On the use of deep learning for ocean SAR image semantic segmentation

Aurélien Colin

To cite this version:

Aurélien Colin. On the use of deep learning for ocean SAR image semantic segmentation. Machine Learning [cs.LG]. Ecole nationale supérieure Mines-Télécom Atlantique, 2022. English. NNT : 2022IMTA0327 . tel-04030623

DOCTORAT
BRETAGNE
LOIRE / MATHSTIC

IMT Atlantique
Bretagne-Pays de la Loire
École Mines-Télécom

THÈSE DE DOCTORAT DE

L'ÉCOLE NATIONALE SUPÉRIEURE
MINES-TÉLÉCOM ATLANTIQUE BRETAGNE
PAYS-DE-LA-LOIRE - IMT ATLANTIQUE

ÉCOLE DOCTORALE N° 601
Mathématiques et Sciences et Technologies
de l'Information et de la Communication
Spécialité : *Signal, Image, Vision*

Par
Aurélien COLIN

On the Use of Deep Learning for Ocean SAR Image Semantic Segmentation

Thèse présentée et soutenue à IMT Atlantique, à Plouzané, le 12/12/2022
Unité de recherche : Lab-STICC, UMR 6285, IMT Atlantique, Brest
Collecte Localisation Satellite, Brest
Thèse N° : 2022IMTA0327

Rapporteurs avant soutenance :

Bertrand LE SAUX	Senior Scientist, Φ-lab, ESA
Shigenori OTSUKA	Collaborative Associate Professor, Kyoto University & RIKEN Cluster for Pioneering Research

Composition du Jury :

Président :	François ROUSSEAU	Professeur, LaTIM, IMT Atlantique
Examinateurs :	Dominique BOUNIOL	Chargée de Recherche, CNRM, Météo-France
	Ronan FABLET	Professeur & Encadrant, Lab-STICC, IMT Atlantique
	Bertrand LE SAUX	Senior Scientist, Φ-lab, ESA
	Shigenori OTSUKA	RIKEN Cluster for Pioneering Research
	Fabio Michele RANA	Research Fellow, Institute of Atmospheric Pollution, National Research Council of Italy
	Stefano ZECCHETTO	Senior Scientist, Institue of Polar Sciences, National Research Council of Italy
Dir. de thèse :	Samir SAOUDI	Professeur, Lab-STICC, IMT Atlantique

Invité(s) :

Romain HUSSON	Ingénieur de Recherche & Encadrant, Collecte Localisation Satellite
Pierre TANDEO	Maître de conférences & Encadrant, Lab-STICC, IMT Atlantique
Alexis MOUCHE	Laboratoire Spatial et Interfaces Air-Mer, IFREMER

CONTENTS

III Collocation-based learning schemes for Earth Observation 89

INTRODUCTION (ENGLISH VERSION)

1 Context

The ocean is the source of most atmospheric processes. Understanding these processes is key to accurately predicting weather and climate patterns, not only over the ocean, but also on land. Due to the difficulty of measuring meteorological conditions on the ocean directly, spaceborne observations have been utilized for many years to monitor the ocean. Various sensors are used to study the ocean, but this thesis focus on C-Band Synthetic Aperture Radar (C-SAR). C-SAR is an active sensor (i.e. self-illuminated) whose signal is not affected by the clouds, on the contrary, for example, to instruments operating in the visible spectrum or infrared. It also has the advantage of providing a very high spatial resolution and is therefore a particularly important oceanographic instrument. C-Band radars are known to be able to observe various phenomena related to the winds [123, 223], the waves [167], the precipitation [232], or various objects on the surface such as icebergs [38], ships [184], biological slicks [185] -the agglomeration of organic matter forming a linear shape and dampening the waves because of the higher viscosity-, or some species of seaweed [170]. Dedicated systems have been developed to study each of these processes.

Various satellites have boarded C-SAR instruments. ERS-1 and ERS-2 [9], were launched respectively in 1991 and 1995 have been operating until respectively 2000 and 2005. ENVISAT [122] also had a ten-year mission from 2002 to 2012. RADARSAT-1 [151] was in operation from 1995 to 2013. Currently, SAR observations routinely acquired by the Sentinel-1 mission of ESA's Copernicus program [10]. Sentinel-1A was launched in 2014 is still observing the Earth. Sentinel-1B, launched in 2016, experienced a technical anomaly at the end of 2021 and stopped its recording, Sentinel-1C is expected to be launched in 2023. RADARSAT-2 [138] has been acquiring observations since 2007 and RADARSAT Constellation [189] since 2019. Since SAR images have been accumulated for decades, the quantity of available observations is extremely high. This stimulates the research of data-driven algorithms able to automatically identify meteorological or ocean information from the observations. Furthermore, the last ten years saw the rise of neural networks. Even though they had existed for decades [226], their use has been driven by the conjunction of both the increasing capacities of GPUs [23] and the raised availability of large datasets. They are now known for their capacity to achieve state-of-the-art performance in various contexts, especially for image processing.

2 Scientific Issues

This thesis is written at the intersection of the three fields of oceanography, remote sensing and machine learning. It aims at segmenting, i.e. computing a pixel-level information, various meteorological and ocean processes (gathered under the adjective "metocean") observed by a Synthetic Aperture Radar. More formally, the question that were trying to answer are the following:

It is possible to segment metocean processes on Synthetic Aperture Radar observations?

The bibliography shows that some information, such as the wind speed, can be computed from the SAR observations at the pixel-level [70]. On the other hand, some information, such as the metocean processes appearing in the observation, can only be resolved at a scale of several dozen kilometres per pixel [207]. This resolution is too low to accurately delimit some phenomena such as icebergs, fronts or rain cells. Furthermore, a single patch can contain several phenomena since it extends for several dozen kilometres. The convective cells containing rain signature and atmospheric fronts are a common example of such co-occurrence. Being able to distinguish the different metocean processes would therefore be beneficial to their understanding.

It is not entirely accurate to refer to the information as being at a "pixel-level" or a "patch-level". In the case of the wind speed, the current computation is performed at a 1 km/px scale. This resolution is low in regards with the maximum resolution of the sensor (a few metres per pixels). In this manuscript, we refer to "pixel-level" information which has a resolution similar to the input used to compute it. To reduce the noise caused by the radar process (particularly, speckle) and lighten the computational load, this information is usually downscaled to a few hundred meters per pixel.

Is it possible to separate processes appearing on the same pixels?

Similar to the co-occurrence of different metocean processes in the same areas, it is possible to encounter them on the same pixels. Examples of such co-occurrence include the appearance of wind-related features with non-wind features such as sea ice, biological slicks, or rain. Understanding the contribution of each phenomenon to the total reflectivity is important to ensure the accuracy of derived information, such as wind speed, which can be difficult to obtain in cases of heterogeneous patterns [232]. It also allows for a better understanding of the interactions between different elements in the environment. By being able to accurately identify and differentiate processes, it is possible to gain insight into their interactions.

Is segmentation achievable using current datasets?

Obtaining large datasets with annotated data is a time-consuming and costly process, so we are exploring ways to minimize the cost associated with dataset building. There are some methodologies available to learn pixel-level segmentation from image-level information, but their effectiveness for our specific problems is still unknown. Therefore, we are conducting experiments to determine which approach yields the best results. Additionally, we are looking into automated solutions to accelerate the dataset building process and reduce manual labor.

If existing datasets are insufficient to provide accurate segmentation, we are considering using collocation with other sensors (ground-stations, other satellites, buoys, etc.). By combining these data sources, it may be possible to reduce the amount of manual annotation required, while also improving the accuracy of the segmentation. This approach can also be used to increase the diversity of the dataset and to ensure that the segmentation can be applied to a variety of conditions. Additionally, it can be used to gain a deeper understanding of the underlying phenomena being observed by cross-referencing and comparing data from multiple sources.

3 Organization of the manuscript

This manuscript is divided in three parts. In the first, we will describe the three main domains aforementioned. Several satellite-based sensors are presented, putting a particular focus on the SAR imagery. As it constitutes the basis of the model presented later, its physical principles are detailed and examples of SAR-derived information are provided.

Then, we introduce the machine learning framework, beginning by the multi-layer perceptron and continuing with the main components of deep learning models and the algorithm necessary to train them. We also provide examples of current use cases of convolutional neural network. This chapter will serve both as an introduction and as a handbook defining the terms and concepts used through the document.

The second part of the manuscript is dedicated to segmentation problems solved with manually annotated datasets. The first subject, the segmentation of biological slicks, presents the deep convolutional neural networks on a simple binary segmentation problem. It highlights the difficulties of collocating phenomena with secondary sensors, which are themselves often indirect information rather than *in situ* measurements. Specifically, the biological slicks are here estimated via both the segmentation of the Sentinel-1 observations and using the ocean colour from Sentinel-3 measurements. The second subject justifies the necessity of pixel-level annotations by comparing the fully supervised frameworks with weakly supervised ones, where only easy to obtain image-level information are used. The dataset imbalance, a common difficulty of machine learning frameworks, is also introduced. It designs the scarcity of some phenomena relatively to others. The last chapter of this part dedicated to manual annotations addresses the multi-scale features of some metocean processes, taking the example of the convective cells. As observations are acquired at a very high resolution, it is sometimes necessary to use both low- and high-resolutions to understand the metocean situation.

Massive collocations are performed in the third part. First, the retrieval of annotation from the Norwegian Meteorological Institute enables the estimation of sea ice concentration. It highlights the complementarities between the categorical and regression aspects of the segmentation losses. Rain rate estimations are performed in the second chapter using collocations with the ground-based radars of the NEXRAD system. Results are validated *via* a comparison with lightning activity, an indirect marker of rain available from the geostationary satellite GOES-16. In the last chapter, SAR observations, SAR-based wind speed, rain rate and an atmospheric model are used to build a dataset and train a rain-invariant wind speed estimator. Though modelled wind speeds are used to train the model, the performance is validated against *in situ* measurements from buoys.

Each chapter in the two last parts cover different challenges, but share a common framework of dataset building, choice of the model configuration and validation through ancillary data. They highlight the capacity of SAR imagery to observe and detect various metocean phenomena.

INTRODUCTION (VERSION FRANÇAISE)

1 Contexte

L'océan est à l'origine de la plupart des processus atmosphériques. La compréhension de ces processus est essentielle pour prédire avec précision l'évolution des conditions météorologiques et climatiques, non seulement au-dessus des mers, mais également sur la terre ferme. En raison de la difficulté de mesurer directement les conditions météorologiques sur l'océan, les observations spatiales sont utilisées depuis de nombreuses années pour surveiller l'océan. Divers capteurs sont utilisés dans ce but, et cette thèse se concentre sur les radars à synthèse d'ouverture en bande C (C-SAR). Les SAR sont des capteurs actifs, c'est-à-dire auto-illuminés, dont le signal n'est pas affecté par la couverture nuageuse, contrairement à d'autres instruments fonctionnant, par exemple, dans le domaine visible ou l'infrarouge. Ils présentent également l'avantage de fournir une résolution spatiale très fine et sont donc des instruments particulièrement intéressants pour l'océanographie. Les radars en bande C sont capables d'observer une grande variété de phénomènes liés aux vents [123, 223], aux vagues [167], aux précipitations [232] ou à divers objets flottants tels que les icebergs [38], les navires [184], le film biologique [185] - une accumulation de matière organique formant de longs filaments et atténuant les vagues en raison d'une viscosité plus élevée - ou certaines espèces d'algues [170]. Des systèmes dédiés existent pour étudier chacun de ces processus.

Divers satellites servent de plateformes pour des instruments SAR. ERS-1 et ERS-2 [9], lancés respectivement en 1991 et 1995, ont été en activité jusqu'en 2000 et 2005. ENVISAT [122] a également eu une mission de dix ans entre 2002 et 2012. RADARSAT-1 [151] était en opération de 1995 à 2013. Actuellement, les observations SAR sont acquises de manière systématique par la mission Sentinel-1 du programme Copernicus de l'ESA [10]. Sentinel-1A a été lancé en 2014 et est toujours en activité. Sentinel-1B, lancé en 2016, a connu un dysfonctionnement à la fin de 2021 et a arrêté d'acquérir des données. Sentinel-1C devrait être lancé en 2023. RADARSAT-2 [138] acquiert des observations depuis 2007 et RADARSAT Constellation [189] depuis 2019. L'accumulation des images SAR depuis plusieurs décennies a entraîné la disponibilité d'une quantité extrêmement élevée d'observations. L'utilisation de méthodes d'apprentissage, capables d'identifier automatiquement les informations météo-océaniques (c'est-à-dire météorologiques ou océaniques), est ainsi particulièrement adaptée à ces applications. Par ailleurs, les dix dernières années ont vu l'émergence de modèles d'apprentissage profond. Bien que les réseaux de neurones existent depuis des décennies [226], leur utilisation a été favorisée par la conjonction de l'augmentation des capacités de calcul des GPU [23] et d'une disponibilité accrue de grands ensembles de données. Ces modèles sont désormais connus pour leur capacité à atteindre l'état de l'art dans différents contextes et démontrent leur versatilité, notamment sur les problématiques de traitement d'image.

2 Problématiques scientifiques

Cette thèse se situe à l'intersection de trois domaines : l'océanographie, la télédétection et l'apprentissage automatique. Elle vise à étudier des problèmes de segmentation, c'est-à-dire à obtenir une information au niveau des pixels, sur différents processus météo-océaniques observables par le SAR. De manière plus détaillée, les questions auxquelles nous tentons de répondre sont les suivantes :

Est-il possible de segmenter les processus métoceaniques dans les observations de radar à synthèse d'ouverture ?

La littérature indique que certaines informations, telles que la vitesse du vent [70], peuvent être calculées au niveau des pixels. Actuellement, certaines informations, telles que le type de processus météo-océanique apparaissant dans l'observation, ne peuvent être calculées qu'à une échelle de plusieurs dizaines de kilomètres [207]. Cette résolution est bien trop faible pour délimiter avec précision certains phénomènes tels que les icebergs, les fronts (atmosphériques ou océaniques) ou les cellules de pluie. À ces échelles, plusieurs phénomènes peuvent exister dans une même cellule de résolution. Les cellules de convection sont une illustration de co-occurrence de processus métoceaniques car elles contiennent simultanément des changements de régimes de vent et, bien que ce ne soit pas systématique, des signatures de pluie. La capacité à distinguer les différents phénomènes à une échelle spatiale plus précise serait donc bénéfique pour leur compréhension.

Est-il possible de séparer les processus météo-océaniques apparaissant sur les mêmes pixels ?

Tout comme les différents processus météo-océaniques peuvent apparaître dans un voisinage restreint, ils peuvent également affecter les mêmes pixels. Des exemples de cette co-occurrence comprennent l'apparition de caractéristiques liées au vent avec d'autres indépendantes du vent et impactées par la banquise, des surfactants tels que les films biologiques ou encore les précipitations. Comprendre la contribution de chaque phénomène à la réflectivité totale est important pour assurer l'exactitude des informations dérivées du SAR. La vitesse du vent peut ainsi être difficile à obtenir dans des cas d'hétérogénéités fortes du signal réfléchi [232]. Résoudre ces co-occurrences permet également une meilleure compréhension des interactions entre différents paramètres environnementaux puisque l'identification précise des différents processus météo-océaniques en jeu rend possible une meilleure compréhension de leurs interactions.

La segmentation est-elle réalisable avec les bases de données existantes ?

Puisque la construction de larges bases de données avec des données annotées est chronophage et coûteuse, nous cherchons à minimiser le coût associé à leur création. Des méthodologies existent pour entraîner des modèles de segmentation à partir d'informations disponibles à des résolutions sous-résolues, mais leur efficacité sur les problématiques spécifiques du SAR pour l'océanographie reste à déterminer. De plus, l'utilisation de solutions automatisées pour accélérer la mise en place des bases de données est une piste supplémentaire à explorer.

De plus, la colocalisation avec d'autres capteurs (stations au sol, instruments spatiaux, bouées...) peut permettre de pallier à l'imprécision des segmentations des bases de données existantes. La combinaison de ces sources de données pourrait permettre de réduire la quantité d'annotations manuelles requises et d'améliorer la résolution des modèles. Cette approche peut également être utilisée pour augmenter la diversité de la base de données et garantir que la segmentation puisse être appliquée à une diversité supérieure de conditions météo-océaniques. Parallèlement, elle peut être utilisée pour mieux comprendre les phénomènes sous-jacents en recoupant et en comparant les données provenant de plusieurs sources, plusieurs capteurs ayant des caractéristiques différentes.

3 Organisation du manuscrit

Ce manuscrit est divisé en trois parties. Dans la première, nous décrivons les trois domaines d'étude mentionnés précédemment. Plusieurs capteurs à bord de satellites sont présentés, en mettant l'accent sur l'imagerie SAR qui constitue le cœur du domaine d'étude. Nous décrivons les principes physiques de cette dernière avant de présenter différentes informations actuellement dérivées de ce type de capteur. Par la suite, nous introduisons le cadre de l'apprentissage automatique en commençant par le perceptron multicouche et en continuant avec les principaux composants des modèles d'apprentissage profond modernes ainsi que l'algorithme nécessaire pour les entraîner. Nous fournissons également des exemples d'utilisation courante de réseaux de neurones convolutifs, lesquels sont particulièrement adaptés pour le traitement d'informations spatialisées telles que les images. Ce chapitre sert à la fois d'introduction et de guide définissant les termes et les concepts utilisés tout au long du document.

La deuxième partie du manuscrit est consacrée aux problèmes de segmentation résolus avec des bases de données manuellement annotées. Le premier sujet, la segmentation des films biologiques, présente les réseaux de neurones convolutifs profonds sur un problème de segmentation binaire simple. Il met en évidence les difficultés de la colocalisation de phénomènes avec des capteurs secondaires, qui sont eux-mêmes souvent des informations indirectes plutôt que des mesures in situ. Plus précisément, les films biologiques sont ici estimés à la fois par la segmentation des observations Sentinel-1 et en utilisant la couleur de l'eau obtenue par les mesures Sentinel-3. Le deuxième sujet de cette partie justifie la nécessité d'obtenir des annotations au niveau des pixels en comparant des cadres d'étude faiblement supervisés à des systèmes entièrement supervisés dont les informations sont de plus grande qualité mais plus difficiles à obtenir. Le déséquilibre des bases de données, difficulté courante des méthodes d'apprentissage automatique et ici induit par des différences de fréquences entre les phénomènes météo-océaniques, est également introduit. Le dernier chapitre de cette partie, consacrée aux annotations manuelles, aborde les caractéristiques multi-échelles de certains processus, les illustrant par l'exemple des cellules convectives. En effet, utiliser simultanément une grande résolution spatiale et une large surface d'observation nécessite des modèles spécifiques qui ne sont pas nécessaires si l'une ou l'autre de ces caractéristiques est superflue.

La troisième partie du manuscrit aborde les colocalisations massives. Dans un premier temps, l'estimation de la concentration de glace de mer est effectuée en tirant parti des annotations massives mises à disposition par l'institut météorologique norvégien (MetNo). Cette étude met en évidence les complémentarités entre les aspects de catégorisation et de régression des fonctions de coût des méthodes de segmentation. Des estimations du taux de précipitation sont réalisées dans le deuxième chapitre en utilisant des colocalisations avec les radars au sol du système NEXRAD, qui disposent d'une portée suffisante pour observer les zones côtières des États-Unis. Les résultats sont validés par une comparaison avec l'activité électrique, marqueur indirect de la pluie disponible à partir du satellite géostationnaire GOES-16. Dans le dernier chapitre, une base de données contenant des observations SAR, la vitesse du vent, la segmentation des régimes de précipitation et un modèle atmosphérique est construite pour entraîner un estimateur de vitesse de vent invariant aux signatures induites par la pluie. Bien que le modèle résultant soit entraîné sur des vitesses de vent issues d'un modèle atmosphérique (donc ayant une résolution spatiale imprécise), ses performances sont validées par des mesures in situ provenant de bouées.

Si chaque chapitre de ces deux dernières parties aborde des défis différents, ils partagent cependant un cadre commun reposant sur la construction d'une base de données, de choix de la configuration du modèle et de validation à travers l'utilisation de données auxiliaires. Ils mettent ainsi en évidence la capacité de l'imagerie SAR à observer divers phénomènes météorologiques et océaniques.

Part I

Background

As this thesis is located at the junction between the fields of Remote Sensing, Oceanography and Machine Learning, it is necessary to present these subjects and the links that bind them together. This first section will therefore be dedicated to an overview of remote sensing and its use in an oceanographic context. We will particularly focus on the Synthetic Aperture Radar in the second section. The last section of this chapter will introduce the machine learning framework, explain the mechanisms of neural networks, and provide examples of current use cases of this technology.

1.1 Oceanography and Remote Sensing

1.1.1 What is Earth Observation ?

Earth Observation refers to the various sensors used to study the Earth system, so as to obtain information on meteorological information (e.g. temperature, wind speed), the ocean (e.g. ocean currents or waves), ecosystems (e.g. chlorophyll concentration, canopy height), and pollution (e.g. oil slicks, fine particles). They encompass mainly two kinds of instruments: in situ and remote sensors. The former defines any kind of sensor providing data in their immediate environment, such as thermometers, anemometers, seismographs. For the ocean, these sensors can be built on buoys (either drifting or anchored) or on ships. As an example, the National Data Buoy Center (NDBC) network contains 112 active buoys, mainly in the United States coastal areas. They provide precious information on wind speed, as in situ measurements are both closest to the real wind speed and acquired with a high temporal resolution. However, the drawback is that they only provide punctual data.

On the other hand, remote sensing instruments are able to acquire information in their vicinity. One example of this family of sensor is the ground-based weather radar network Next Generation Weather Radar (NEXRAD). This instrument, more detailed in Chapter 6, is a Doppler radar measuring the reflectivity of the air volume. At the lowest inclination (5°), its measurement can be used up to three hundred kilometers away (with the altitude and the resolution of the observation varying with the distance).

1.1.2 Satellite-based observation

One specificity of ground-based sensors is their immobility: they observe a fixed area. This particularity is useful for providing continuous local observations and making the collocations easier, but it makes it impossible to provide global observations without an extensive sensor network. In particular, it restricts the diversity of the observed scenes as the topography and the bathymetry are often important components of the metocean processes. Satellite-based sensors solve this issue by deploying their instruments in orbit. Those in geostationary orbit (∼36 000 km), such as GOES-16 (used in Chapter 4 and 6) and GOES-17 that host the Geostationnary Lightning Mapper (GLM) [61]. They observe almost a full hemisphere continuously at the expense of lower resolution. In low orbit, the platform is closer to the area it observes, meaning both a higher resolution and a better revisit time (because of an increased orbital speed). GPM-core, which hosts the Dual-frequency Precipitation Radar (GPM-DPR) [168], orbits at 407 km. It provides swaths of 245 km width with a spatial resolution of 5 km/px and a revisit time of 2.5 days.

It is also important to differentiate between active and passive sensors. Active sensors have an emitter that sends the signal to be reflected. On the other hand, passive sensors need the observed area to be illuminated by another source. In the case of optical imagery, such as Sentinel-2's sensors, the illumination is provided by the Sun. The observation is therefore time-dependent. Besides, the different position between the source and the receiver can induces shadows. One example is the dark area observed behind a cloud on optical imagery. Some passive sensors, however, such as thermal infrared, measure the signal emitted by the target itself, and are therefore less time-dependent.

Finally, we can differentiate between nadir and off-nadir instruments. Nadir instruments observe directly at the vertical, such as GPM-DPR. When the target is observed with a non-null angle with the vertical (meaning a non-null incidence angle), the observation is said to be off-nadir. Sentinel-1's radars are examples of off-nadir instruments.

An overview of the different sensors used in this document is present in Table 1.1. The core of this document is based on the Sentinel-1 missions, which characteristics are indicated in section 1.2.2. However, before detailing its specifics, we will first present the SAR process as to familiarize the reader with this particular imagery process.

Instrument	**Elevation**	**Swath width**	**Wavelength**	**Resolution**
Sentinel-1 C-SAR	693 km	250 km (IW) ~20 km (WV)	5.5 cm	50 m (IW GRDH) ~5 m (WV SLC)
Sentinel-3 OLCI	814 km	1270 km	400 to 1020 nm	300 m
GPM-DPR	407 km	245 km	2.2 cm (Ku) 0.84 cm (Ka)	5 km
GOES-16/17 GLM	35790 km	full-disk	777.4 nm	8 km (nadir) to 14 km
GOES-16/17 ABI's band 14	35790 km	full-disk	11.2 µm	2 km

Table 1.1: Main characteristics of the satellites used in this document.

