jianhua Xu, Yaning Chen, and Weihong Li. Downscaling simulation of groundwater storage in the Tarim River basin in northwest China based on GRACE data. Physics and Chemistry of the Earth, Parts A/B/C, 123 :103042, October 2021. ISSN 14747065. doi : 10.1016/j.pce.2021.103042.