1.2 Synthetic Aperture Radar - Physical Principles & use for Ocean Studies

1.2.1 Physical Principles

To understand SAR observations, it is necessary to understand the physical principle behind the SAR instruments. The goal here is not to delve deep into the physic (especially since books already cover this subject [158, 199, 81]), but to provide insights into the relationship between the radar observation of the ocean processes. First, we will explain the SAR principles and the different processes that increase, or decrease, the received signal. In a second part, we will clarify the terminology used to describe SAR products. Lastly, existing information derived from SAR observations is presented.

Radar physic and implication on the ocean

Microwaves sensors, such as Sentinel-1's C-band radars, are active sensors. They send a pulse towards the surface and measure the signal reflected in their direction. The reflected signal can be theoretically computed as in Eq. 1.1 [158].

$$P_{out} = P_{in} \frac{(G\lambda)^2 \sigma}{(4\pi)^4 R^4} \tag{Eq. 1.1}$$

- P_{in} is the signal received by the target;
- P_{out} is the signal reflected to the sensor;
- G the antenna gain;
- λ the wavelength;

- R the distance to the target;
- σ the scattering cross-section.

Eq. 1.1 does not take into account the attenuation of the signal during the propagation. The signal of C-band radar is not attenuated by the gas in the atmosphere [78]. However, some scatterers -objects scattering the radar signal, for example the hydrometeors- can be present in the air column and disturb the signal during the observation [3]. This is especially when the rain rate is very high, leading to the appearance of dark areas on the observations (since the radar signal is absorbed) [35].

However, in most cases, the primary factor of the radar equation is the scattering cross-section σ. Then, we will provide an example of SAR observation and analyse the different ocean and atmospheric processes appearing on the scene.

Surface Scattering

The scattering cross-section is impacted by both surface and volume scattering. Surface scattering happens on the interface of two homogeneous areas, such as the atmosphere and the ocean. It can be divided into specular reflection (Fig. 1.1.a, b and d), which follows the Snell's law, and the Bragg scattering (Fig. 1.1.c). The latter happens when the surface is irregular on a scale similar to that of the incident wavelength. The signal will therefore be reflected in all directions.

The Snell's law is recalled in Eq. 1.2:

$$\begin{aligned} \theta_1 &= -\theta_2 \\ \frac{\sin(\theta_1)}{\sin(\theta_3)} &= \frac{n_1}{n_2} \end{aligned} \tag{Eq. 1.2}$$

Where:

- n_1 is the refractive index of the medium 1;
- n_2 is the refractive index of the medium 2;
- θ_1 is the incident angle;
- θ_2 is the reflection angle;

- θ_3 is the refraction angle;

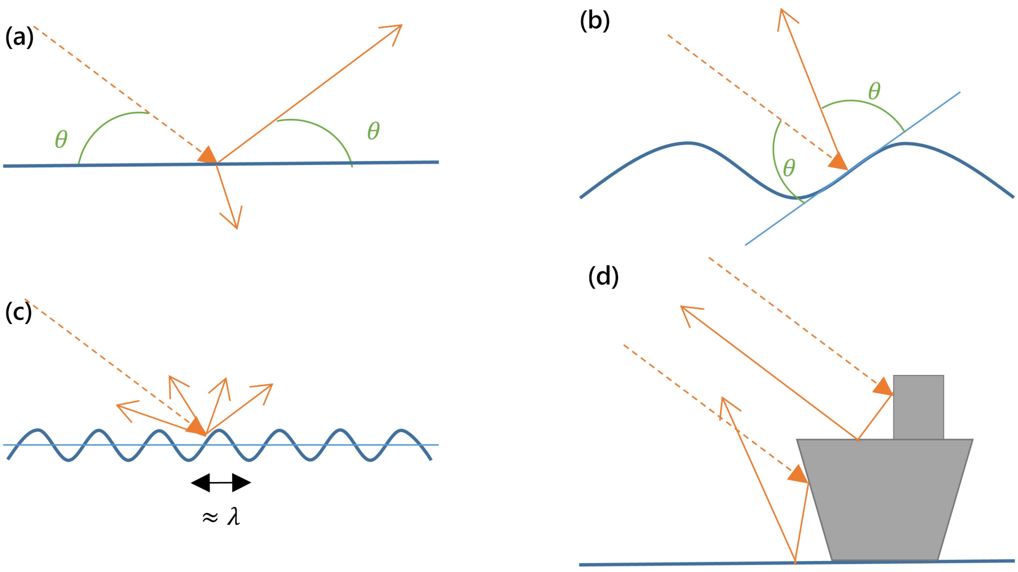


Fig. 1.1: Examples of surface scattering. (a): reflection of the signal on a smooth surface. (b): reflection on a wavy surface, smooth in relation to the emitted wavelength. (c): reflection on a surface rough in relation to the emitted wavelength λ. (d): Multiple reflections caused by complex objects in the observed area. For simplicity, refraction is only indicated in (a).

In the context of oceanography and C-band radars, low wind areas appear as dark areas. As there is no perturbation on the sea surface in this case, most of the signal is in the specular component. On the contrary, when the wind speed increases, the sea surface roughness, along with backscattering, increases too [70].

The specular/Bragg scattering behaviour also explains the different appearance depending on the viscosity of the surface layer. Oil surfactants, for example, have a higher viscosity and dampen the small surface waves. They are therefore observed as dark slicks. It can be noticed that the viscosity of the water depends on multiple factors such as the salinity, the temperature or the pressure [156].

Icebergs, because of their irregular shape, generate multiple reflections in every direction, which explains why they appear as bright targets.

As depicted in Fig. 1.1.a, some of the radar signal is refracted in the observed surface. As per the Fresnel equations, the amount of refracted signal decreases as the dielectric permittivity of the target increases. This is especially true in the case of sea ice: upon formation, first-year sea ice as a similar salinity as the ocean waters, and therefore a high dielectric permittivity. On the contrary, multi-year sea ice has a low salinity and absorb a larger part of the signal.

Volume Scattering

Volume scattering is the dispersion of the emitted signal when passing a heterogeneous medium. It is a well-known phenomenon in the context of meteorological observations as it is used to estimate the distribution of raindrops with weather radars [129], especially in the S-band. It can also be caused by vegetation [85], snow [158], or sea ice [81]. The latter is typically characterized by the presence of gas bubbles embedded in the ice and produces a high backscattering. Volume scattering in C-band also happens in the air column, especially in the melting layer -i.e. the atmosphere layer where snow and ice melt into droplets- which can cause misinterpretation of rainfall at the surface. Volume scattering is depicted in Fig. 1.2.

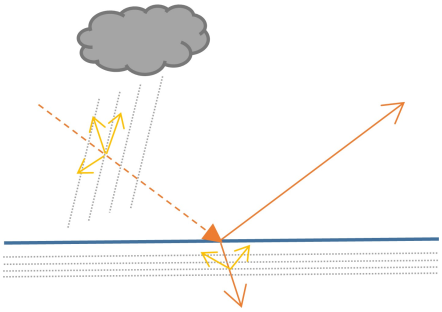


Fig. 1.2: Volume scattering as the signal is transmitted in heterogeneous media.

Sigma nought

Knowing the emitted and received signal, it is possible to compute σ from Eq. 1.1. However, this variable is affected not only by the geophysical processes on the observed surface, but by the radar polarization and the incidence angle [108]. Therefore, various calibrations are performed to obtain a Normalized Radar Cross Section (NRCS), also called σ_0.

Real and Synthetic Aperture Geometry

The resolution of a real-aperture radar is given by two different processes. In terms of range, radars work by sending several pulses. Therefore, the range resolution depends on their pulse (spatial) period, which is the product of c, the speed of light, and τ, the pulse duration. Since the range resolution is projected on the range/azimuth plan with an incidence θ, it is computed as in Eq. 1.3.

$$r_{range} = \frac{c\tau}{2\sin\theta} \tag{Eq. 1.3}$$

On the other hand, the azimuth resolution depends on the horizontal antenna angular resolution β_{real} and the distance R to the target. The latter is related to the incidence angle while the former is limited by the diffraction of the optical system and therefore is bounded by $\frac{\lambda}{L}$. The azimuth resolution is given in Eq. 1.4. The geometry of the Real Aperture Radar is depicted in Fig. 1.3.

$$r_{azimuth,real} = \beta_{real} R = \frac{\lambda h}{L\cos\theta} \tag{Eq. 1.4}$$

However, two issues arise with real aperture radars. The first one is related to the aperture size L. To achieve a resolution of 5 meters at an altitude of 693 km (the elevation of the Sentinel-1 satellites), an antenna with a wavelength of 5 cm would need to be 6 km long (Sentinel-1's antenna is 12.3 m long). The second issue is that azimuth resolution decreases with the incidence angle, while range resolution increases.

As the satellite moves in its orbit, every position in the measurement can be observed during $t = \frac{r_{azimuth,real}}{v}$, where v is the speed of the satellite. It is possible to coherently sum all the measurements of each position to simulate an antenna array with aperture $r_{azimuth}$. Therefore, the resolution of the synthetic aperture is indicated in Eq. 1.5.

$$\begin{aligned} r_{azimuth,SAR} &= \beta_{SAR} R \\ &= \frac{\lambda}{2 r_{azimuth,real}} \cdot R \qquad [199] \\ &= \frac{\lambda}{2\frac{\lambda}{L}R} \cdot R \\ &= \frac{L}{2} \end{aligned} \tag{Eq. 1.5}$$

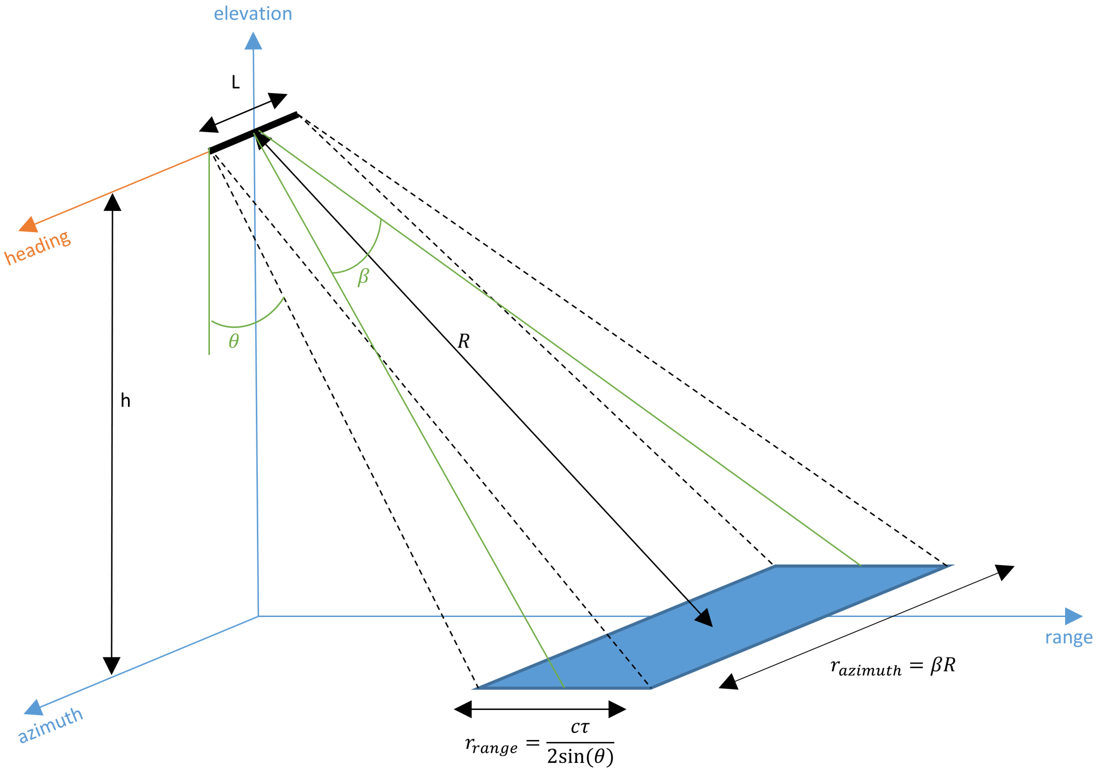


Fig. 1.3: Geometry of the Real Aperture Radar. h is the satellite elevation, L its aperture, R the distance to the target, β_{real} the angular azimuth resolution, θ an incidence angle, c the speed of light and τ the pulse duration.

Under this formulation, the azimuth resolution is independent of the incidence angle and inversely proportional to the length of the antenna. However, it assumes that the observed scene does not change during the integration time t, which is not strictly true in the context of earth observation, especially in oceanography.

Observation Example

Fig. 1.4 is an example of a SAR observation. In this image, where σ_0 has been thresholded to highlight ocean processes rather than land, a large dark area can be observed to the east of Corsica. It corresponds to a low-wind area. At the borders of this area, thin dark lines can be seen spreading. These are biological slicks, oil-like surfactants produced by plankton. They often reveal upwelling, as cold and nutrient-rich waters replace the warm coastal waters under the action of either wind or coastal currents. They are visible because the surfactant increases the viscosity of the water and decreases the Bragg scattering. If the wind speed is too low, they can be difficult to separate from the background because of the smoothness of the sea surface. On the other hand, if the wind speed is high, the slicks dissipate quickly without reaching a high enough concentration to be visible.

Some ships are visible, especially to the east of Cap Corse. Their high dielectric permittivity, their irregular shapes allowing for multiple reflections (and especially reflections in the direction of the radar), and their edges increasing the scattering, make them appear as bright dots even though the resolution (in this case, 100 m/px) is lower than their size.

In the direction of the Strait of Bonifacio, the wind speed increases because of the Venturi effect [55]. Gravity waves can be observed in this area because of they modulate the wind speed [200]. Other orographic processes, in particular lee waves can be seen behind Monte Argentario and the Giglio Island. They are caused by variations in elevation. The wind direction can be estimated to be orthogonal to these gravity waves. On the other hands, on the south of this area, wind streaks are caused by roll vortices in the atmospheric boundary layer and are parallel to the wind direction [231].

1.2.2 SAR products from the Copernicus project

Copernicus is an Earth Observation program coordinating several remote sensing missions [32]. In particular, it encompasses the different Sentinel satellites of which the Sentinel-1 constellation is a part.

The Sentinel-1 constellation is composed of two satellites (Sentinel-1A and Sentinel-1B) launched respectively in April 2014 and April 2016. They are expected to be joined by Sentinel-1C at the beginning of 2023 and by subsequent launches to ensure the continuity of the mission. Each of these satellites follows a heliosynchronous dawn-dusk orbit at 693 km in altitude and carries a C-band synthetic-aperture radar. The orbital period is around 100 minutes, with a revisit time of 12 days (175 orbits). Unfortunately, Sentinel-1B has been out of commission in December 2021 following a technical anomaly.

Product Levels

Several SAR products are created from the observations acquired by the Sentinel satellites. Level-0 products are the raw data acquired by the satellite. Level-1 products are processed products, after calibration and Doppler centroid estimation. Level-1 products can be either Ground Range Detected (GRD), using multi-looks to reduce the impact of speckle (a multiplicative noise) at the expanse of a smaller resolution. Single Look Complex (SLC) preserves the full resolution and the phase information, but are more affected by the

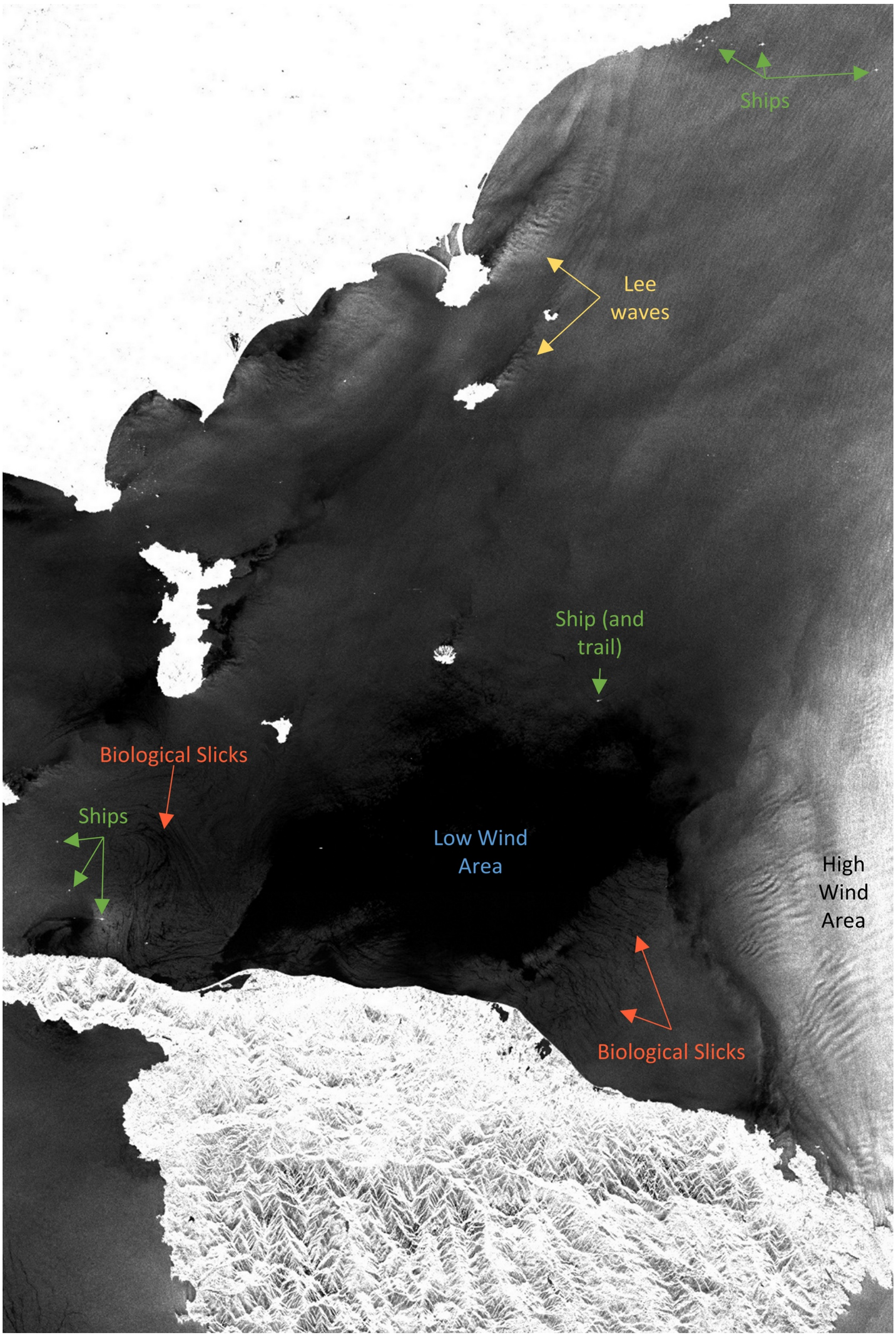


Fig. 1.4: SAR observation from the 07th of June 2021 at 17:14:15.

speckle. Level-2 products are geophysical products derived from Level-1. They contain various information than can be obtained from the SAR (e.g. the wind speed estimate from a geophysical model function), collocation (e.g. the wind direction from ECMWF's atmospheric model) or metadata about the acquisition (e.g. the satellite heading).

Observation Modes

Sentinel-1 products also differ by their observation modes. Four different observation modes exist:

- Interferometric Wideswath (IW) is the default observation mode over coastal areas. It produces swaths of 250 km in range and several hundred kilometers in azimuth. It is composed of three subswaths whose incidence angle ranges are respectively [29.16°, 34.89], [34.77°, 40.15°] and [40.04°, 44.28°]. In Level-1 GRD product, IW exists either at 10 m/px (GRDH, for high-resolution) or at 40 m/px (GRDM, for medium resolution). GRDM is usually a merge of multiple consecutive GRDH to produce a track of several thousand kilometers. The Equivalent Number of Looks (ENL), which can be used to estimate the impact of the speckle, is respectively 4.4 and 81.8.

- Extra-Wide Swath (EW) is the default observation mode over polar areas. It produces swaths 410 km in range. It is composed of five subswaths which incidence angle ranges are respectively [18.22°, 27.57°], [27.38°, 33.42°], [32.65°, 38.05°], [37.84°, 42.53°] and [42.08°, 45.16°]. In Level-1 GRD products, EW exists in GRDH at 25 m/px or in GRDM at 40 m/px. The ENL is 2.7 and 9.6, respectively.

- Wave Mode (WV) is the default observation mode over the open ocean. It produces imagettes of 20x20 km. Wave Mode images are observed either in [21.68°, 23.22°] (WV1) or in [34.88°, 36.13°] (WV2). Wave Mode only exists in SLC at a resolution of 1.7x4.1 m/px (WV1) and 2.7x4.1 m/px (WV2).

- Stripmap is not routinely acquired. We omit this mode for simplicity.

We note that some observation modes have overlapping incidence angles, which indicates that a method trained on a particular mode could be extended to another. As an example, the third and fourth subswaths of EW are contained in IW's incidence angle range, so are the WV2. However, such extension is a study field in itself as other characteristics, e.g. the ENL (and therefore the noise) are different.

Polarization

The Sentinel-1 product also differs by its polarization channels. In polar areas, EW and IW have a primary horizontal polarization. They are therefore denoted as either SH (containing one HH channel) or DH (containing both HH and HV channels). At other latitudes, EW and IW are observed with a vertical polarization. They are denoted as SV (VV channel) or DV (VV and VH channels). WV is always observed in SV. As a side note, it implies that methods relying on the cross-polarization cannot be used on WV imagettes.

The geographical distributions of each observation mode, and each polarization, are presented in Fig. 1.5.

1.2.3 Example of SAR-Derived Information

In this section, we will present some information derived from SAR observations.

SAR has been known for decades to be suitable for the problem of oil slick detection, a form of pollution often left by ships cleaning their tanks in open sea. Observations as early as ERS-1, launched in 1991, were used to build oil detection algorithms [12]. This is especially true in the case of Sentinel-1, as the incidence angle that maximizes the visibility of the slicks (between 20° and 45° [57]) is similar to that of the two satellites. [197] found better categorization performance with C-band instruments, rather than L- and X-band. The oil slick detection continues to be refined up to now [6].

SAR is also used to detect ships. They appear distinctly on the acquisitions, due to both their angular shapes and the change in the dielectric permittivity of the observed surface. Furthermore, the use of the Vessel Monitoring System (VMS) and the Automatic Identification System (AIS) allows for the easy creation of datasets to evaluate detection methods [121]. Deep learning architectures such as RCNN [109] or YOLOv2 [17] are also being used to train ship detectors, with dedicated datasets regularly published [213, 184, 212]. Ship detection shares similarities with iceberg detection, which can also be performed on SAR observations [38]. The discrimination between the two is also researched [76].

The study of the cryosphere, in particular, sea ice. The estimation of the sea ice concentration is the subject of the dedicated section 5, but the elaboration of sea ice charts primarily use SAR observations [82], along with other information such as visible and infrared imagers, microwave radiometers, scatterometers, or altimeters [82].

Another main field using SAR observations is the estimation of wind characteristics.

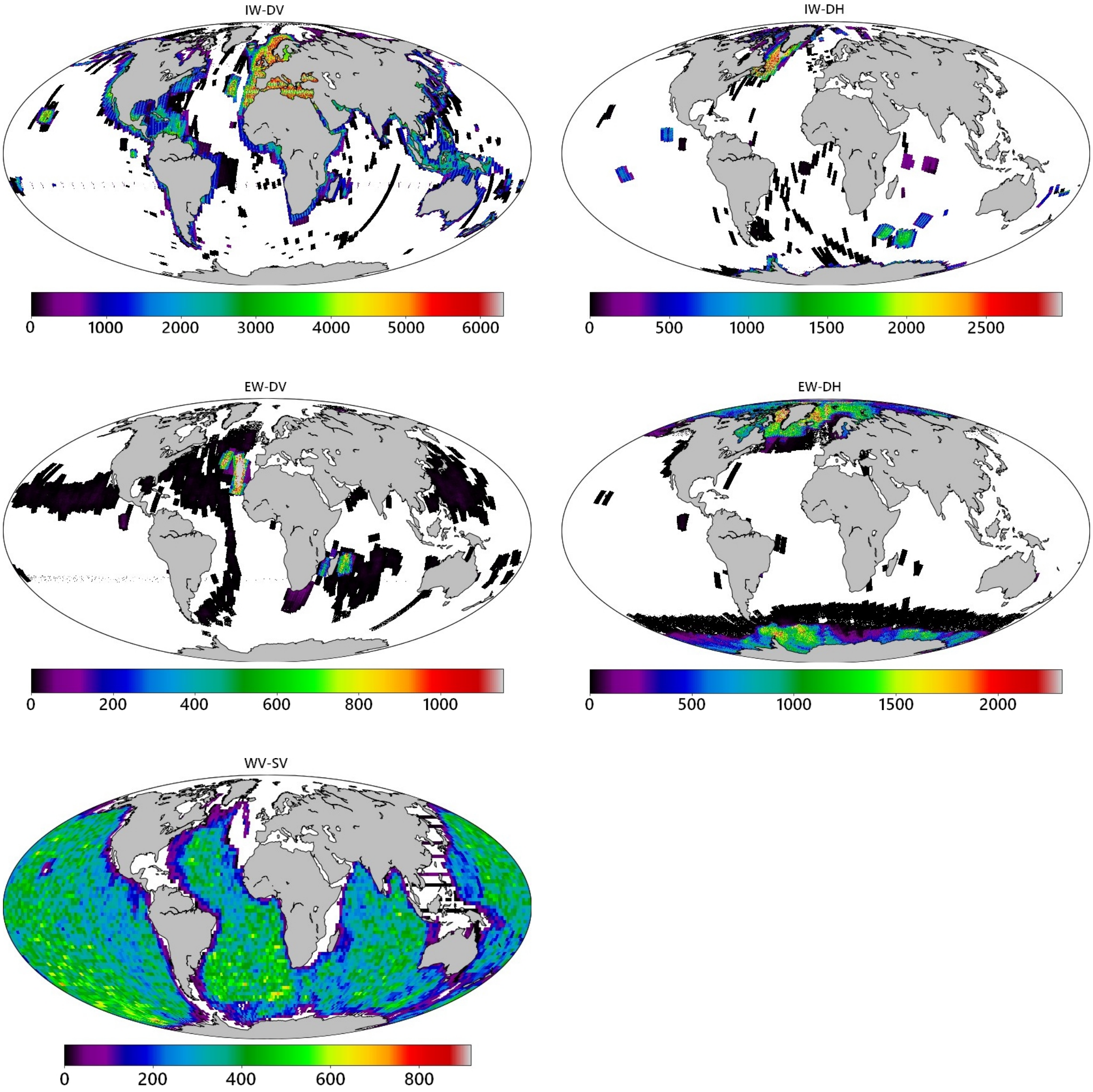


Fig. 1.5: Geographical distribution of Sentinel-1 observations between 2019-01-01 and 2021-31-12. From top to bottom IW-DV, IW-DH, EW-DV, EW-DH and WV-SV, based on the L2-OCN products.

Notably, a range of C-band geophysical models have been developed to estimate the wind speed from SAR observations (CMOD3 [119], CMOD4 [179], CMOD5 [69], CMOD5.N [70], CMOD6 [47], CMOD7 [180], C_SARMOD2 [123]). The geophysical models rely on a prior wind direction to estimate the speed. Thus, efforts have been made to also estimate this prior from the SAR observation and it has been proved to be possible in certain meteorological conditions, especially when atmospheric roll vortices are present [44]. Vortices rolls can be observed if the wind speed is higher than 2.5 m/s [206, 231], but convection and rain signatures can hinder the direction retrieval. Deep learning methods are also being explored to estimate the wind direction despite the lack of wind streaks [223].

Surface waves have been a prominent subject of SAR studies as several of their characteristics, such as the wave spectrum can be estimated from SAR observations [4, 220]. Their height [167] and their period [181] can be estimated even in complex situations such as marginal sea ice areas [8] or cyclones [169]. Internal waves -often caused by tidal currents [2] on contrary to surface winds generated by the wind- are also studied via SAR observations [1]. As they propagate in the whole water column rather than the surface, they impact different processes such as the turbulence exchange [94] and therefore the development of biological slicks [172].

1.3 Machine Learning

1.3.1 Introduction to Neural Network and Classic Structure of Convolutional Neural networks

Machine Learning has been rising since the end of the 20th century, particularly since 2012, when AlexNet [95] won the ImageNet LSVRC-2012 challenge [164] outperforming other challengers. In the case of neural networks, it mostly relies on three major components:

- a dataset of observations ;
- a model to compute information from the observations ;
- a loss function to estimate the quality of the model.

The specifics of these three points differs depending on the learning scheme used. We can distinguish four learning schemes depending on the dataset characteristics:

- Fully-supervised learning relies on the association of a groundtruth to each observation. It is particularly suited for estimation of a continuous value (a regression) or of an integer value (a class).
- Weakly-supervised learning also relies on groundtruths, but assume that their quality is low. Examples are wrong or imprecise labels.

- Semi-supervised learning relies on both annotated and unannotated observations. It aims at addressing the difficulties to build large annotated datasets. Sometimes, semi-supervised is include as a weakly-supervised framework.
- Unsupervised learning does not use associated groundtruths. Examples of applications include clustering (grouping of similar data points into distinct categories), anomaly detection, and dimensionality reduction. Self-supervised learning, under which the truth is directly created from the observation, is a type unsupervised learning and can be applied, for example, for noise reduction. Unconditional Generative Adversarial Networks are another example of a popular unsupervised method.

In the following, we will mainly spoke about fully-supervised methods.

When using a neural network, the idea is to optimize the parameters of the model to minimize a cost function, also called a "loss." This can be done by following a backpropagation algorithm. The neural network is composed of several layers that compute a derivative function. For example, a fully connected layer, also called a "dense layer," computes the function described in Eq. 1.6.

$$\hat{y} = L(x_1, ..., x_{n_1}) = f\Big(b + \sum_{i=1}^{n_1} x_i \cdot a_i\Big) \tag{Eq. 1.6}$$

where y is the output of the layer, a vector of n_2 elements, x its input, a vector of n_1 elements, a a weight matrix of shape (n_1, n_2) and b a vector of n_2 elements, also called "bias". a and b are the weights of the layer and adjusted during training. The f function is called the "activation function". It is often chosen to be non-linear because of the universal approximation theorems [34, 124]. Three common activation functions are the sigmoid (Eq. 1.7) and ReLU (Eq. 1.8), and the linear function (Eq. 1.9).

$$sigmoid(x) = \frac{1}{1 + e^{-x}} \tag{Eq. 1.7}$$

$$relu(x) = max(0, x) \tag{Eq. 1.8}$$

$$linear(x) = x \tag{Eq. 1.9}$$

For the output layer, the rule of thumb is to use the sigmoid function on classification or segmentation problems when the output needs to be in the [0, 1] interval. On regression problems, the linear function allows for results in the $]-\infty, +\infty[$ interval. However, if the

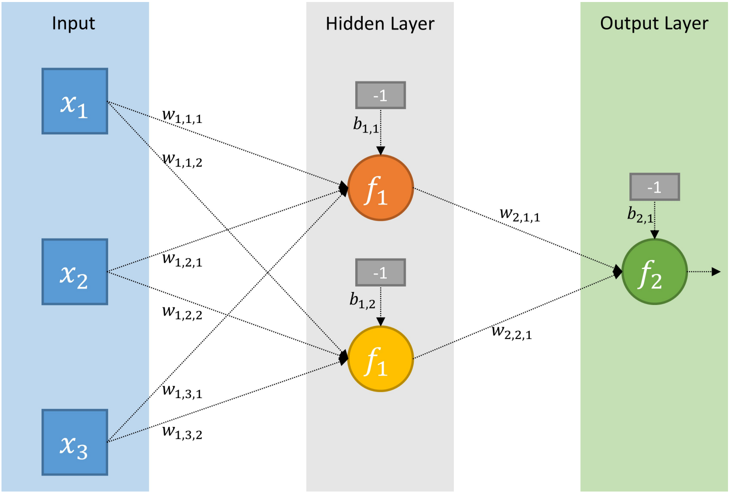


Fig. 1.6: Architecture of a Multi Layer Perceptron with three inputs, one hidden layer of two neurons, and one output.

interval is $[0, +\infty[$, as with many geophysical parameters such as wind speed, the ReLU function can also be used.

For the other layers, usually called hidden layers, the ReLU function is preferred as the computation of its derivative is quick and not subject to the vanishing gradient problem [150]. Vanishing gradient is an issue that occurs when several successive layers decrease the gradient close to zero, preventing the model from learning anything. It happens with the sigmoid function because its derivative is always lower than 1.

A simple neural network: the Multi Layer Perceptron

The architecture of a simple Multi Layer Perceptron is presented in Fig. 1.6. In this example, it is composed of one input layer of dimension 3, one hidden (or intermediate) layer with two neurons, and an output layer with a single neuron. Each neuron -or perceptron- computes a weighted sum of its inputs, which include a bias, and applies an activation function. The neurons of the same layer share the same activation function.

The orange and yellow neurons compute respectively the y_1 and y_2 values as described in Eq. 1.10.

$$\begin{aligned} y_1 &= f_1(x_1 \cdot w_{1,1,1} + x_2 \cdot w_{1,2,1} + x_3 \cdot w_{1,3,1} - b_{1,1}) \\ y_2 &= f_1(x_1 \cdot w_{1,1,2} + x_2 \cdot w_{1,2,2} + x_3 \cdot w_{1,3,2} - b_{1,2}) \end{aligned} \tag{Eq. 1.10}$$

Then, the output layer computes z as indicated in Eq. 1.11.

$$z = f_2(y_1 \cdot w_{2,1,1} + y_2 \cdot w_{2,2,1} - b_{2,1}) \qquad \text{(Eq. 1.11)}$$

Denoting i the index of the layer, k the index of the input of the layer i, and l the index of the neuron in the layer i, the training consists in optimizing the bias $b_{i,l}$ and the parameters $w_{i,k,l}$ to minimize the error between the output z and the ground truth.

Back propagation algorithm

The neural network is the composition of several layers, so that the output of the intermediate layer i would be computed recursively as indicated in Eq. 1.12.

$$\mathcal{M}_i(x) = L_i \circ \mathcal{M}_{i-1}(x) \qquad \text{(Eq. 1.12)}$$

In this equation, $\mathcal{M}_i(x)$ is the output, at the layer i, computed with an input x, and L_i is the function computed by the layer i itself.

The goal of the back propagation algorithm is to compute the gradient of the loss function relatively to the last layer, and then propagate it to the previous layers. Computing the loss of the last layer is easy. Denoting the error $e = \mathcal{L}(y, \mathcal{M}(x))$ (i.e. the result of the loss function), the weight gradient is computed in Eq. 1.13.

$$\frac{\delta e}{\delta w} = \frac{\delta e}{\delta \mathcal{M}(x)} \frac{\delta \mathcal{M}(x)}{w} \qquad \text{(Eq. 1.13)}$$

For an arbitrary layer, i, J_i denotes the set of layers that use the output of the layer i as an input. Assuming $\frac{\delta e}{\delta \mathcal{M}_j(x)}$ is known for any layer j in J, we can compute the gradient for the weights w_i of the layers i as indicated in Eq. 1.14.

$$\begin{aligned} \frac{\delta e}{\delta w_i} &= \frac{\delta e}{\delta \mathcal{M}_i(x)} \frac{\delta \mathcal{M}_i(x)}{\delta w} \\ &= \frac{\delta \mathcal{M}_i(x)}{\delta w} \sum_{j \in J} \frac{\delta e}{\delta \mathcal{M}_j(x)} \frac{\delta \mathcal{M}_j(x)}{\delta \mathcal{M}_i(x)} \end{aligned} \qquad \text{(Eq. 1.14)}$$

By recurring from the end of the network, it is therefore possible to compute the loss gradient for each layer, and for each weight.

Delta rule/weight update/optimizer

The last part of the back propagation algorithm is to apply to the delta rule (also called optimizer) to adjust the weights using gradient descent, as expressed in Eq. 1.15.

$$w_{i,t+1} = w_{i,t} - \lambda \frac{\delta e}{\delta w_{i,t}} \tag{Eq. 1.15}$$

Here, t is the index of the weight update. It expresses that the weights are iteratively optimized until convergence of the error e.

In this manuscript, we will particularly use the RMSprop and the Adam (ADAptative Momentum) [91] optimizers. Both include a normalization by the square of the gradient to avoid it turning too big or too small. Adam use both the first and the second moment, the former being used to accelerate the training while the slope is in the same direction. RMSprop only use the second moment. In both cases, the moments are accumulated inline with an exponential decay, meaning they are of the form $x_t = a \cdot x_{t_{-1}} + f_t$. The update rule corresponding to RMSprop is indicated in Eq. 1.16, the update rule for Adam is Eq. 1.17.

$$\begin{aligned} v_{i,t} &= 0.9 v_{i,t-1} + 0.1 \Big(\frac{\delta e}{\delta w_{i,t}} \Big)^2 \\ w_{i,t+1} &= w_{i,t} - \frac{\lambda}{\epsilon + \sqrt{v_{i,t}}} \frac{\delta e}{\delta w_{i,t}} \end{aligned} \tag{Eq. 1.16}$$

$$\begin{aligned} \hat{m}_{i,t} &= 0.9 m_{i,t_1} + 0.1 \frac{\delta e}{\delta w_{i,t}} \\ \hat{v}_{i,t} &= 0.999 v_{i,t-1} + 0.001 \Big(\frac{\delta e}{\delta w_{i,t}} \Big)^2 \\ w_{i,t+1} &= w_{i,t} - \frac{\lambda}{\epsilon + \sqrt{1000 \hat{v}_{i,t}}} 10 \hat{m}_{i,t} \end{aligned} \tag{Eq. 1.17}$$

Classic Layers of a CNN

With the dense layer, the loss function, the back propagation, the delta rule, and the dataset, it is possible to design basic neural networks. More advanced networks use additional layers such as convolution layers, pooling or concatenation.

Convolution Layer

Dense layers are often inadequate as they discard spatial information (the proximity between two pixels is not taken into account in Eq. 1.6) and the number of weights -or parameters- increases as the product of both the input and the output length. Convolutional neural networks are neural networks that rely on convolution layers. On the contrary, with dense layers, where the output of a neuron depends on every value of the input, convolutional layers retain spatial coherence as these neurons only use the values of the pixels in the neighborhood. This is done by applying the formula in Eq. 1.18:

$$y_{i,j}(x) = f(\sum_{m=0}^{c} \sum_{k=\lceil -\frac{n}{2} \rceil}^{\lfloor \frac{n}{2} \rfloor} \sum_{l=\lceil -\frac{n}{2} \rceil}^{\lfloor \frac{n}{2} \rfloor} x_{i+k,j+l,c} \cdot w_{k,l,c}) + b \tag{Eq. 1.18}$$

where y is the output of the convolution layer, f is the activation function, b the bias, x an image input with c channels, n the size of the convolution kernel, and w the weight at position (i, j, k). Therefore, a single convolution defines $c \cdot n^2 + 1$ parameters, which is significantly lower than the classic perceptron that would define $c \cdot I \cdot J$ where I and J are the width and height of x. One particularity of the convolutional layers is that their weights do not depend on the input size (except for the number of channels), unlike dense layers. This property can be used to run the model on images with various shapes.

Fig. 1.7 depicts a convolution layer with two convolution kernels running on an input of shape (5, 5, 3). As such, each convolution kernel is composed of 27 weights plus a bias. Following equation Eq. 1.18, the pixel $y_{1,3,0}$ is given by equation Eq. 1.20.

$$y_{1,3,0}(x) = f(\sum_{m=0}^{c} \sum_{k=\lceil -\frac{n}{2} \rceil}^{\lfloor \frac{n}{2} \rfloor} \sum_{l=\lceil -\frac{n}{2} \rceil}^{\lfloor \frac{n}{2} \rfloor} x_{1+k,3+l,c} \cdot w_{k,l,c}) + b \tag{Eq. 1.19}$$

$$\begin{aligned}
&= f(a_{0,0,0} \cdot x_{0,2,0} + a_{0,1,0} \cdot x_{0,3,0} + a_{0,2,0} \cdot x_{0,4,0} \\
&+ a_{1,0,0} \cdot x_{1,2,0} + a_{1,1,0} \cdot x_{1,3,0} + a_{1,2,0} \cdot x_{1,4,0} \\
&+ a_{2,0,0} \cdot x_{2,2,0} + a_{2,1,0} \cdot x_{2,3,0} + a_{2,2,0} \cdot x_{2,4,0} \\
&+ a_{0,0,1} \cdot x_{0,2,1} + a_{0,0,1} \cdot x_{0,3,1} + a_{0,2,1} \cdot x_{0,4,1} \\
&+ a_{1,0,1} \cdot x_{1,2,1} + a_{1,0,1} \cdot x_{1,3,1} + a_{1,2,1} \cdot x_{1,4,1} \\
&+ a_{2,0,1} \cdot x_{2,2,1} + a_{2,0,1} \cdot x_{2,3,1} + a_{2,2,1} \cdot x_{2,4,1} \\
&+ a_{0,0,2} \cdot x_{0,2,2} + a_{0,0,2} \cdot x_{0,3,2} + a_{0,2,2} \cdot x_{0,4,2} \\
&+ a_{1,0,2} \cdot x_{1,2,2} + a_{1,0,2} \cdot x_{1,3,2} + a_{1,2,2} \cdot x_{1,4,2} \\
&+ a_{2,0,2} \cdot x_{2,2,2} + a_{2,0,2} \cdot x_{2,3,2} + a_{2,2,2} \cdot x_{2,4,2} \\
&+ k_a)
\end{aligned} \tag{Eq. 1.20}$$

The number of weights in a convolution layer, 56 in the case of Fig. 1.7, depends on the size of the kernels, their number (which is the number of output channels), and the number of input channels. In particular, and in contrast to "dense" layers composed of perceptrons, it does not depend on the dimensions of the input. For comparison, an input of shape (5, 5, 3) and an output of shape (5, 5, 2) would require the optimization of 3800 parameters.

Pooling

Pooling is a shrinking of the input vector space dimension. Given a stride s and a function f whose domain has a higher dimension than its codomain, pooling consists of applying f every s pixels. A maximum or a mean over a close neighborhood are common choices for the pooling function. Pooling is often used to decrease the size of the activation maps and increase the spatial range of the model (also known as its receptive field). A so-called global average pooling can also be used in the categorization task to entirely pool the spatial dimension and leave the activation as an array whose length is the number of channels. However, to pool over the channels, it is often preferred to use a convolution kernel which acts as a weighted mean.

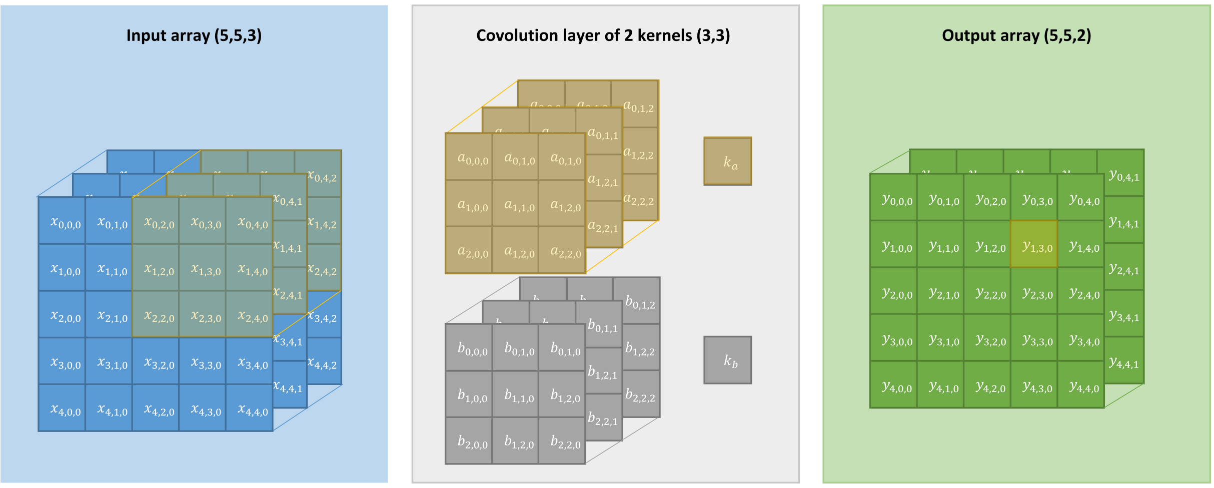


Fig. 1.7: A convolution layer with 2 kernels takes a 3-channel image of 5 by 5 pixels as input and outputs two activation maps of 5 by 5 pixels. The yellow overlay represents the values used to compute $y_{1,3,0}$.

Concatenation

Sometimes, two layers of the model have to be combined within a single layer. This typically occurs when one of the layers is passed through a pooling layer, then upscaled, and combined with a layer that remained at the same resolution. This concatenation is used to take a set of vectors as input and return a single vector whose length is equal to the sum of the length of each input.

Transposed Convolution

Eq. 1.18 assumes a stride equal to one pixel, meaning that the convolution window, from which the input pixels are multiplied by the kernels, moves by only one pixel. As a result, it produces an output with the same spatial dimensions. However, the stride can be greater than 1. In this case, the convolution is computed every s pixels, where s is the stride, and the spatial dimensions are reduced. The stride can also be lower than 0. Operatively, it can be done by padding in between the actual values [45] and artificially inflating the input array. In this case, the spatial dimensions of the output are higher than those of the input. This process, called "transposed convolution" or "deconvolution", is used to upscale the activations map.

1.3.2 Translation Equivariance & Receptive fields

Translation Equivariance

Eq. 1.18 implicates the property of translation equivariance since, if we note $\dot{x}$ the translation of x by the two integers α and β so that $\dot{x}_{i,j} = x_{i+\alpha,j+\beta}$, we have:

$$\begin{aligned} y_{i,j}(\dot{x}) &= f(\sum_{m=0}^{c} \sum_{k=\lceil -\frac{n}{2} \rceil}^{\lfloor \frac{n}{2} \rfloor} \sum_{l=\lceil -\frac{n}{2} \rceil}^{\lfloor \frac{n}{2} \rfloor} \dot{x}_{i+k,j+l,c} \cdot k_{k,l,c}) + b \\ &= f(\sum_{m=0}^{c} \sum_{k=\lceil -\frac{n}{2} \rceil}^{\lfloor \frac{n}{2} \rfloor} \sum_{l=\lceil -\frac{n}{2} \rceil}^{\lfloor \frac{n}{2} \rfloor} x_{i+k+\alpha,j+l+\beta,c} \cdot k_{k,l,c}) + b \\ &= y_{i+\alpha,j+\beta}(x) \end{aligned} \qquad \text{(Eq. 1.21)}$$

Eq. 1.21 indicates that the translation of the input is equivalent to the translation of the output[1]. This is not entirely true since both Eq. 1.18 and Eq. 1.21 ignore cases where $i+k$ (resp. $j+l$) is lower than 0 or higher than I (resp. J). In these cases, the missing values need to be filled. This is commonly done by padding x with zeros. Other solutions include reflecting the values from the images to obtain $x_{-a,-b} = x_{a,b}$.

Translation equivariance expresses the homogeneity of the model's processing, as, with border effects excluded, two areas in different positions will be computed with the same weights.

Receptive Field and Border Effect

Previously, it was highlighted that the acquired images have thousands of pixels in both directions. Due to limitations on GPU memory, it is not possible to feed the entire observations to the model. To overcome this, the inputs must be divided into tiles and the model must be run on each tile. Fig. 1.8.b illustrates an example of this tiling, or mosaic. The original image Fig. 1.8.a is divided into tiles of 128x128 pixels. However, the left eye of the woman, located in the center of the original picture, is near the intersection of the tiles. This means that the predictions of the pixels at each corner of the intersection

---

[1] As a side note, CNNs that are not fully convolutionnal (meaning that they contain dense layers in addition to convolutionnal layers) are often translation invariant as the use of a global pooling layer averages the activation maps on all spatial dimensions. Thus, any translation of the input will result in the same output.

will be computed with different contextual information, leading to a loss of translation equivariance demonstrated in the previous section. Furthermore, the lack of contextual information in one or two directions will degrade the performance of the models. Fig. 1.8.c shows tiles generated with a stride of 64 pixels. The left eye of the woman now appears on four tiles, and is especially in the center of the center tile. In this example, one pixel of the original image appears at more than 32 pixels from the border in at least one tile.

Fig. 1.8: Example of the creating of an overlapping mosaic. (a) is the original image of 256x256 pixels. (b) is a mosaic of tiles of 128x128 pixels build with a stride of 128 pixels. (c) is a mosaic of tiles of 128x128 pixels build with a stride of 64 pixels.

Fig. 1.9 illustrates the propagation of information through a convolutional neural network. The input for this illustration is an image with all pixels set to zero except for the center pixel, which has a value of 1. Fig. 1.9.1 indicates the output of a CNN when fed this almost-null input. Column (a) shows the output of a CNN when given this almost-null input. In column (a), the model is randomly initialized with no bias. From Eq. 1.18, we can see that if both the bias and the input are zero, the output will also be equal to zero. Therefore, non-zero values indicate the range of information propagation from the center pixel. Line (2) is the polar projection of line (1) and shows a propagation of around 30 pixels. The anisotropy observed is caused by the random initialization.

In columns (b) and (c), the model has been trained on a SAR dataset (from Chapter 6). In column (b) and (c) the model is run on 128x128 pixels and 256x256 pixels respectively. As it is a fully-convolution model, with only the weights of the convolution kernels and their bias defined, this model can be used on inputs of different sizes. It can be observed that the input center pixel has a larger impact on the output, extending up to around 80 pixels. The proximity of the border also has a significant impact on the output value, up to more than 64 pixels, which is half the size of the image. In column (c), the border has the same extent, but the larger image dimensions result in more pixels having accurate values. The differences between columns (a) and (c) suggest that the same architecture can have a different spatial information propagation depending on the weights, making it difficult to estimate this characteristic beforehand.

However, it is possible to compute a higher bound of this spatial information propagation, also called the "receptive field" of the model. It can be computed iteratively on simple models by considering the following rules:

- Convolution layers increase the local receptive field by $\lceil \frac{k}{2} \rceil$ where k is the width of their kernel. They divide the local resolution by their stride s.
- Dense layers do not change the local resolution but, as they discard spatial information, they have an infinite receptive field.
- Pooling layers reduce the local resolution by dividing it by their stride s. They also marginally impact the receptive field, as the local receptive field, after the pooling operation is $\lceil \frac{e_{i-1}}{s} \rceil$.

The iterative formula for the receptive field at layer n, depending on the stride s_n and the kernel width k_n, is indicated in Eq. 1.22.

$$e_n = \begin{cases} \left\lfloor \frac{e_{n-1}}{s_n} \right\rfloor & \text{if pooling} \\ \left\lfloor \frac{e_{n-1}+k_n}{s_n} \right\rfloor & \text{if convolution} \\ +\infty & \text{if dense} \end{cases} \tag{Eq. 1.22}$$

Since transposed convolutions have a stride lower than 1, they are a major contributor to the receptive field. A large number of pooling/transposed convolution increases the size of the contextual information accessible to the network, but it can also lead to border effects. The higher number of weights can also lead to overfitting, which is detrimental to the model's ability to generalize beyond its training set. It is therefore necessary to balance the receptive field to retrieve enough spatial context without being hindered by drawbacks. Some hints on how to solve this issue are presented in Chapter 4.

1.3.3 Convolutional Neural networks for Earth Observation

One of the strengths of the CNN is their ability to be used on a wide range of problems. They have long been used in segmentation, as the estimation of the land use in optical

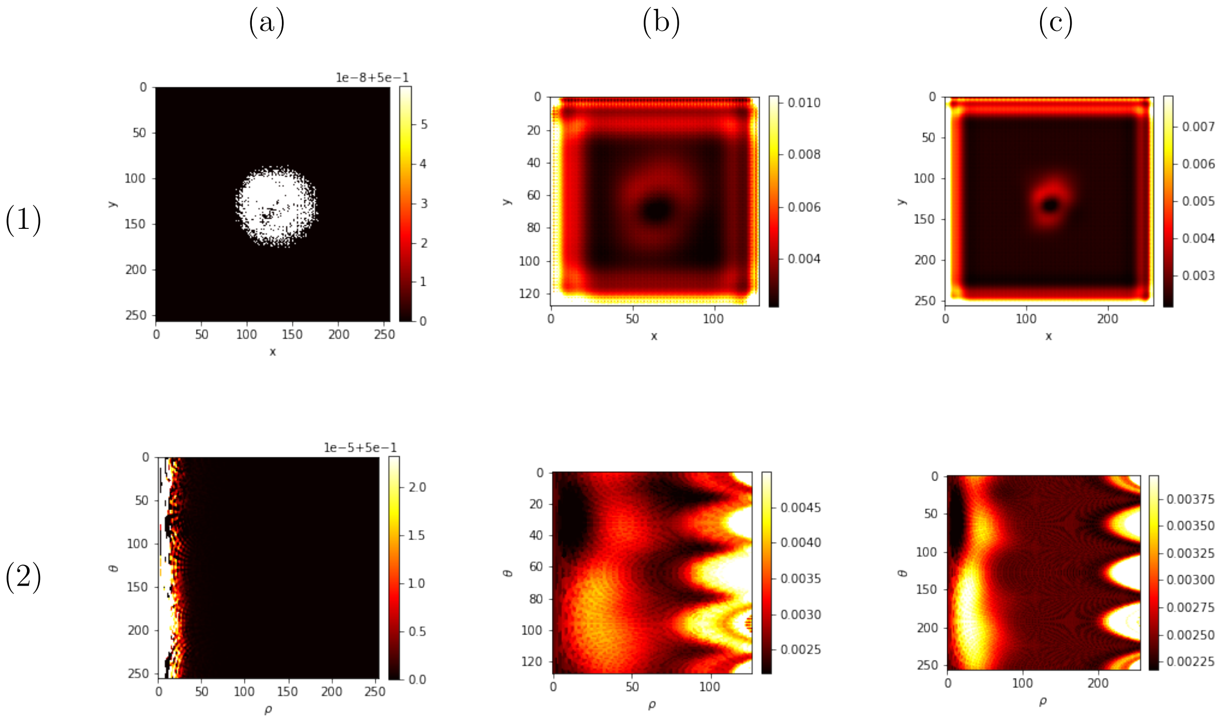


Fig. 1.9: Output of a model using an image with every pixel at 0 except at the center. The center pixel of the input is set at 1. The model is randomly initialized (a), trained for images of 128x128 pixels (b) and transferred for images 256x256 pixels (c). (1) is the output of the model, (2) the polar projection from the center of the output.

imagery is a popular research subject [86, 135] and benefit from the accumulation of data through the years [58]. Segmenters can also be used with other sensors, for example on SAR to delimit rivers [146], floods [142], or roads [68]. On multispectral observations, they are used to segment clouds [43] or survey crops [145, 229, 90] among other applications.

Since the acquisitions are geo-referenced, it is possible to use multiple observations of the same area to detect change. This problem is close to segmentation but differs by the use of multiple inputs acquired at different times. It is particularly used to detect change of buildings [152, 5, 53], forests [63, 79] or land cover in general [139].

Spatial observations and deep learning models have also been used to detect ships [230, 216] in which case the output isn't a segmentation but a set of coordinates, or to classify thumbnails [75] by a single-vector outputs. In particular, the TenGeoP-SARwv [208] is a dataset of ten metocean processes visible on SAR images.

Deep learning can also be used to enhance the observation, by reducing noises [103, 50] or reconstruct parts of the observation that could either be masked clouds [133, 225].

It can be used to increase the resolution [51, 162, 102, 7], or to convert one observation in the style of another sensor that could be in the same domain but different resolution [126] or of a different kind [159]. Synthetic observations can also be generated [33] and used for data augmentation. In this case, the model often uses a GAN framework [60] to add an implicit regularization term in the loss function.

PART II

Deep learning schemes for the semantic segmentation of the ocean surface from SAR observations

FOREWORD

As raised in the previous chapter, deep learning models rely on groundtruth to learn the transformation from the input manifold -in our case the space of the SAR observations- to the desired segmentation. For some problems, the groundtruth can be obtained through collocation with secondary instruments. This typically happens when data from one source are available on one area, but can not be used globally. In such case, the observations by this third-party instrument are collocated with Sentinel-1 observation to produce groundtruths. Learning the transformation from the SAR image to the third-party sensor allows to extend the area on each the collocated information is available.

However, this is not always possible. Some processes of interest need to be manually annotated since remote information is difficult to obtain. It can also happens than only punctual information is available, meaning that the information is not spacialized. If an underlying implicit spacialized information can be infered from the punctual data, a framework called Weakly-supervised learning can be used. Another solution is to manually annotate the desired groundtruths. This process is particularly time-consuming, often leading to a smaller (but cleaner) dataset.

Chapter 2 introduces the context and the difficulties of semantic segmentation. It focus on the biological slicks. Chapter 3 is a study of ten different classes of meteorological and ocean processes and a comparison of the the manual pixel-level supervised learning with weakly-supervised frameworks using image-level information. It demonstrates the need of manual annotations. Finally, Chapter 4 focus on the multiscale aspect of some metocean processes, namely the convective cells, for which both high-resolution and large spatial context are needed to obtain satisfactory segmentations.

Chapter 2

Context and Challenge: Biological slicks as a case-study

2.1 Introduction

In the previous chapter, we introduced some examples of machine learning applications to SAR imagery. In the case of the TenGeoP-SARwv dataset, it could be summarized by the question "What is the primary metocean process in this sample?" By design, we assume that one image has only one associated class, but this is not always the case: one image can have multiple labels. One example is the *Biological Slicks*. *Biological Slicks* are large areas of the ocean surface that are covered with a thin film of living organisms. They are usually created by the accumulation of phytoplankton that form the base of the marine food web [77]. They are usually found in areas of upwelling, where currents bring nutrient-rich waters to the surface. The waves, in areas covered by the slicks, are damped by the high viscosity of the microfilm over the surface, leading to a decrease in Sea Surface Roughness. As they provide an important source of food for marine animals, fishing vessels can often be seen in their vicinity. As such, the corresponding observation could correspond to either the "*Bright Target*" or the "*Biological Slicks*" classes. This categorization only provides image-level information. It does not indicate the location of the processes. Segmentation is akin to categorization since it provides binary information about a class, but it is estimated at the pixel-level rather than the image-level.

The slicks have a biological interest as they are associated with a higher density of fish (and plankton) [92], as such, they are searched by the fisheries. Their morphology can help to locate sub-mesoscale marine eddies since their shape is impacted by the surface currents [84]. Their dispersion is also affected by the wind speed [171] and by the Stokes drift caused by the waves [22]. The shape of slicks can give clues about how they were produced. For instance, annular-shaped slicks may point to coral spawn [83]. From a pure image processing point of view, they often appear to be similar to oil spills, which is a similar problem on which SAR imagery has been applied [195].

This chapter aims to illustrate the methodology used to train a segmentation model. To simplify the problem formulation, we follow a fully-supervised framework, meaning that the annotated groundtruth is exactly what the model is expected to return. This chapter also highlights the difficulty of interpreting a SAR observation with different processes appearing with the same low-level features. Comparison with Sentinel-3 ocean colour measurements are used to confirm the result by comparing the probability of slick appearance with the estimate of the chlorophyll concentration.

2.2 Dataset

For this short introduction, seven manually annotated IW, acquired between June 02nd 2018 and June 11th 2021 are annotated at the pixel-level. These IW are selected independently of their position through manual selection to maximize the appearance of the slicks. Two IW were observed by Sentinel-1B, the remaining five were obtained by Sentinel-1A. The small number of IW is explained by the small scale of the slicks and their extent on the observations. It leads the annotation task to be particularly time-consuming (around 45 minutes per observation). The resolution of the observations, and their corresponding groundtruth, is 100 m/px. Samples of these IW are depicted in Fig. 2.1. After the annotation, 1032 patches of 25.6 x 25.6 km are extracted and divided into training, validation, and test subsets. The split is performed at the IW-level rather than the patch-level to avoid any risk of data leakage.

When wind speed is lower than 2 m/s, it is difficult to detect Biological Slicks due to low backscattering. Though they are present, the contrast with the surrounding waters is too low to make them visible. Additionally, at moderate or high wind speed, the slicks are dispersed and disappear quickly. The annotated dataset reflects this distribution, as shown in Fig. 2.2, with the wind speed range corresponding to the atmospheric conditions restricted to an interval between 2 m/s and 5 m/s.

The trained model is also applied to observations of the Mediterranean and Black Seas. In this case, as manual groundtruths are not required, it is possible to use a higher number of images. For this analysis, 92422 IW acquired between 2015 and 2021 (both included) are used.

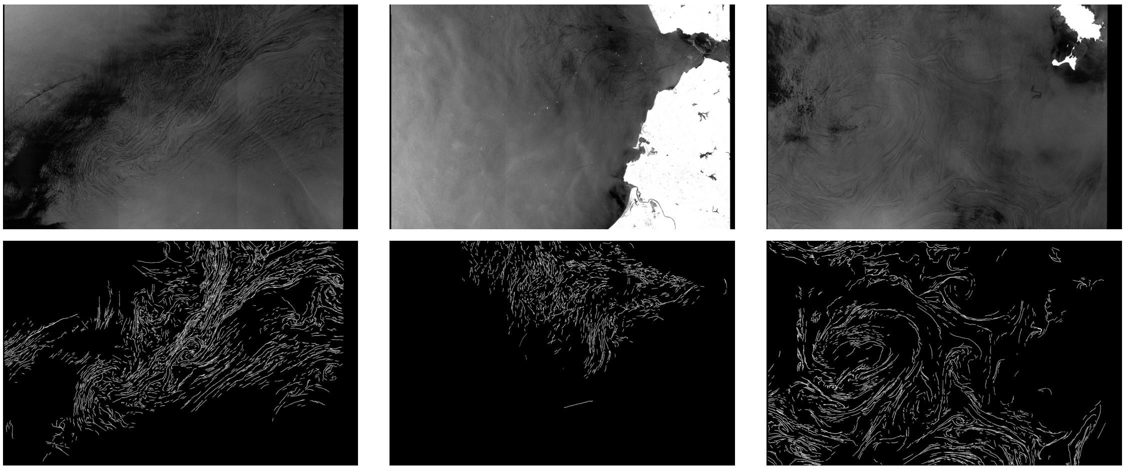


Fig. 2.1: Samples for the *Biological Slick* dataset. Observations were performed respectively on September 24th 2018 at 05:45:47, March 05th 2019 at 18:25:40 and March 16th 2019 at 05:53:54. The first line is the SAR observation, the second line is the manual segmentation.

2.3 Convolutional Deep Learning Model

From the training set, a U-Net [234] model is trained to minimize the Binary Cross Entropy between the prediction from the model and the groundtruth. Though cross-entropy loss would be standard for categorical segmentation (and binary segmentation in particular), the MSE has the advantages of being symmetric, meaning that $MSE(a,b) = MSE(b,a)$, and of being less prone to numerical instability. The training used Adaptive Momentum, a learning rate of 10^{-5} and is performed over 360 epochs.

The model architecture, presented in Fig. 2.3, is composed of an encoder and a decoder. The encoder, made up of connected convolution blocks, reduces the input size by performing convolutions and shrinking the image. While the spatial dimensions are reduced, the spectral dimension is increased as the processed data progresses further into the network. The second part of the architecture is a decoder. It retrieves the activation maps generated by the encoder and increases their size through transposed convolutions. The difference between U-Net and a classic autoencoder is the presence of transversal connections between corresponding blocks rather than only at the lowest resolution. This facilitates the preservation of finer details and reduces the likelihood of the "vanishing gradient" issue that is often encountered with deep learning models.

After training, the model can be applied to other SAR observations. To process IW products, the wide-swath image is divided into tiles of 20x20 km, each tile having overlap

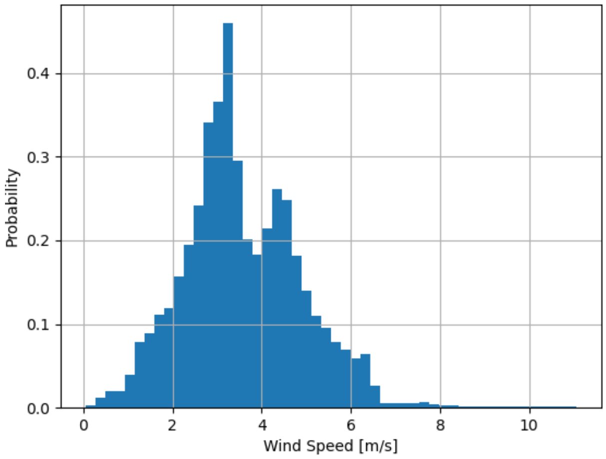


Fig. 2.2: Wind Speed distribution of the seven IW of the dataset, as obtained from the ECMWF atmospheric model.

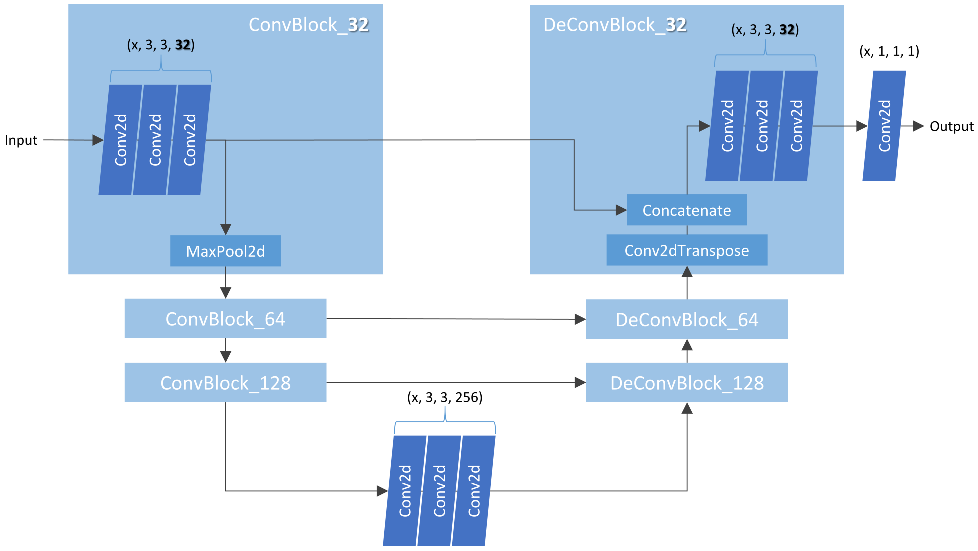


Fig. 2.3: Architecture of the U-Net model used to segment *Biological Slicks* using Sentinel-1 ocean surface roughness.

with its neighbors. We can note that, since the segmentation model is fully convolutional, it can be applied to larger images as long as the resolution stays at 100 m/px. However, for simplicity's sake (and to reduce the strain on the GPU memory), we often choose to stay at the same resolution, even if this is detrimental to the segmentation quality (mitigated by the overlap strategy) and the processing speed. For the latter, we recorded a typical duration of a dozen seconds to process a full IW on a GTX 1050 Ti. With more modern GPUs, the processing of the IW is mainly affected by the download of the L1

GRD product and its preprocessing from 10 m/px to 100 m/px, effectively decreasing the file from several hundred Mo to several Mo.

2.4 Results

To assess the validity of the deep learning model, the predictions on the test set are compared to the groundtruths. The test set is made up of elements that were not used for either training or validation. Validation refers to the evaluation of the model during training to survey overfitting. As the test set is composed of data ignored during training, it provides information on the model's capacity to generalize beyond its training set. In this study, due to the lack of annotated data, only one IW (134 patches) was used.

Due to the difficulty of obtaining data that has been manually annotated, predictions are made on a large pool of data that has not been annotated, but whose metocean context is known. First, the frequency of *Biological Slick* segmentation is compared to the estimate of the chlorophyll using the ocean colour in the Alboran Sea, located at the East of the Gibraltar Strait. Second, the capacity of the model to segment pollution -whose SAR signature is similar to Biological Slicks- is assessed on the east coast of the Black Sea, where oil is known to surface from the sea bed.

2.4.1 Results on the test set

Fig. 2.4 illustrates the result of the segmentation over the IW of the test set, which was not used during either the training or the validation. On this observation, acquired in the vicinity of the Mississippi Delta, several bright objects concentrate in an area of low wind speed. In this area and the surrounding waters, *Biological Slicks* are being segmented by the model. The manual annotation also indicates the presence of these slicks. The bright targets (fishing vessels) are explained by the abundance of plankton that draws the fish to it. The model reached a Binary Cross Entropy of 0.1919.

2.4.2 Seasonal biological activity in the Alboran Sea

As the *Biological Slicks* are indicators of plankton, they are expected to correlate with the chlorophyll concentration [114]. The latter can be estimated from the ocean color [190]. Figure Fig. 2.5 demonstrates that ocean color and the detection of Biological Slicks are

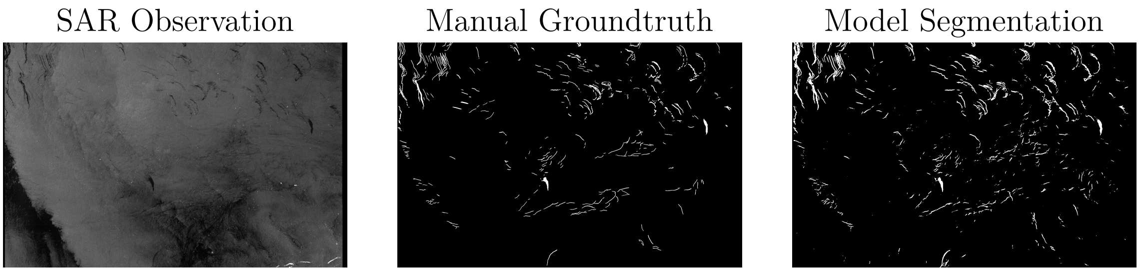


Fig. 2.4: Segmentation of the IW of the test set, observed on the 11th of June 2021 at 00:10:00.

correlated, as the gradient of both reflects the same current patterns. The two observations were separated by a 17-hour interval.

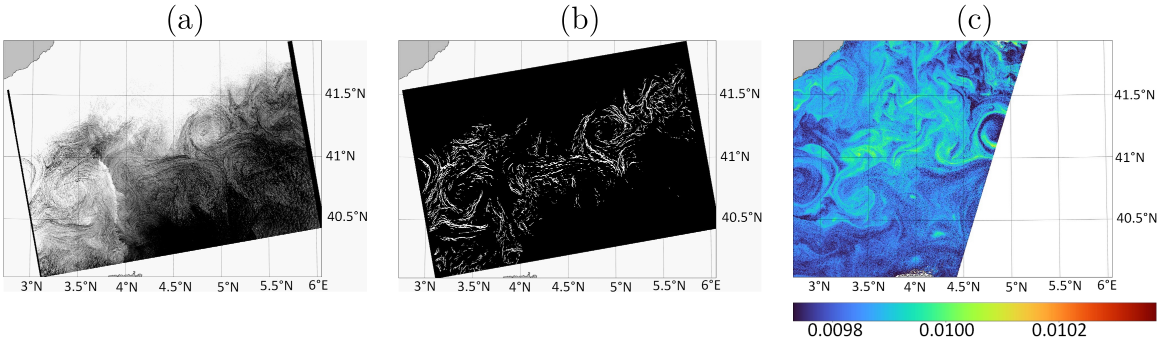


Fig. 2.5: SAR observation acquired on the 21st of February 2019 at 17:38:18(a). Segmentation of the *Biological Slicks* (b). Chlorophyll concentration (mg/m^3) estimated from Sentinel-3 OLCI obtained on the 22nd of February 2019 at 10:24:14.

This correlation is visible in the Alboran Sea. Overall, the Alboran Sea is known to be an area with particularly high primary production compared with other regions of the Mediterranean Sea [198]. This sea, at the East of the Gibraltar Strait, is characterized by a seasonal gyre -the Western Alboran Gyre- appearing during spring and stable during summer [202]. As depicted in Fig. 2.6.b4, the chlorophyll concentration obtained from the ocean colour is lower in the center of the gyre. These figures are obtained from the L2 products of the Ocean and Land Color Instrument (OLCI) sensor aboard the satellites of the Sentinel-3 serie. The measurement period is from 2018 to 2021 (both included).

The Sentinel-1 SAR observations over the Alboran Sea are systematically segmented for the period between 2018 and 2021 (to correspond to the availability of Sentinel-3 data). They aggregated on a grid of $\frac{1}{25}$ degree. Comparison with Sentinel-3 data highlights the correlation between the presence of the slicks and an higher chlorophyll concentration. The Eastern Alboran Gyre is also visible in summer (Fig. 2.6.4) and autumn (Fig. 2.6.5). To obtain these figures, observations with wind speed higher than 6 m/s were removed. The wind speed information is obtained from ECMWF atmospheric model.

As a side note, we can observe a decrease of the estimated *Biological Slicks* distribution on a line going approximately east-north-east from the Gibraltar Strait. One possible cause for this is the wake generated by ships passing through the Gibraltar Strait, which can dislocate the slicks and make them more difficult to see on SAR observations.

The same study can be performed on the entire Mediterranean Sea. To draw the maps presented in Fig. 2.7, 92422 IW-GRD acquired from 2015 to 2021 (both included) are processed to a resolution of 100 m/px and fed to the model. The aggregated maps are mean-pooled to a resolution of 0.1°. The results are coherent with chlorophyll maps [13], indicating a lower primary production on the eastern basin. It interesting to note that the difference of ocean colour in the Black Sea and Azov Sea is not reflected in the estimate from Sentinel-1 SAR measurements. We assume that the difference of salinity or/and temperature are too important to allow a simple comparison. Further studies should be conducted to better understand the discrepancies between Sentinel-1 and Sentinel-3 observations of biological activity.

2.4.3 Detection of cold seeps in Black Sea

Although the model is trained to recognize *Biological Slicks*, it may also return some phenomena with similar features as biological slicks, such as oil spills. Most of the time, oil spills are caused by ships cleaning their tanks. However, oil spills can also occur naturally, such as through cold seeps in the Black Sea [104]. When oil emerges from a cold seep, it can move depending on the wind and currents.

The area around the Kobuleti Ridge has been particularly studied because of the Pechoti Mound and the Colkheti Seep, as both locations are known to leak oil [48, 157, 98]. Examples of observations from this area, along with the associated segmentation, are presented in Fig. 2.8. In the first observation, the oil slick is easily visible on the SAR image and is correctly segmented. The second example shows an often-occurring wind gust descending from the Rioni River. The strong winds disperse the oil, making its signature less visible. In the last two examples, *Biological Slicks* are present throughout the entire area and highlight the presence of submesoscale currents. It is worth noting that, in Fig. 2.8, the coast appears on the the left despite being on the east side of the Black Sea. This

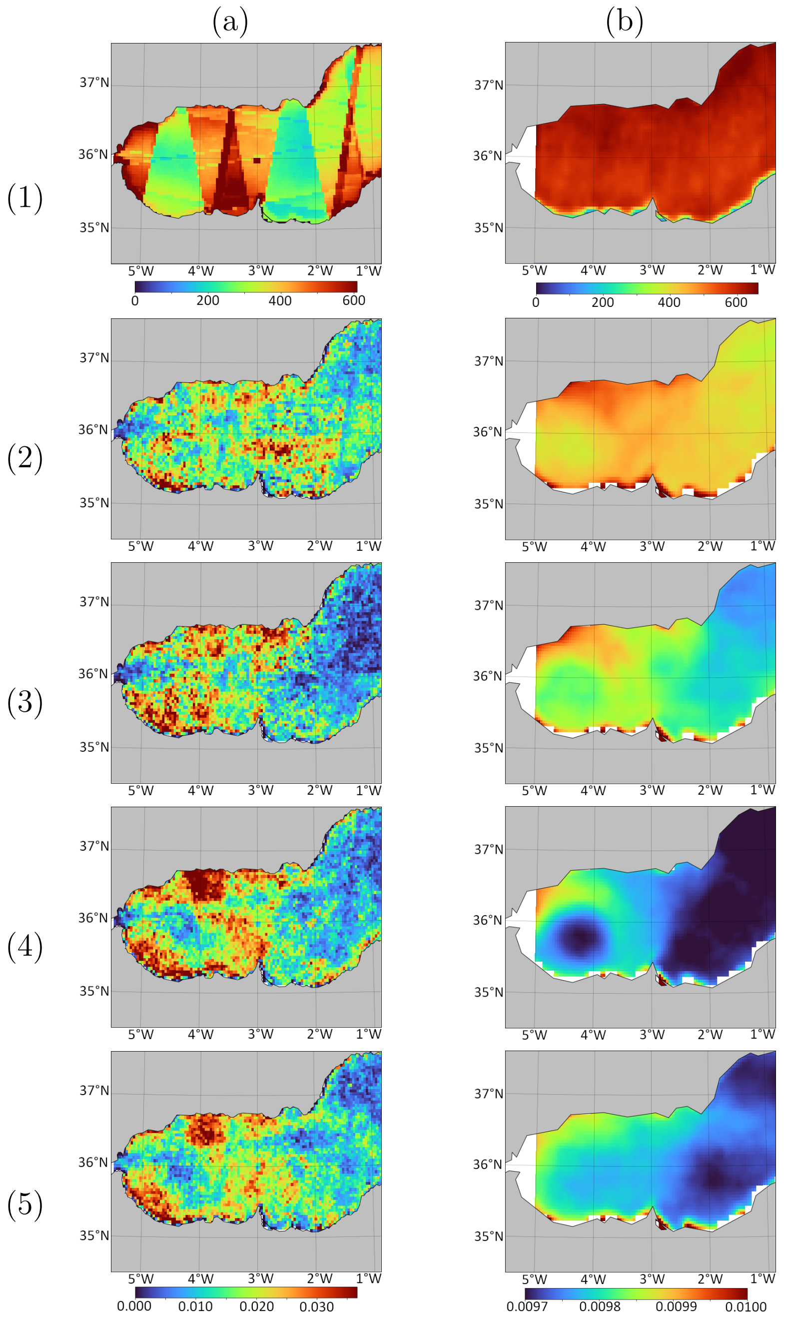


Fig. 2.6: Comparison of the probability of *Biological Slick* detection from the segmentation of Sentinel-1 IW (column a) and the chlorophyll concentration from Sentinel-3 OLCI (column b). Line 1 is the sum of available IW. Line 3 to 6 are the seasonal mean to respectively December to February, March to May, June to August and September to November, computed for 2018 to 2021 (both included). Points hidden by clouds (for Sentinel-3) or at wind speed higher than 6 m/s (for Sentinel-1) are removed.

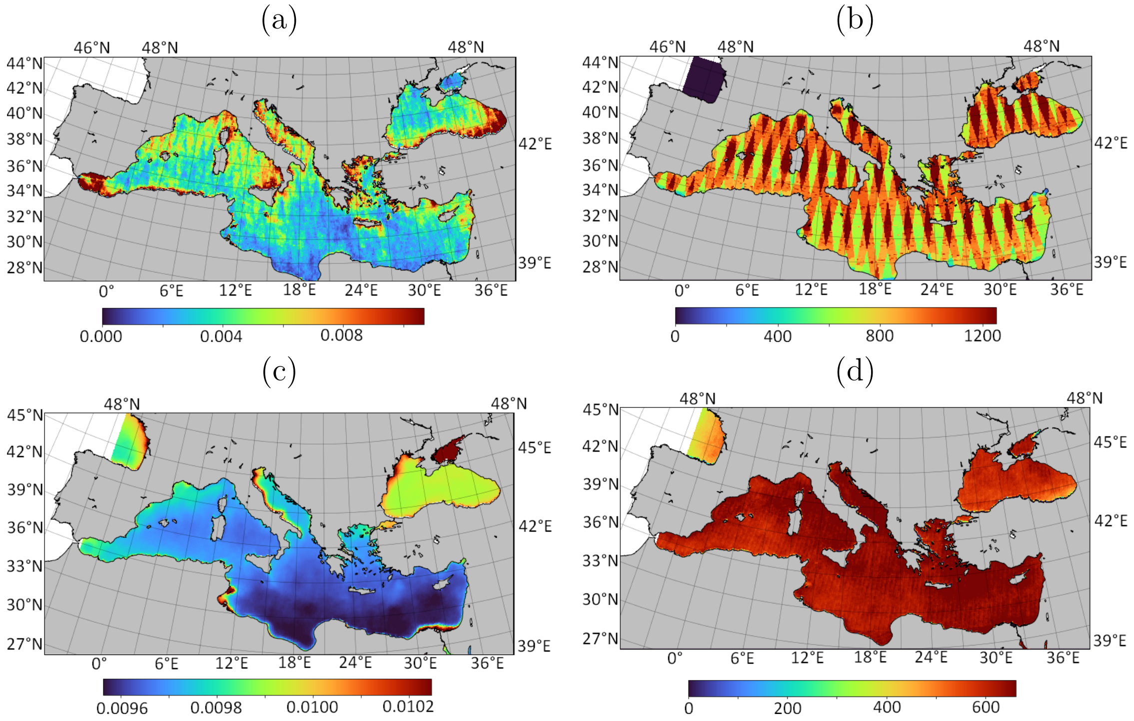


Fig. 2.7: Mean probability of *Biological Slicks* detection from SAR images over the Mediterranean Sea between 2015 and 2021 (a). Number of IW used to compute the mean (b). Mean chlorophyll concentration obtained from Sentinel-3 OLCI between 2018 and 2021 (c) and number of swath acquired by Sentinel-3 (d).

is because the observation is in sensor geometry, with the minimal incidence angle on left side of the observation.

It can be observed that the oil associated with cold seeps has slightly different features than *Biological Slicks*. In the fourth observation, it appears thicker and easier for a human operator to detect as a similar phenomenon. It also seems more resistant to higher wind speeds. When caused by the cleaning of a ship's tank, the human origin of the oil pollution can be inferred by the proximity of a ship. Therefore, it is important to consider the context when identifying processes with similar low-level features.

In Fig. 2.9, the probability of *Biological Slicks* detection, computed over the year 2018 to 2021 indicates the mean segmentation of all available IW between January 2015 and June 2022. A peak is visible at the position of the Pechoti Mound and Kolkheti Seep (resp. 41.123E, 41.983N and 41.104E, 41.969N [48]). Oil leaks appear to be lower during the winter, but this could be due to the higher mean wind speed during this time, which can dislocate the slicks [144].

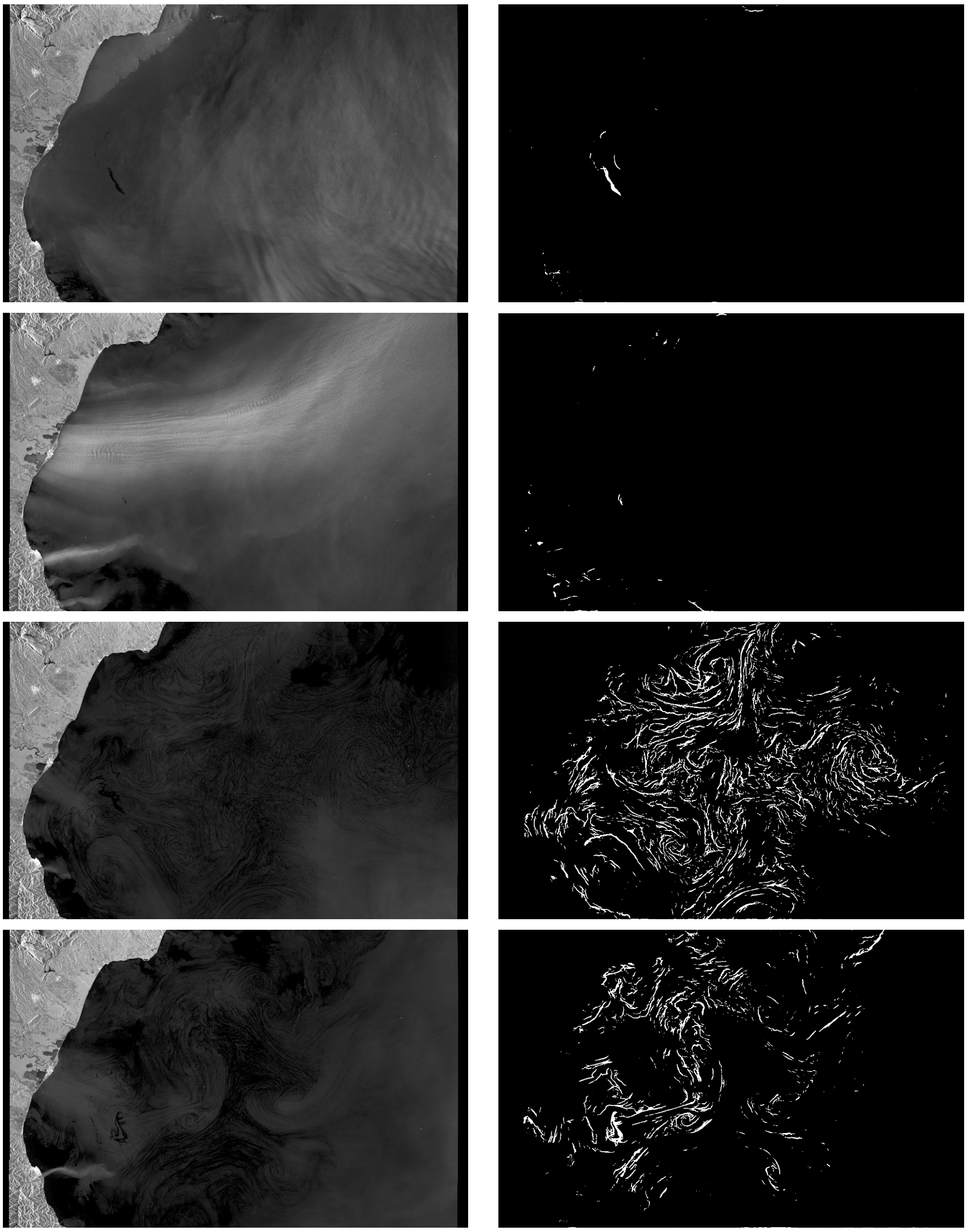


Fig. 2.8: Observations and segmentation of the Kobuleti Ridge were acquired on April 20th 2018 at 03:24:33, November 05th 2019 at 03:24:48, April 23rd 2022 at 03:24:58 and May 15th 2020 at 03:24:47

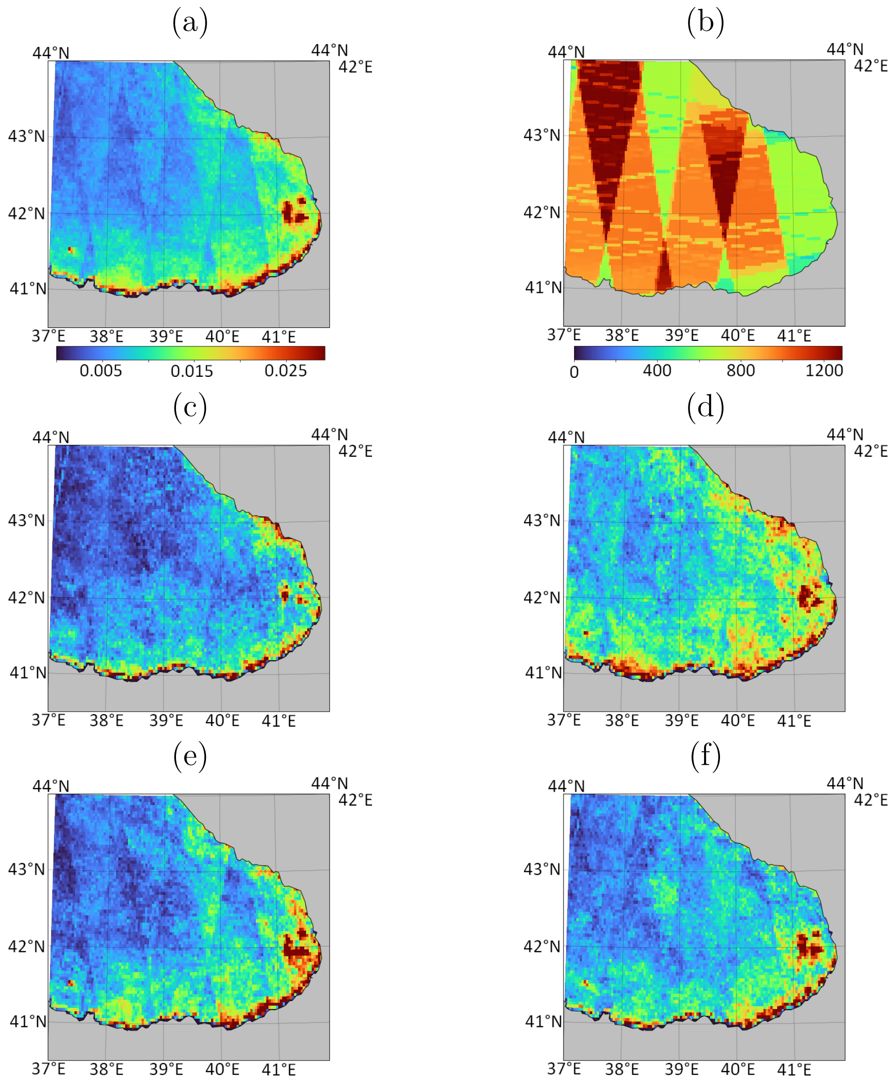


Fig. 2.9: Probability of *Biological Slicks* detection over the Kobuleti Ridge. (a) is the mean over the years 2015 to 2021 (both included). (b) is the number of IW used to compute (a). (c) to (f) are the restriction to respectively December to February, March to May, June to August and September to November.

2.5 Conclusion

Biological Slicks are oily surfactants floating on the ocean surface that dampen waves due to their high viscosity. On SAR images, they appear as areas of low backscattering because of the lower surface roughness.

The methodology presented in this section helps to discriminate Biological Slicks from surrounding waters by using a deep learning model (U-Net) and manually annotated SAR observations from Sentinel-1. The results were found to be consistent with those obtained from ocean color measurements using Sentinel-3. It is possible to identify ocean gyres by the gradient of biological matter, which is lower in the center of the gyres and higher at their borders. However, manual annotations are difficult to collect, resulting in datasets with limited meteorological and oceanographic diversity.

This studies raise several perspective. First, the annotation of additionnal observation is necessary for a more reliable quantitaive assessement of the model capacities. This a time-consuming but necessary process. Second, as hydrocarbons from natural and anthropogenic pollution can have similar signatures, it is necessary to be able to differentiate them. In the case of human-related activity, this distinction is also usefull for pollution surveillance and could be inferred from contextual information such as the presence of ships. Finally, it would be beneficial to research the complementarity between Sentinel-1 and Sentinel-3 to understand the discrepancies in their estimates of biological activity.

Though this chapter exclusively used fully-supervised learning, it highlighted the difficulty of acquiring segmentation datasets. Chapter 3 compare the fully-supervised framework with weakly-supervised methods using image-level categorizations.

Part III use groundtruths provided by colocation with other sensors in order to automatically build large datasets without the need of manual annotation.

CHAPTER 3

SEMANTIC SEGMENTATION OF METOCEAN PROCESSSES USING SAR OBSERVATIONS AND DEEP LEARNING

The following section is the reproduction of [27] of the same author.

3.1 Introduction

Since neural-networks-based algorithms have become the state-of-the-art framework for a wide range of image processing problems in the 2010s, the use of deep learning approaches has been extended to many kinds of remote sensing data, including for instance infrared imagery [183], land applications of SAR [236] and SAR-optical fusion [165]. By contrast, the segmentation of ocean processes from SAR images has been little addressed. Synthetic Aperture Radar (SAR) imagery relies on physical basis very different from optics. Operating in the microwave band, SAR images result from a complex physical imaging process compared with optical imaging, which results in high resolution images (typically few meters resolution). Over the oceans, the SAR imagery is sensitive to the surface roughness which can be impacted by the wind, the waves, the presence of ships or icebergs, a surface viscosity difference caused by oil or biological slicks, the precipitations or by sea ice in polar regions [81]. Contrary to optical images, SAR can be acquired in almost all conditions, especially over cloudy regions and at night. For these reasons, SAR images can be processed into a wide variety of geophysical products such as wind maps, wave spectra, surface currents, and its full potential still remains unexploited, especially in terms of scientific and operational services [111]. Among the instruments in flight, the C-band Sentinel-1 constellation operated by the European Space Agency (ESA) is one of the main real-time providers of SAR images and products. The segmentation of ocean processes from SAR has already been addressed, but on specific cases such as oil spills

[15], ships (though as a detection problem as they appear as ponctual bright areas) [36] or sea-ice [54]. These studies found out that the Convolutional Neural Networks (CNN) [163, 66] architecture is the most adapted. Especially, [15] found that they achieve better results than other usual segmentation methods such as logistic regression or k-Means.

The present study aims to extend these previous works and more specifically to assess whether deep learning approaches may address the joint semantic segmentation of a wide range of metocean processes in SAR ocean images.

Deep learning for semantic segmentation tasks mainly relies on the fully-supervised framework, that requires the manual segmentation of hundreds up to thousands of images to build a representative training dataset. For SAR imaging over the ocean, the difficulty to access ground-truthed data sets may partly explain the small amount of studies on this topic. First, this task is quite complex and should be carefully performed by domain experts. On the SAR observations, the signature of the geophysical phenomena is complex and requires specific knowledges of both SAR imaging and of their impact on the sea-surface roughness. Second, this is a time-consuming task. The Common Object in COntext data set [115], also known as COCO, stated that drawing segmentations required 800 s per object, while the categorization task at image-level requires only 30. A second hardship was the high variance in the annotation, especially when multiple annotators were involved in the data set creation. This problem was also raised by [147] in the context of sea ice charts.

Fortunately, the first annotated SAR images data set for classification was released by [208]. It consists in a collection of more than 37,000 20 × 20 km SAR imagettes for 10 categories of geophysical phenomena: Atmospheric Fronts, Biological Slicks, Icebergs, Low Wind Area, Micro-convective Cells, ocean Fronts, Pure Ocean Waves, Rain Cells, Sea Ice and Wind Streaks. It was later supplemented by another 10,000 images. In [207], deep learning models trained from this dataset led to promising results with an excellent classification performance [206]. However, this effort could not generalize to imagettes with at least two phenomena, and, as a categorization task, is inherently limited in resolution. Still, even image-level annotation integrates implicit information on the represented features, based on the annotator knowledge, and building up on this valuable effort would be of great interest to train semantic segmentation models [67, 217]. This is referred as weak supervision. Most of the works studying weak supervision are based on photographs and make assumptions to leverage the segmentation, typically on image contrast and homogeneity [154], on boundary pixels [64] or on the object compacity [118]. Another possibility is to turn classifiers into segmenters [204, 224].

The general objective of this study is to evaluate whether deep learning models and strategies may address the semantic segmentation of SAR ocean images in terms of metocean processes. Specifically, we aim to assess the relevance for this task of both fully-supervised and weakly-supervised schemes. The proposed approach combines a selection of state-of-the-art deep learning frameworks, the creation of a reference ground-truthed dataset of annotated SAR ocean images and the design of a benchmarking setting, based on a reference ground-truthed dataset.

In this study, the SAR observations used as groundtruth will first be presented in Section II. Then, the different algorithms able to obtain the segmentation will be presented in Section III. The segmentation of metocean phenomena is addressed through both the weakly supervised framework, using the 37k image-level (one label per image) annotations of the TenGeoP-SARwv dataset, and a fully supervised of 1k manual annotated groundtruth (2D annotations).

3.2 Data

The image-level dataset is built from the TenGeoP-SARwv dataset described in [208]. It contains more than 37k Sentinel-1 Wave Mode (WM) acquisitions categorized at image-level for ten metocean processes. Wave Mode observations are acquired on an area of approximately 20 km by 20 km. Though initially at a resolution of 5 m per pixel, the images are downsampled to 50 m per pixel to reduce the speckle and for computational reasons. The resolution can be further decreased down to 100 $\mathrm{m.px}^{-1}$ without impeding the accuracy of the categorization as shown in Table 3.1. However, it decreases the resolution of the weakly supervised segmentation, as they commonly use a categorizer as a backbone of the segmentation. The scenes are observed under two incidence angles of 23.8° and 36.8° in co-polarization polarization. It has to be noted that the distribution of the classes in the TenGeoP-SARwv dataset varies with the incidence angle. For example, 93% of the Pure Ocean Waves observations were obtained at an incidence angle of 23.8° where 92% of the Micro Convective Cells correspond to observations at 36.8°. Though these two classes are the extreme cases, it indicates that the SAR signature of the phenomena can be more visible at particular incidence angles.

Table 3.1: Accuracy of the categorization task depending on the resolution.

Resolution	**Accuracy (Validation)**	**Accuracy (Test)**
50 $\mathrm{m.px}^{-1}$	99%	77%
100 $\mathrm{m.px}^{-1}$	98%	75%
200 $\mathrm{m.px}^{-1}$	93%	66%
340 $\mathrm{m.px}^{-1}$	88%	59%

Interferometric Wide Swath images (IW) are used for testing purpose only. In contrary to WM, they cover large areas of hundred or thousand of kilometers high and a few hundred of kilometers wide. The incidence angle varies between 32.9° and 43.1°. Also, if the WM observations are acquired with a minimum coastal distance of hundreds kilometers, IW product are obtained in coastal area. In this study, IW samples are preprocessed to match the aspect of the WM images described before. The main preprocessing is the computation of the normalized radar cross section and a de-trending to remove the effect of the variable incidence [205].

The ten metocean processes studied are Atmospheric Fronts (AF), Biological Slicks (BS), Icebergs (IB), Low Wind Areas (LWA), Micro-convective Cells (MCC), ocean Fronts (OF), Pure Ocean Waves (POW), Rain Cells (RC), Sea Ice (SI) and Wind Streaks (WS). Fig. 3.1 indicates the number of observation of these phenomena, over the globe, as determined from the categorization labels from the TenGeoP-SARwv dataset.

For the fully suppervised method only, 100 manually annotated samples per class of TenGeoP-SARwv's are provided for the training set, and 10 per class for the validation set. All algorithms are evaluated on an independant test set of 10 samples per class.

These groundtruths were produced at the pixel-level by manually indicating the area covered by each phenomena appearing in the observations. To account for the fuzziness of the boundary of some phenomena and the difficulty to accurately delimit them, the segmentations are at a resolution of 400 m.px^{-1}. The SAR observation is downscaled to 100 m.px$^-1$ following the results in Table 3.1 as using a resolution of 50 m.px$^-1$ would increase the time and memory consumption with a marginal performance increase. The segmentation dataset is available on kaggle[1].

To artificially increase the data, rotation of $\pm\frac{\pi}{2}$ and symmetries (both vertical and horizontal) are used. The quantity of input is thus multiplied by 8. This data augmentation implicitly assumes that, on the scale of one observation, the variation of the incidence angle has negligible effect against the meteorological state.

---

[1] https://www.kaggle.com/rignak/sar-wv-semanticsegmentation

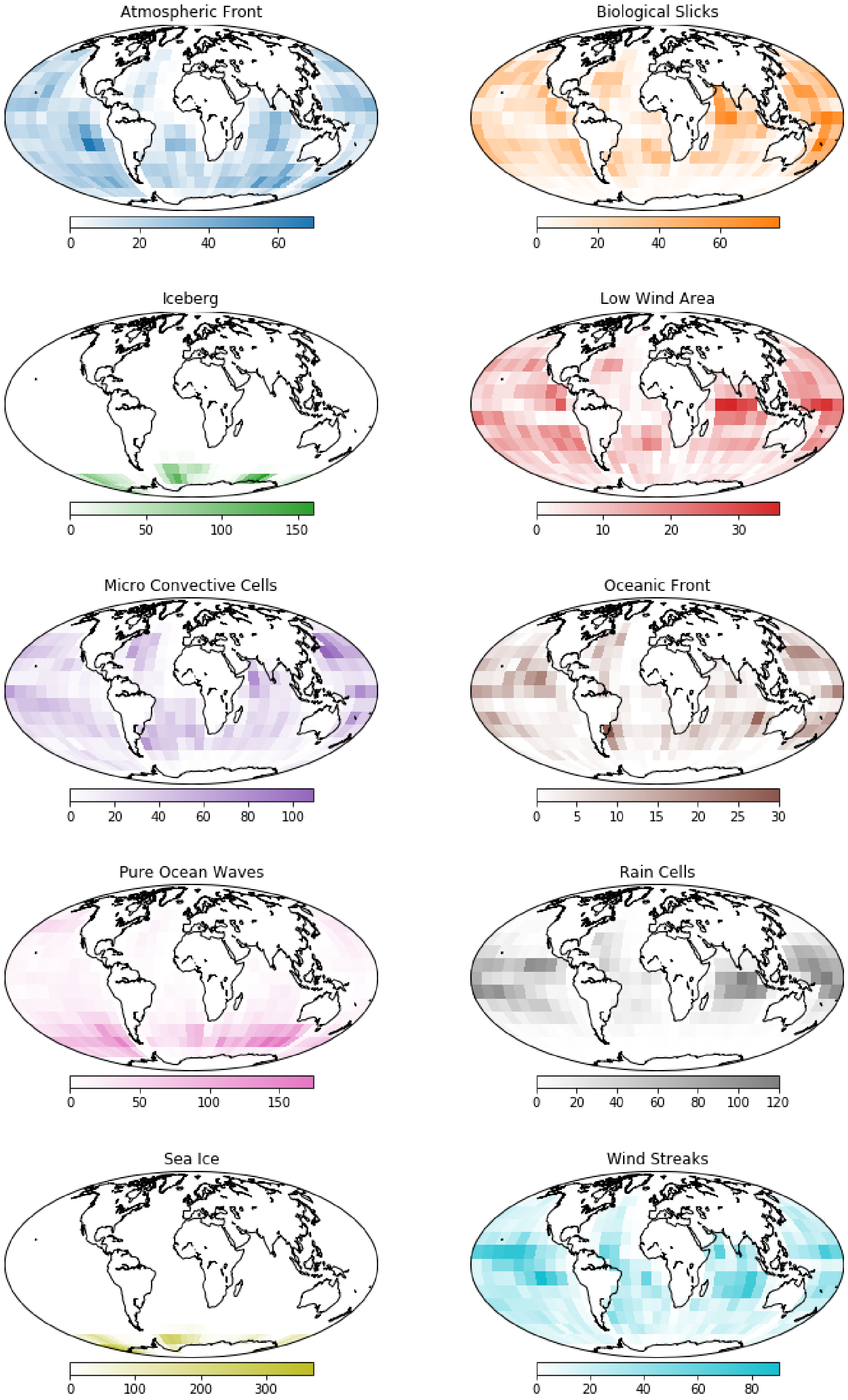


Fig. 3.1: Number of TenGeoP-SARwv's images for each phenomena, projected on grid of 10° by 10°. These distributions are directly obtained from the groundtruth contained in [208].

3.3 Methodology

Five families of algorithms are tested. The first one is trained using the 1000 manual segmentations under the fully supervised framework. The last four rely on the categorization dataset, using the 37k pairs of SAR observation/classification to learn a pseudo-segmentation in a weakly-supervised framework. Specifically:

A uses the small dataset of pixel-level segmentation in a fully supervised framework;

B calls a categorizer multiple times while masking a small part of the input;

C uses the constant model size property to run a categorizer on small portions of the image and get partial categorizations;

D exploits the conservation of spatial information using the Class Activation Maps;

E considers the image-level annotation as a heavily noised pixel-level segmentation and set constraints on spatial information propagation in the model.

The scripts used to train the categorizers and the segmenters are available on GitHub[2].

3.3.1 Fully Supervised Segmentation

All the weakly-supervised methods use the full TenGeoP-SARwv dataset and its more than 37k image-level annotations. However, some previous work stated that supervised learning was successful in SAR imagery for segmentation. Ref. [15] found that the convolutional neural networks achieve better results than other usual segmentation methods such as logistic regression or k-Means to segment pollution on acquisitions from Sentinel-1 and Envisat. Ref. [54] shows that the regression of the sea ice concentration, seen as a segmentation problem with classes representing the concentration intervals, can be solved with a model's architecture similar to U-Net [163]. Ref. [223] used Residual Networks [66] to regress the wind direction on tiles of 5 km by 5 km extracted from interferometric wide swath products from Sentinel-1.

This method relies on the manually pixel-level annotated dataset containing one thousand wave mode in the training set. The segmentation model follows the U-Net architecture [163] with a four-story model but a cut-off of the decoder to get a output of 64x64

---

[2] https://github.com/Rignak/DeepLearning/tree/oceanix

pixels. The architecture is depicted in Fig. 3.2. The model is trained under the Weighted Binary Cross Entropy function defined as:

$$WBCE_c(y_c, \tilde{y}_c) = -\frac{K_c C}{\sum K} \cdot \left[\frac{\sum y_c}{64^2} y_c \log(\tilde{y}_c) + (1 - \frac{\sum y_c}{64^2})(1 - y_c) \log(1 - \tilde{y}_c) \right] \qquad \text{(Eq. 3.1)}$$

This loss is quite different as the one described in [73] as it uses the vector K which is the inverse of the pixel-wise *a priori* probability of each phenomena to appear (computed on the training set). C stands for the number of classes and is needed to conserve the same learning rate. y_c and $\tilde{y}_c$ are respectively the groundtruth and the predicted segmentation for the class c.

$$K = (34.5, 11.8, 250.0, 14.1, 8.2, 71.4, 2.3, 50.0, 9.1, 8.3) \qquad \text{(Eq. 3.2)}$$

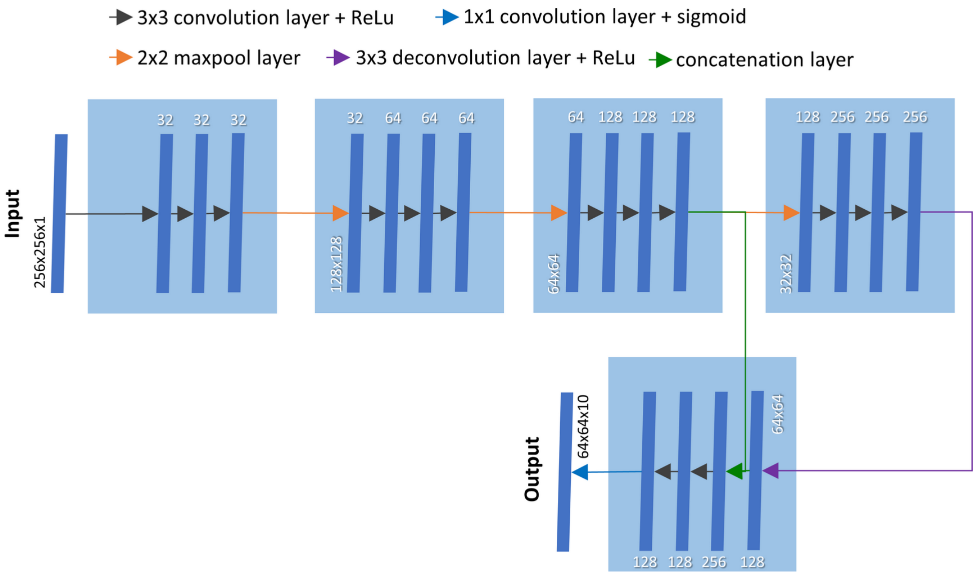


Fig. 3.2: Architecture of the fully-supervised model. It is a U-Net model [163] with a input shape fourfold the output to account for the imprecision of the annotation.

3.3.2 Masking the Input (MASK)

This method, as suggested by [224], assumes that the class-output of a categorizer will be affected negatively if a part of the researched object is hidden. Usually, this is done by moving a square mask over the input. With X the input, the masked input $X'_{c,a,b}$, where

a and b denote the position of the mask, and c its size, is defined by:

$$X'_{c,a,b}(x,y) = \begin{cases} X(x,y) & \text{if } |x-a| > \frac{c}{2} \text{ and } |y-b| > \frac{c}{2} \\ \\ v & \text{else,} \end{cases} \tag{Eq. 3.3}$$

with v is the value taken by the input when the mask is over it. As one of the TenGeoP-SARwv's class, namely the *Low Wind Area* class, can be recognized by an overall low intensity, setting $v = 0$ could lead to an increase of the categorization score for this class. Incidentally, it introduces high spatial frequencies that could, a priori, be seen as either an *Atmospheric Front* or an *ocean Front*. Setting v at the mean value of the image could mitigate this behaviour.

Once the edited inputs are generated, the pseudo-probability of the phenomena i to be depicted on the input X is denoted as $\mathcal{M}(X)_i$, with $\mathcal{M}$ a model trained on the categorization problem. The pseudo-segmentation is defined by:

$$Y(x,y) = \frac{1}{\#A_c(x) \cdot \#B_c(y)} \sum_{\substack{a \in A_c(x) \\ b \in B_c(y)}} \max(0, \mathcal{M}(X)(x,y) - \mathcal{M}(X'_{c,a,b})(x,y)) \tag{Eq. 3.4}$$

$$A_c(x) = \left\{ a \in \mathbf{N}, |x-a| < \frac{c}{2} \right\} \tag{Eq. 3.5}$$

$$B_c(y) = \left\{ b \in \mathbf{N}, |y-b| < \frac{c}{2} \right\} \tag{Eq. 3.6}$$

However, this computation has a complexity of $O(h \cdot w)O(\mathcal{M})$. To decrease it and be able to run it in an acceptable time, we use a stride s and compute $Y'(x,y)$ only when both indices are multiples of s. Thus, the complexity is reduced to $O(\frac{h \cdot w}{s^2})O(\mathcal{M})$.

In the experiments, $c = 75$ px (3.75 km) and $s = 25$ for a 512×512 px-wide input. The resolution of the pseudo-segmentation is thus of 16x16px and required to run the categorizer on $256 + 1$ images of shape $(512, 512)$.

3.3.3 Partial Categorization (PART)

Using partial categorization to build a pseudo-segmentation relies on the opposite paradigm. Following [204], the categorizer runs on multiple locations and only take a small part of the input at a time.

Some categorizers (such as InceptionV3 [187]) have a number of weights independent of the input size. Indeed, they do not include fully connected layers before a global pooling. Therefore the number of weights is only defined by (1) the number and the shape of the convolution kernels and (2) the number of classes, and of neurons of the fully connected classes. This property enables a categorizer trained for a shape S_1 to categorize inputs of shape S_2.

A minimal shape can be defined by the architecture of the network. In the case of InceptionV3, the last activation map shrink to a single pixel for an input shape of $(75, 75)$. In other words, an InceptionV3 trained on $(512, 512)$-input can provide categorization down to inputs of shape $(75, 75)$ (given that the resolution does not change). Therefore, it is possible to divide several tiles and obtain the categorization for each of them. The pseudo-segmentation is therefore defined by:

$$Y(x,y) = \mathcal{M}\left(X\left[x - \frac{S_2}{2} : x + \frac{S_2}{2}, y - \frac{S_2}{2} : y + \frac{S_2}{2}\right]\right) \tag{Eq. 3.7}$$

Fully-convolutional categorizers also possess the property of translation invariance: if the effect of the borders are considered as marginal, the output of the last convolutional layer, the activation map, will be equivariant with the input. However, as a spatial global average is computed after these convolutional layers, localized information will be lost and the equivariance will become an invariance. In the context of searching the position of a phenomenon, it means that running the categorizer on *every* possible location is unnecessary: the result on an image after a translation of a few pixels will not change.

In the experiments, the tile shape S_2 is either of 129×129 or 75×75, and the stride of $\frac{S_2}{3}$. As with the previous algorithm, a categorizer has to be run on dozens or hundreds of inputs (depending of the resolution) but these inputs are several times smaller than the full observation.

3.3.4 Class Activation Maps (CAM)

Even without doing assumption on the scale, shape or intensity of the SAR signatures of the metocean processes, it is possible to obtain a pseudo-segmentation. By essence, a convolutionnal network keeps the localization properties of the input image through its layers. The stride of the convolutions, the size of the kernel and the poolings build a filter system that recognizes shapes [105] and textures [52] and returns activation maps, i.e., the presence of the features recognized by the filters. However, a categorizer has to finally fuse the information of these activation map to obtain an image-level answer. This is usually done using global pooling [186] or dense layers [95, 173, 65]. The localization information is thus conserved until the last convolutional layer, whose outputs are called Class Activation Maps (or CAM). After that point, localization information is discarded by the use of either dense (also called fully connected) layers or by a global polling. It has been shown that it was possible to use CAM, before the destruction of structural information, to obtain the segmentation of the objects [233].

To get CAM from a categorizer, the global average pooling is removed from the network. Then, the output of the last convolution layer -of shape (w, h, f)- is multiplied by the weights of the dense layer -of shape (f, n)- to produce a (w, h, n) output, where w and h are respectively the width and height of activation maps, f the number of filters and n the number of classes. The biases are applied if available.

In SAR domain, CAM has been used for segmentation of crop areas under weak supervision on image-level labels [211]. To this end, a categorizer from U-Net architecture [163] is built by adding a global pooling layer after the U-Net module and a dense layer to translate the 32 channels of the output into a single value (the binary classification of the input). This classifier is called U-CAM and achieve a categorization accuracy of around 98%. However, this study was done with low-resolution pictures (50×50 pixels, with 200k samples at a resolution of 30 m.px^{-1}) with only two classes. In contrast, the TenGeoP-SARwv dataset contains images of 512×512 pixels, categorized in ten classes. Since the categorization task is more complex, the attempts to train a U-Net classifier did not reach satisfactory result (best accuracies at 40%).

However, since the shape of TenGeoP-SARwv's images is higher than InceptionV3's minimal shape, it is possible to use this architecture to build CAM. Similarly to the U-CAM network, this model ends with a global average pooling layer followed by a dense layer. Thus, it is feasable to use the activation map of the ante-penultimate layer to obtain segmentation of the metocean process. As it does not include unpooling layer, the activation maps have a lower space resolution than the input. For an input of shape $(512, 512)$, the activation maps have a size of $(14, 14)$, representing a spatial resolution of approximately 1.8 km.px^{-1}, whereas U-CAM could keep the same resolution as the input.

This algorithm remind those described in the previous section as it is based on the output of a convolutional layer previously to a pooling phase. However, the spatial propagation of information through the convolution is not clipped to the width of a tile. It can be computed that it is equal to $\frac{1255}{64} \approx 19.6$ pixels in the final activation map (superior to its own width on inputs of 512 × 512).

3.3.5 Image-Level Classes as Noisy Pixel-Level Segmentation (IM2PX)

Neural networks have been proved to be able to train under very noisy datasets [161]. As such, the image-level labels used in the weak-supervision can be interpreted as segmentations with heavy spatial noise. Under this paradigm, the $(1, 10)$ label vector is converted into an image of shape $(h, w, 10)$, the value of the vector being used for each pixel. Then, an image-to-image model is given the task to solve the segmentation problem.

Assuming that (1) the segmentation model respect constraints on the spatial information propagation (the value of an output pixel is defined by the value of the input pixels in a constrained neighbourhood), it is possible to infer implicit information on the position of the phenomena. In layman's term, if a pixel is located far from the features of a class, the segmenter (whose work is to do a pixel-level categorization) won't be able to assign this class to the pixel, as no pixels of the neighbourhood *should* (2) give information about it. Thus, if (3) the network has been trained without overfitting, the only pixels with a significative value for the categorization will be those under the neighbourhood of a phenomenon.

Condition (1) can be ensured by studying the model architecture. The spatial propagation of a classic autoencoder is $\rho = \frac{k}{2}(2^n(n+4)-2)$ with n the number of max pooling layer and k the number of successive convolutional layer (of kernel size 3) at each stage. This can be demonstrated by remarking that at the first layer of the stage i of the spatial propagation is $\rho_i = 2(\rho_{i-1} + \frac{k}{2})$, and that $\rho_0 = n\frac{k}{2} + k$ (this is the spatial propagation of the encoding network). To prevent ρ from rising, and for simplicity's sake, an encoder model was used with $n = 3$ and $k = 3$. Condition (2) depends on the comprehension of the phenomena. It is not impossible that long-range correlation could provide information on the main phenomena on the picture, but they are considered small enough to be marginal.

3.4 Results

3.4.1 Quantitative Results on the Wave-Mode Test Set

Quantitative results are obtained by computing the Dice index (also known as F1-score) over a test set of manually annotated segmentations. The Dice index is defined as:

$$D(x_i, y_i) = \frac{1}{\#S} \sum_{e \in S} \frac{2 \cdot x_{e,i} \cdot y_{e,i}}{x_{e,i} + y_{e,i}} \tag{Eq. 3.8}$$

In this equation, S is the set of all the pixels, $x_{e,i}$ is the predicted segmentation of the label i at the pixel e and $y_{e,i}$ the corresponding groundtruth. The Dice index gives values between 0 and 1, the latter being obtained on perfect agreement between the groundtruth and the prediction. Another (equivalent) formulation of the Dice index is as the harmonic mean of the precision and recall, which increase when respectively the false positive and false negative decrease. As such, the Dice index would always be equal to 0 in case of non-detection whereas the range of a mean square error or a cross entropy would depend on the surface covered by the phenomena.

These groundtruths are independent of the datasets used either in the weakly-supervised framework or in the fully supervised one. SAR observations, used as inputs of this test set, originate from a supplementary dataset to TenGeoP-SARwv and contains 100 manual segmentations (with 10 observations from each class).

According to Table 3.2, the supervised framework outperforms all the other methods. Compared with the other methods, it gives especially good results for the *Rain Cells*, the *Atmospheric Fronts*, the *ocean Fronts*, the *Pure Ocean Waves* and the *Icebergs*, with a Dice index of respectively 31.1%, 29.5%, 33.7%, 47.9% and 13.5% for four classes. Incidentally, it has the lowest inference time, requiring only to run a U-Net model a single time.

However, other methods have their merits on specific phenomena. In particular, partial categorizations are powerful to segment *Wind Streaks* and *Sea Ice*. Though not obtaining higher performances than the fully-supervised framework, they produces good results on *Pure Ocean Waves* with a 75 px-wide tile (40%). The Low Wind Area and Micro Convective Cells requiring enough contextual information to be discriminated, 129 px-wide tile is more adapted to phese phenomena, obtaining a Dice index of respectively 63.2% and 34.2%, the latter being the best segmentation over all methodologies.

Table 3.2: Dice index for each algorithm and phenomena computed over the test set. A Dice equal to 1 means perfect overlap between the truth and the prediction. Values are given as mean and standard deviation over 10 trainings to account for the random initialization and order of the groundtruths used in the epochs.

Method	**AF**	**BS**	**IC**	**LWA**	**MCC**	**OF**	**POW**	**RC**	**SI**	**WS**	**Mean**
A - Fully supervised	**29.5%** **[4.4%]**	24.7% [8.4%]	**13.5%** **[2.1%]**	69.5% [4.5%]	30.8% [7.4%]	**33.7%** **[6.2%]**	**47.9%** **[6.5%]**	**31.1%** **[3.5%]**	63.3% [11%]	61.1% [7.5%]	**40.5%** **[6.1%]**
B - MASK ($v = \mu$)	9.6% [2.1%]	16.1% [2.9%]	4.9% [0.9%]	15.2% [2.8%]	20.6% [3.7%]	18.9% [1.5%]	9.6% [2%]	13.1% [2.3%]	17.2% [4%]	27.2% [8.8%]	**15.2%** **[3.1%]**
B - MASK ($v = 0$)	10.4% [1.1%]	17.2% [1.6%]	1.2% [0.5%]	5.9% [1.7%]	3.28% [5%]	1.3% [1.2%]	14.8% [2%]	8.2% [1.6%]	19% [8.2%]	51.7% [11.1%]	**17.4%** **[3.4%]**
C - PART ($c = 75$, $s = 25$)	6.9% [1.3%]	**29%** **[1.2%]**	2% [0%]	39.2% [9.8%]	23.2% [3.4%]	6.7% [0.9%]	40% [4.9%]	19% [6.2%]	73.3% [4%]	66.9% [5.1%]	**30.6%** **[3.7%]**
C - PART ($c = 129$, $s = 43$)	14.6% [1.7%]	28.9% [0.9%]	2% [0.2%]	63.2% [2.6%]	**34.2%** **[6%]**	11.8% [1.5%]	19.2% [4.4%]	23.5% [0.6%]	**77.5%** **[3.5%]**	**70%** **[4.8%]**	**34.5%** **[2.6%]**
D - CAM	16.6% [3.4%]	14.8% [3.4%]	2.2% [0.3%]	47.6% [5.8%]	12% [3.2%]	20.4% [3.2%]	7.3% [1%]	17.9% [2.5%]	40% [9.4%]	14.2% [8%]	**19.3%** **[4%]**
E - IM2PX	12.1% [1.9%]	26.9% [1.8%]	1.9% [0.5%]	**69.6%** **[1.5%]**	26.5% [5.2%]	10.9% [1.1%]	17.5% [1.5%]	18.8% [3.7%]	69.8% [12.5%]	27.7% [5.6%]	**28.2%** **[3.5%]**

The noisy segmentation paradigm is competitive on *Sea Ice*. However, this particular class was defined, in the TenGeoP-SARwv dataset by the presence of sea ice over the whole observation, with no open water area. With this definition, the categorization annotation is effectively the same as the segmentation. This paradigm also obtains the best result for the segmentation of *Low Wind Area* but, given the standard deviation, should be considered equivalent to the fully-supervised framework.

Mask-based methods and Class Activation Maps, on the other hand, are inefficient on most classes. If we exclude the supervised framework, CAM perform well on *ocean Front* and *Atmospheric Fronts* whereas the Masking method outperformed the other unsupervised methods on *Icebergs*.

As the quality of the groundtruth in the fully supervised framework is higher, it is possible to obtain better results with much less annotations if they are at pixel-level, as depicted in Fig. 3.3. Assuming the performance follows a power law [72], the fully supervised framework could provide even better results. On the other hand, all the weakly supervised algorithms are converging to much higher loss values. An increase of the groundtruth quantity would not lead to enhanced performances.

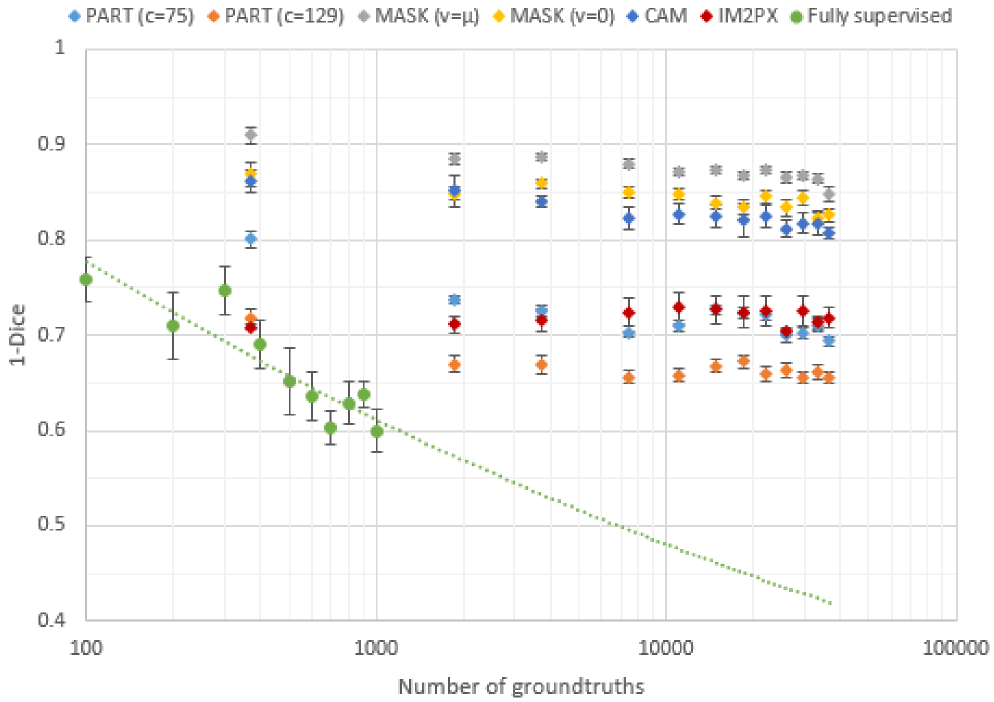


Fig. 3.3: Evolution of the Dice index on the test set in regards with the quantity of available groundtruth. The error bars correspond to the standard deviation over ten models (random initialization, random selection of groundtruths). The green doted curve correspond to a modelisation following [72].

3.4.2 Visual Inspection of Wave-Mode Results

The visual inspection of the segmented wave mode images are given in Fig. 3.4. It confirms that the most promising method is the fully supervised trained U-Net model. The segmentations are sharp enough to have a reliable accuracy on *Icebergs*, *Oceanic* and *Atmospheric Fronts* while retaining the ability to segment global phenomena such as *Pure Ocean Waves* and *Micro Convective Cells*. These observations are coherent with the Dice index values.

On the contrary, the partial categorization methods suffer from high noise since the area covered by the tiles are often too small to contain enough context. As visible in the Fig. 3.4, some tiles, scattered over the image, are estimated as *Wind Streaks*. It also generates confusion between the *ocean Fronts* (and the boundaries of the *Rain Cells*) and ice leads (and thus the overestimation of the *Sea Ice* class).

The noisy segmentation paradigm, as per the results compiled in Table 3.2, was supposed to give good results on the segmentation of *Sea Ice*, though the result are not especially good on the other phenomena (with the exception of *Low Wind Area*). However, the aforementioned examples show that this method leads to repeated overestimation of the Sea Ice class, which mitigate the relevance of this method.

The algorithms relying on the Class Activation Maps show capabilities to segment some phenomena, but are impacted by the high spatial propagation of data, as previously discussed. This lead to an overestimation of local phenomena (*Icebergs*, *Atmospheric Fronts*, *ocean Fronts*). CAM are also enduring difficulties when multiple phenomena are present, as the more prominent monopolize the activation. This explains the large black areas, which are portions of the observations where the predictions are below a threshold.

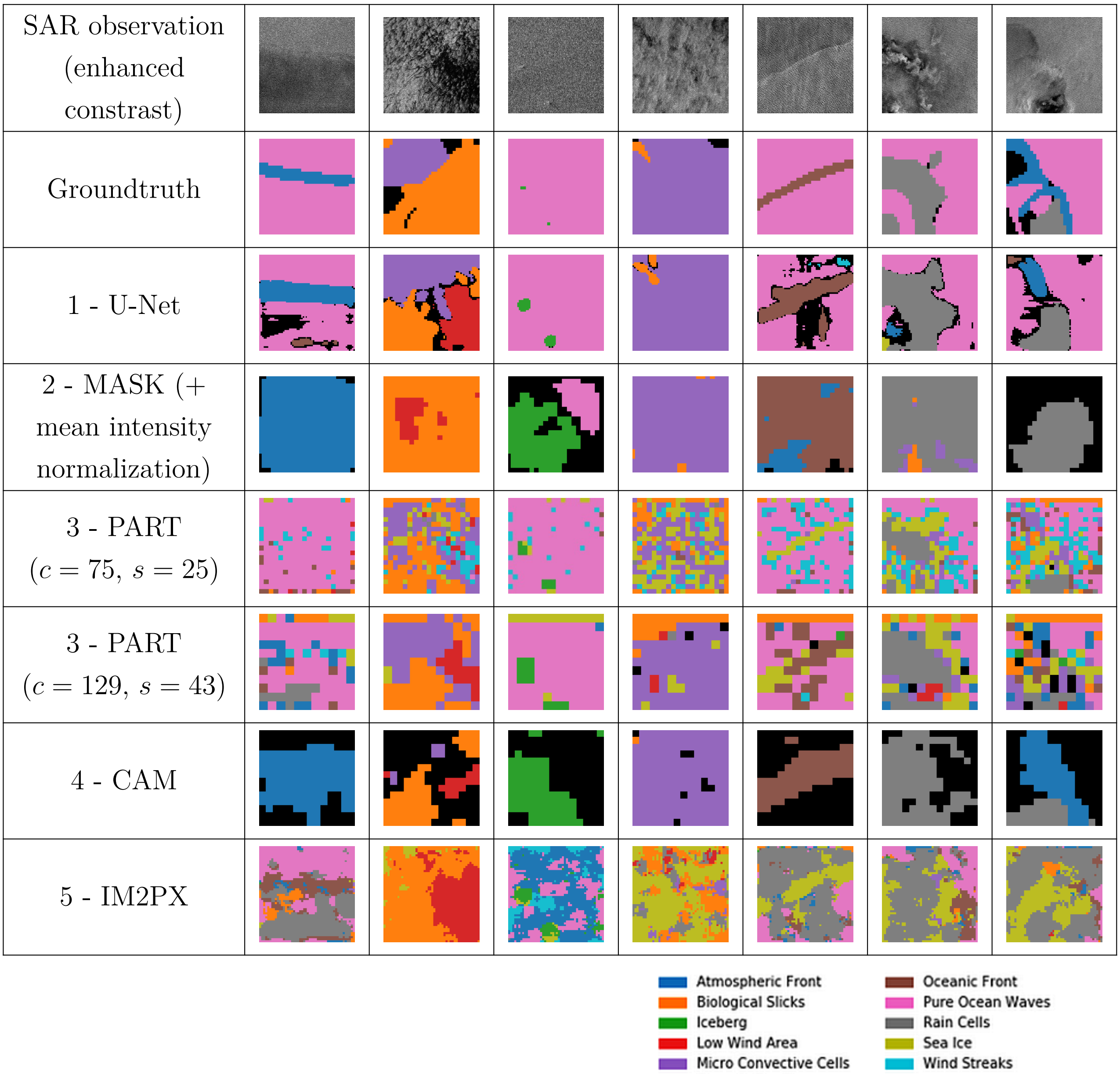


Fig. 3.4: Examples of output on elements of the test set for each segmentation method. Pixels in black are below a decision threshold of 0.5 and left undecided.

3.5 Discussion and Perspectives

All previous results rely on TenGeoP-SARwv dataset and same mode segmentation groundtruth, generated from the Sentinel-1 Wave Mode which is a 20x20 km acquisition mode. However, SAR observations can cover a much larger area whenever acquired in IW Mode, 250 km wide. They cover a wider incidence angle range, from 29 to 46 degrees (w.r.t 23 and 36 for WV) and are acquired over coastal regions. Though they differ from the Wave Mode images, Interferometric Wide Swath can be preprocessed in a similar way in terms of pixel size reduction and incidence angle dependent normalization.

Even though the occurrence of observable classes highly depends on the incidence angle [208], this information is not used in the segmentation process. The possibility to extend the developed methodology to a wider range of incidence angles thus appears realistic.

To this end, Wide Swath observations are divided in tiles of 20 km by 20 km. These mosaics are constructed by ensuring an overlap between each tile. With a construction stride equal to the half of the tile width, a pixel on the border of a tile will be contained in the central part of the next one. Taking advantage of the equivariant property of the fully convolutionnal networks such as U-Net, reconstruction of the mosaic with the central part of predicted tile provide continuity over the whole Wide Swath observation. The segmentation of areas of several hundred kilometers wide can be computed in a few seconds (depending of the available GPU) with the supervised method. However, weakly supervised algorithms relying on multiple categorization for one tile cannot be realistically used for such applications.

At first sight, Fig. 3.5, with a model trained under the fully supervised framework, shows promising results for the segmentation of Wide Swaths, with most metocean processes being well detected. The *Rain Cells* appearing in the south of the Tuscan Archipelago (in the yellow rectangle) are correctly returned despite their small scale. An Atmospheric Front can be observed at the south-east of Corsica (in the green rectangle). This front delimit two areas of different wind speed and continue in the south-west of the island. The area in the south, with more intense wind, contains *Wind Streaks* which are nicely highlighted. Ships are also appearing in several places of the observation and are returned as *Iceberg* as they have a similar signature. This behaviour reflect a contextual difference between the wide swaths acquisition (which are taken in coastal areas) and the wave modes -on which was trained the network- where the probability to have a ship is lower.

Other geographical differences include the orographic waves depicted in the east of Sardigna. They are induced by the western winds and the upwind topography. These orographic waves appear as large scale wind disturbances perpendicular to the wind flow. Such coastal interactions are absent of the TenGeoP-SARwv dataset. Given their similar aspect, they are mis-classified as successive atmospheric fronts. Their classification as a separate new phenomenom would probably require their investigation at a larger scale than 20x20 km to identify several wave patterns and the coast vicinity. It can also be noted that, in this area, orographic waves are co-occurring with *Wind Streaks.* This raises concerns about the implicit segmentation assumption that, at pixel-level, metocean processes are mutually exclusive.

The influence of the coast is also visible in the Gulf of Genoa. There, a different regime with strong North-Eastern offshore winds is influenced by the rapidly changing topography. It leads to a strong turbulent flow and some wind discontinuities aligned with the flow. It is segmented as *Rain Cells* though the visual inspection does not highlight precipitations.

Overall, even though the classes may not be adapted to all the phenomena met over these coastal regions, the algorithms behaves well in its ability to find the closest one among the possible choices. Introducing new classes, adding several scales of predictions can probably leviate the identified issues.

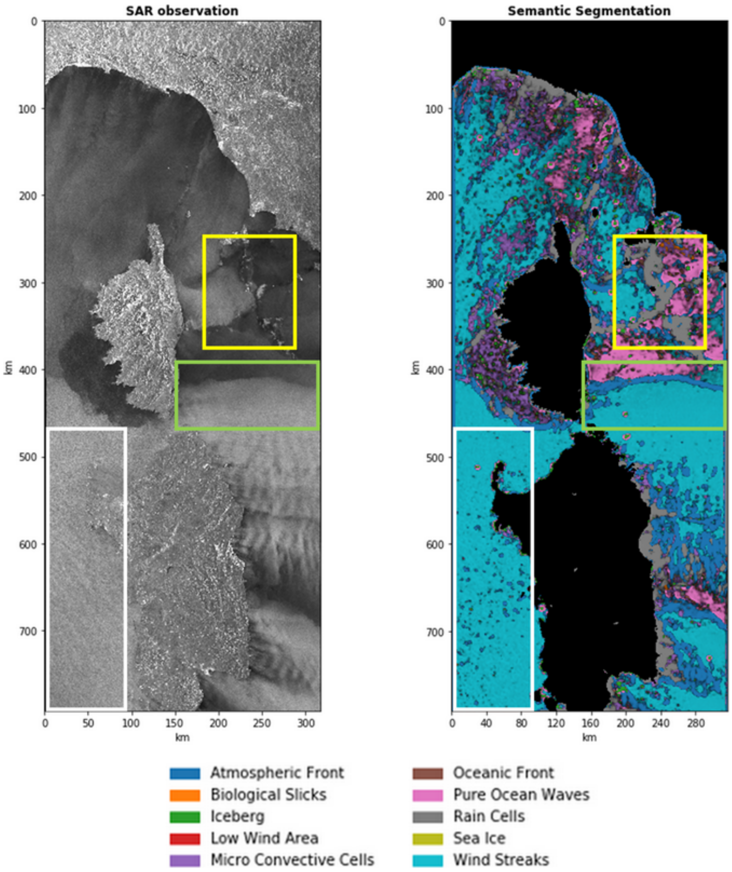


Fig. 3.5: Interfermetric Wide-swath acquired the 20 November 2020 at 05:28 and its segmentation with the fully-supervised method. To account for geographical considerations, a model excluding the *Sea Ice* class is used.

3.6 Conclusions

Fully supervised and weakly supervised segmentation are two paradigms under which it is possible to obtain segmentation of metocean processes. Fully supervised segmentation needs pixel-level annotations that are more time consuming to produce than image-level annotation used in weakly supervised frameworks. However, our experiments show that the data quantity required to outperform weakly supervised techniques can be reduced to a few dozens per class. More data leads to even better results. The extrapolation of our results suggests that building a fully groundtruthed dataset of dozen thousands segmented images may result in a significant gain in the segmentation performance.

The metocean phenomena studied in this work are open water processes contained in the TenGeoP-SARwv dataset. They are thus restricted to the ten classes described and suffer from several limitations mainly linked to the representativity of the data. First, as this dataset was created for categorization, the diversity of the phenomena are not as high as they could be in reality. In fact, the *Sea Ice* class mostly contains images wholly covered by ice, for which the categorization information is effectively equivalent to the segmentation. Not only the diversity could increase with the introduction of the sea/ice boudnary, but this class could also be divided in multiple subclasses [81] depending on the development stage or ice concentration. In the same way, *Rain Cells* typically occur with three components: a front, a down-burst area corresponding to the descending wind, and a smaller splash area generated by the precipitation. A more detailed analysis of this phenomena could be given by additional sources, for example weather radars.

The representativity of the ocean situation can be impeded by the instrument itself. The area covered by Wave Mode being limited to 20 km by 20 km, some phenomena, such as *Wind Streaks* with long wavelength are difficult to acquire. A multi-scale approach using IW groundtruths is believed to be beneficial to the semantic segmentation. The extension of the segmentation to coastal areas will also require the integration of new classes such a orographic waves, offshore winds, sea/land breezes or internal waves that do not appear on the Wave Mode acquisitions.

The categorization task assume that a wavemode can belong to only one class. The segmentation decrease this constraint by considering multiple areas assigned to different classes inside one image. Still, an assumption that the phenomena are mutually exclusive at the pixel level is made. If this supposition holds on the standardized TenGeoP-SARwv dataset, it becomes difficult to maintains it on Wide Swath observations as orographic waves or internal waves can be superimposed with *Wind Streaks*. A multi-label estimation (as opposed with “categorization”) would need a more precise delimitation of each phenomenon.

- END OF THE ARTICLE -

Chapter 4

THE PROBLEM OF VARIABLE SCALE PHENOMENA. EXAMPLE OF THE COLD POOLS

4.1 Introduction

Cold Pools are large convective systems (ranging from several kilometers to several dozen) associated with descending winds, visible on SAR observations [99]. They are key elements of moisture evolution [188, 100], inducing the generation of new cold pools on their borders [196]. However, their dynamics are not well understood, leading to activities dedicated to characterizing their associated wind speed, temperature difference, extent, and duration [42, 97]. Convective systems are notably the subject of the EUREC4A campaign [178].

On SAR images, Cold Pools have distinctive signatures as the vertical component of the wind speed is null at the sea surface. This causes the wind to radially disperse as the altitude decreases, resulting in a change of the wind direction or a modification of the wind speed. Interactions with the EUREC4A campaign [178] highlighted that the semantic segmentation, initially developed for individual 20x20 km WV observations in Chapter 3, was insufficient due to the difference in scale of the processes. In Fig. 4.1, four *Convective Cells* ranging in size from 864 km^2 to 12 km^2 are highlighted. As shown by the pink square, a model that uses the entire information in the 20x20 km area (i.e. a maximum receptive field) would still only be able to provide partial observations of the larger *Convective Cells*. Fig. 4.2 focuses on areas of 20x20 km and illustrates that it is impossible to recognize a convective process using WV-like images. However, the size of some *Convective Cells*, particularly the younger ones, can be small. Reducing the resolution to enable the segmenter to run on wider images means losing the ability to recognize the smallest *Convective Cells*.

Multiscale process is a known problem in the remote sensing domain and various

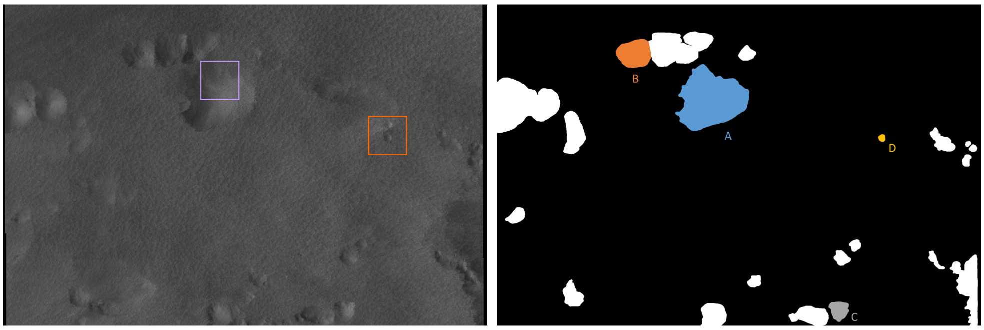


Fig. 4.1: On October 23rd, 2021 at 17:07:49, convective cells were observed (as shown on the left in the image). The corresponding segmentation is on the right. Cells A, B, C, and D cover 864 km^2, 221 km^2, 82 km^2, and 12 km^2 respectively. The purple and orange squares indicate the areas extracted in Fig. 4.2.

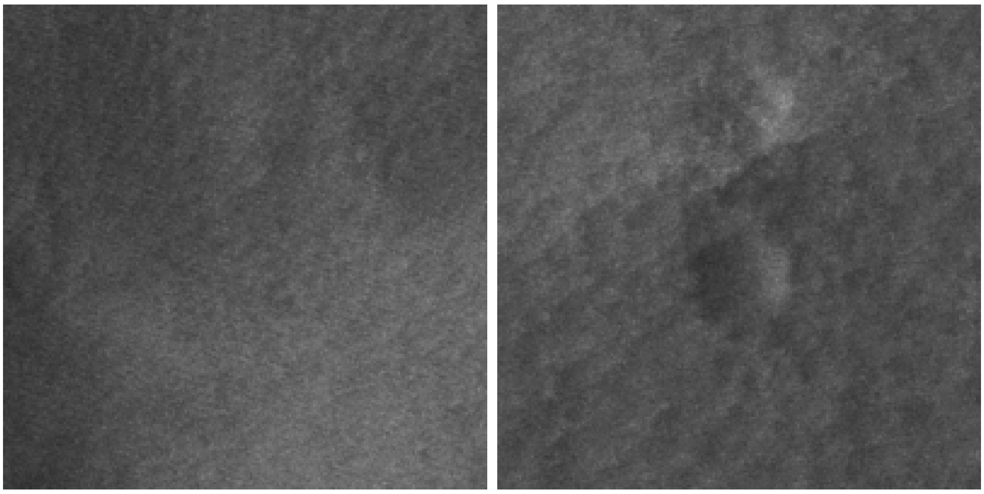


Fig. 4.2: 20x20 km tiles extracted from Fig. 4.1, centred on the top of *Cell* A (left) and on *Cell* D (right).

architectures have been developed to tackle this issue. Among the techniques used to enable a multiscale model, attention-based architectures [203] have been recently used [18]. Atrous convolutions [19] are also used to increase the receptive field and limit the number of weights [110]. Another method is to crop and resize to obtain samples of same pixel size but different spatial resolution [227, 112, 237]. As large IW observations (of several thousand pixels width at 100 m/px) are available from Sentinel-1 products, we design a U-Net architecture [163] augmented by "crop and resize" modules. Using a manually annotated convective dataset, the model is trained to segment both the content and the border of the *Convective Cells*. The model is compared to the original U-Net architecture at different scales. Collocations with IR sensors confirm its ability to detect cold pools at an early stage of development.

4.2 Dataset

The dataset is composed of 25 manually annotated IW at 100 m/px. Both the *Convective Cells* and the associated *Fronts* are annotated. The *Fronts* are important information as they help to differentiate *Cells* close to each other. However, as shown by Fig. 4.2, these *Fronts* are sometimes too fuzzy to locate and are therefore left without annotation. Out of the 25 acquisitions, 4 are used for testing, 4 for validation, and the remaining 17 images are left in the training set. The model only uses the VV-channel as input, as higher dimensional data would increase the risk of overfitting. During the training, patches of 20x20, 40x40, 80x80 and 160x160 km are extracted at random locations. The IW is mirrored on its borders to ensure that the patches will contain data rather than a fill value. Data augmentation is used by randomly applying vertical and/or horizontal symmetry and/or rotation of $\frac{\pi}{4}$, $\frac{\pi}{2}$ or $\frac{3\pi}{4}$ radians. A second data augmentation consists in a brightness adjustment by a random variable following the uniform distribution between 0.8 and 1.2. The brightness adjustment is only performed on the training set and does not affect neither the validation nor the test.

The dataset is available on kaggle[1].

4.3 Proposed Architecture

Since only IW and EW observation mode can be used to segment the *Convective Cells*, it is possible to take advantage of the large area covered and adapt the tile extraction scheme. Rather than extracting a single tile X of shape (w, h) at a resolution of r, for each index i of the architecture level, one tile X_i of shape (w, h) is extracted at a resolution $r \cdot 2^i$. These supplementary inputs pass convolution blocks and are cropped to be concatenated to the regular path. In the decoder, those secondary inputs are used to produce segmentation maps at each resolution. The activation maps before the final segmentation are also cropped and forwarded to the regular path. The layer connected to the secondary outputs but not to the regular path are discarded after the training similarly to the auxiliary outputs used, for example, in the Inception model [187].

---

[1] https://www.kaggle.com/datasets/rignak/convectivecells

Under this scheme, the effective field of the model can be larger than the size of the initial 20x20 km wide tile and conserve the continuity between adjacent tiles. The number of levels in the architecture can also be increased, though it increases the risk of overfitting. It should be noted that the continuity between the tiles is not explicitly enforced, but relies on the assumption that the intermediary outputs will act as regularizers.

Since the annotated processes, the *Fronts* and *Cells*, have different extents on the image, the loss function to minimize is balanced to their distribution. In particular, the *Fronts* cover only 3% of the observations. We consider the balanced mean square error described in Eq. 4.1.

$$MSE_b(\hat{y}, y) = 0.5\left(\frac{y}{k} + \frac{1-y}{1-k}\right) MSE(\hat{y}, y) \qquad \text{(Eq. 4.1)}$$

y is the ground truth segmentation, $\hat{y}$ the corresponding output of the model, and k the *a priori* probability of a pixel to be segmented, which is computed prior to the training on the training set. $\mathcal{L}$ is balanced in a way that if a segmentation only concerns a few pixels (typically in the case of *Fronts*), the weight of the MSE on these pixels will be high.

During the training, the formulation of Eq. 4.1 changes to retrieve the real-world distribution at the batch n, as indicated in Eq. 4.2. i is the index of the current batch, and n is the batch after which $\mathcal{L}$ becomes equal to the MSE.

$$\mathcal{L} = \frac{n - \min(i, n)}{n} MSE_b + \frac{\min(i, n)}{n} MSE \qquad \text{(Eq. 4.2)}$$

Finally, as $\mathcal{L}$ is computed at different resolution levels, the overall loss is a weighted mean of every loss, with the weights following a geometric progression of rate 0.5 so that the sum of the weights below a resolution level k would be equal to the weight at the resolution level k. The use of intermediary loss, for outputs discarded at inference time, is known to facilitate the training of very deep models [187] or constrain the final output layer [54].

4.4 Results

After training, the proposed multi-scale architecture is compared with single-scale U-Net models using the training set. The results on wide-swaths from the test set are described to qualitatively assess the segmentation and confirm lower border effects.

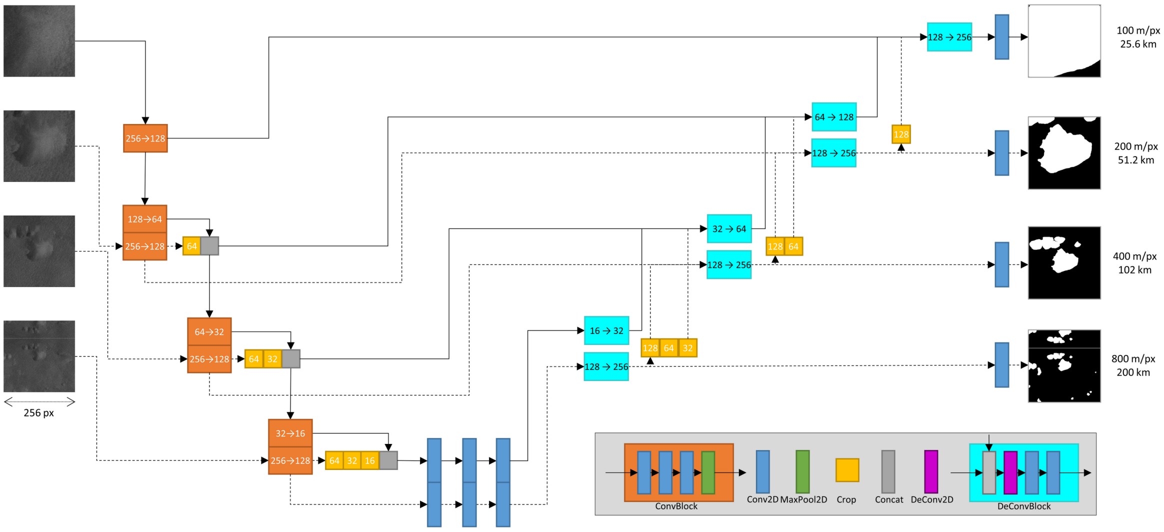


Fig. 4.3: Architecture of a multiscale segmenter. The dotted lines indicate the additional elements that turn a U-Net segmenter into a multiscale model. The numbers on the convolutional blocks represent the input and output matrix sizes. For the cropped layers, only the output size is indicated.

4.4.1 Quantitative comparison with single-scale models

The multi-scale CNN is compared with classical models of similar effective field, such as U-Net architectures with 5, 4, and 3 levels and resolution of 100 m/px, 200 m/px and 400 m/px, respectively. Three functions are used for the comparison: the Mean Squared Error MSE, the balanced MSE described in Eq. 4.1 and the F$_1$-score. The F$_1$-score is defined by computing a threshold α defined in Eq. 4.3.

$$\alpha = \underset{\tilde{\alpha}}{argmin} \quad P(f(x) > \tilde{\alpha}) < k \qquad \text{(Eq. 4.3)}$$

Since the segmentations are binarized, it is possible to compute the number of True Positives (TP), False Positives (FP), False Negatives (FN) and True Negatives (TN). The F$_1$-score is defined as in equation Eq. 4.4. The F$_1$ score can also be considered as the harmonic mean of the precision and the recall.

$$F_1 = \frac{2TP}{2TP + (FP + FN)} \qquad \text{(Eq. 4.4)}$$

The results of the tests are presented in Table 4.1. On the segmentation of the *Cells*, the multiscale model (with inputs starting from 100 m/px) is able to outperform the classical U-Net model at resolutions of 100 m/px, 200 m/px, or 400 m/px. The results are coherent for both the MSE and the balanced MSE (both of which are to be minimized) and the F_1-score (which is to maximize). On the other hand, on *Fronts*, the performances of the multiscale architecture are similar to that of the U-Net model at a resolution of 100 m/px, which has the same number of layers. This is consistent with the observation that convective *Fronts* can be recognized without the need for extended contextual information.

Loss	**Dataset**	**MS 100 m/px**	**U-Net 100 m/px**	**U-Net 200 m/px**	**U-Net 400 m/px**
MSE	*Cells*	★ 0.089	0.096 (0.0079)	0.103 (0.0061)	0.106 (0.0045)
	Fronts	★ 0.011	★ 0.011 (0.00079)	0.013 (0.00075)	0.012 (0.00045)
MSE_b (Eq. 4.1)	*Cells*	★ 0.42	0.46 (0.054)	0.50 (0.049)	0.49 (0.0098)
	Fronts	0.36	0.36 (0.028)	★ 0.35 (0.014)	0.37 (0.022)
F_1-score (Eq. 4.4)	*Cells*	★ 0.63	0.54 (0.055)	0.51 (0.055)	0.53 (0.029)
	Fronts	0.21	0.22 (0.0089)	★ 0.19 (0.0071)	★ 0.19 (0.0052)

Table 4.1: Quantitative results of the segmentation of convective processes. Results are given as mean and standard deviation over 5 trainings except for the multiscale model because of time consideration. ★ indicates the best method for each loss.

4.4.2 Qualitative comparison

The examples of the segmentations are presented in Fig. 4.4 (segmentation of the *Cells*) and Fig. 4.5 (segmentation of the *Fronts*). The border effect is visible on the segmentations from the U-Net model at 100 m/px as the prediction is subject to discontinuities at the edges of the patches. This issue, though still ocurring on the segmentations from the multiscale model, is mitigated. It can be noted that the segmentation of the *Cells* appears to loose its precision as the wind speed decreases, as visible in the bottom-right quarter of the observation from January 04th 2018.

4.4.3 Comparison with GOES measurements

Convective process can be observed with IR sensors such as the ABI aboard the GOES-16 and GOES-17 satellites [166, 106]. In particular, the 14th band, centred at 11.2 µm, [40] is sensitive to both clouds and water vapour. The latter is particularly associated with

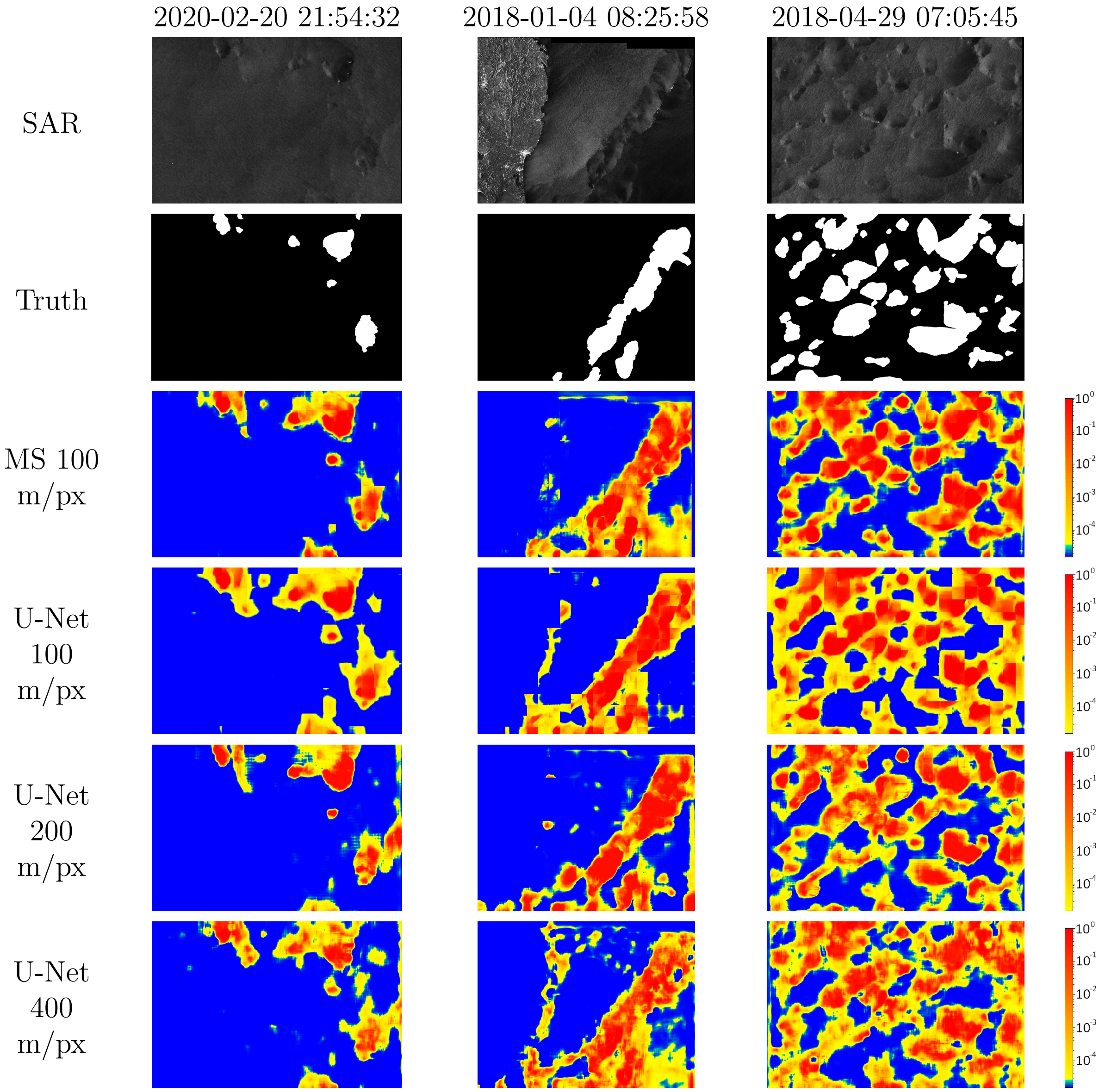


Fig. 4.4: Results of the segmentation of the *Convective Cells* on IW from the test set.

convection since cold pools are generated by rainfall evaporation [191]. SAR observations, and their subsequent segmentations as convective processes, have been co-located with the GOES-16 and GOES-17 satellites. Examples are displayed in Fig. 4.6. The advantage of such co-location is the possibility to observe the evolution of the *Convective Cells* since GOES-16/17 observations are available with a time resolution of 10 minutes. The spatial resolution for the 11.2 µm band is 2 km per pixel.

In Fig. 4.6.a, the correlation between the *Convective Cells* segmentation and a bright-

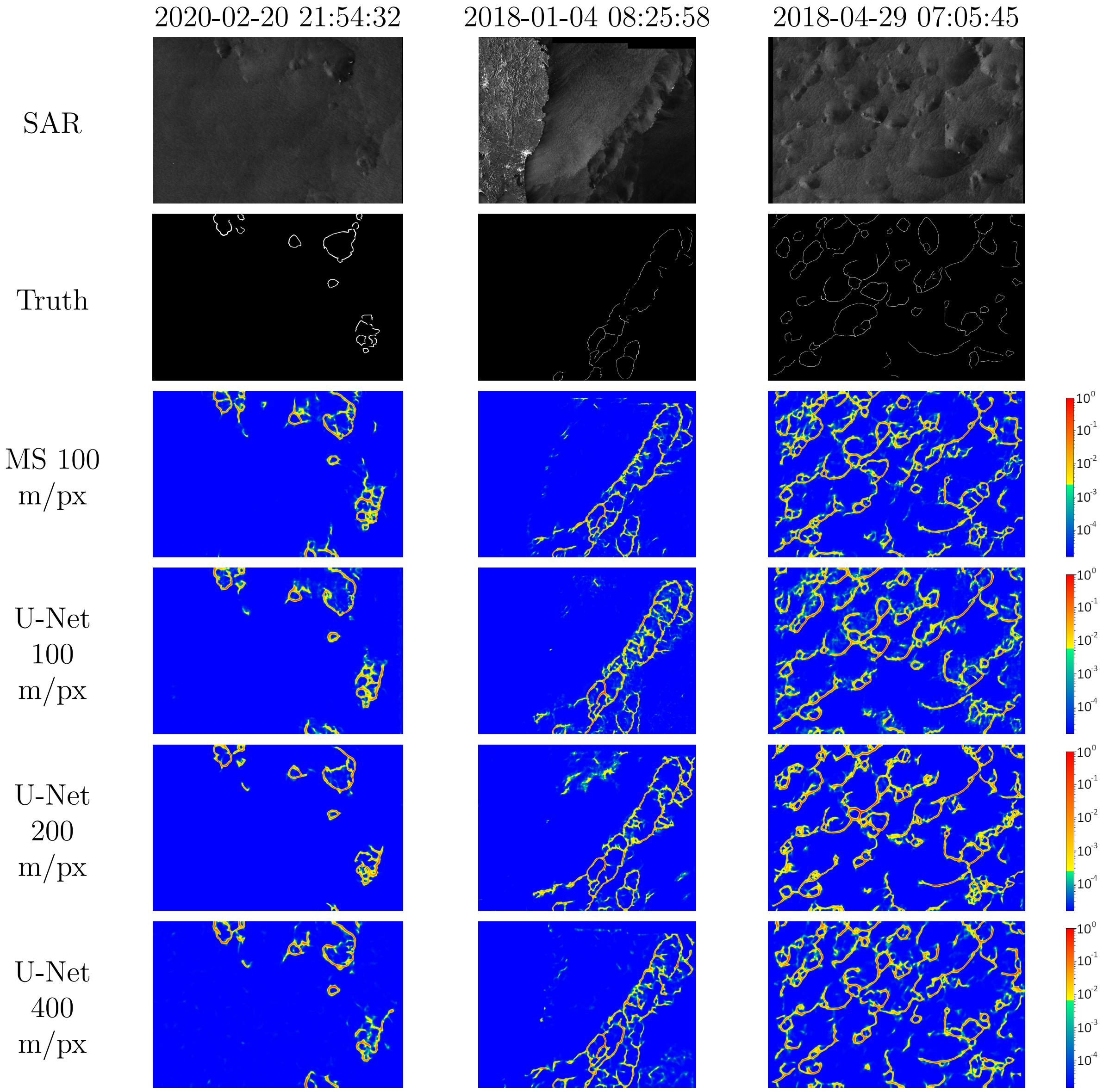


Fig. 4.5: Results of the segmentation of the *Fronts* on IW from the test set.

ness temperature around 280 K in GOES-16/17 Band 14 can be observed. In particular, *Cells* of various development stages are present. In the top-left corner, a large (and likely old) *Convective Cell* is dissipating. Though being easily visible at $T_0 - 4$, it has disappeared at $T_0 + 96$. In the green square, two adjacent *Cells* with strong rain signatures in their center are developing, reaching a maximum surface at $T_0 + 36$ and still being present at $T_0 + 136$ though having been separated into two isolated *Cells* around $T_0 + 56$. Interestingly, they are aligned with other small-scale *Cells* containing rain signatures. It should be noted that intersecting cold pools are known to last longer and grow larger and more intense than isolated ones [49]. In the orange square, a young *Cell* is developing into a large cold pool at $T_0 + 136$. Finally, the cyan square sees the birth of a strong *Cell*, with a low brightness temperature at $T_0 + 76$ in the trail of a *Cell* within which a rain signature was present in its back. This collocation indicates that the SAR observation, and its associated segmentation, can provide hints to determine the development of the convective processes.

As shown in Fig. 4.6.b, the 11.2 µm band is limited when at least two layers of cloud or water vapour are present. In this collocation, large cold pools are easily visible in the SAR observation. They are segmented as convective processes though their size, and the lack of gradient at the side opposed to the rain signatures, makes the model fail to delimit their whole extent. On observations from ABI's 14th band, these convective pools are hidden by an upper cloud layer moving toward the top-left. However, the existence of the second layer moving toward the top right is perceptible, especially at $T_0 + 40$, $T_0 + 60$ and $T_0 + 80$. It should be noted that the presence of rain signatures hinder the segmentation, as the upper *Cell* is only partially delimited. Furthermore, in Fig. 4.6.a, the *Cell* at the left of the blue square and those inside the green square, are not detected either. This is likely to be a bias in the dataset as most of the groundtruth *Cells* does not contain rain signatures. This model is also impacted by the lower wind speed of these two collocations in comparison to the wind speed distribution of the dataset.

4.5 Conclusion

As SAR observations of ocean surfaces are sensitive to variations in the wind, both for its direction and its speed, they can be used to study convective processes. The very high resolution of Sentinel-1 sensors enable to image even fine-scale processes, though the acquisitions used in this study are downscaled to a resolution of 100 m/px. With the appropriate dataset, it is possible to train a deep learning to segment the cold pools, and delimit their *Fronts*. However, the manual annotation is often difficult since the borders of the *Cells* can be too fuzzy to be identified. Besides, the *Cells* have a wide range of extent, notably depending on their age. To account for the specifics of these convective processes, we proposed a multi-scale architecture able to segment large *Convective Cells* without being subject to border effect or losing the ability to detect fine scale (younger) *Cells*. The results of the segmentation are compared to the manually annotated test set

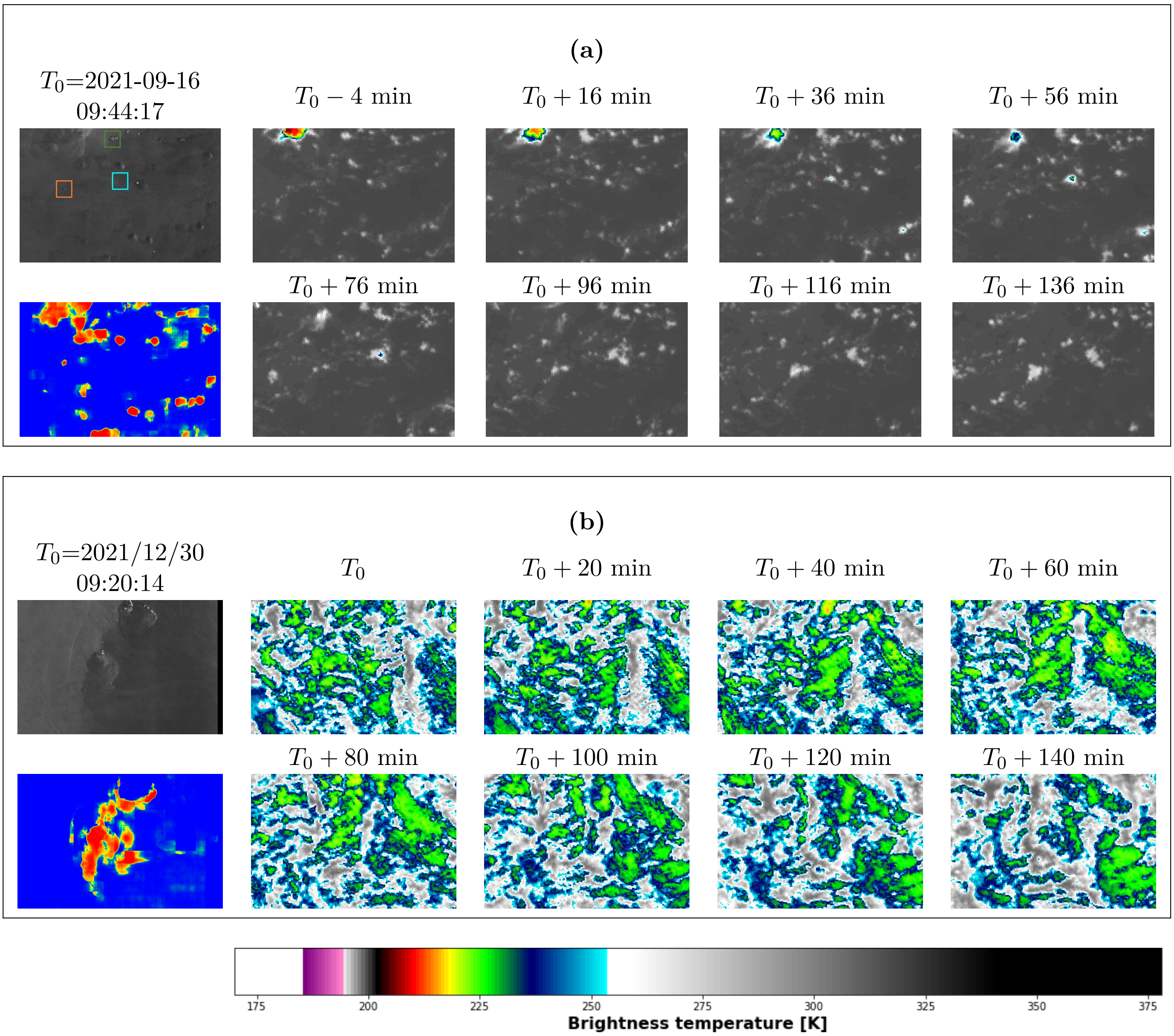


Fig. 4.6: In each box, the top-left image is the SAR observation. Bottom-left is the output of the convective cell segmenter. The other observations are the Brightness Temperature estimated from GOES-16 Advanced Baseline Imager (ABI) 14th band (11.2 µm), which indicates water vapor in the atmosphere.

and compared to the observation from GOES-16 and GOES-17 IR Longwave Band. The main limitation of the segmentation of the convective process appear to be the size of the dataset, which is too small to represent the whole diversity of cold pools. In particular, we identified lower performances at low wind speed and in presence of rain cells.

Part III

Collocation-based learning schemes for Earth Observation

FOREWORD

In the previous chapter, segmentation was performed to retrieve purely semantic processes. Groundtruths were obtained through manual annotation, although the underlying processes could be expressed with metocean properties, such as a wind direction discontinuity for convective processes or a variation in surface viscosity for biological slicks. These properties are difficult to retrieve automatically, leading to a need for manual annotations and small datasets.

In the following chapter, groundtruths are obtained from third-party sources. In the first section, sea ice concentration is studied. Data is manually annotated by the Norwegian Meteorological Institute, but their sheer size dwarfs previous datasets. The other two subjects are rainfall estimation, obtained through third-party sensors (NOAA's NEXRAD) and the estimation of wind speed when SAR observations are contaminated by rain signatures. In this case, groundtruths are obtained by colocalizing with an atmospheric model.

Thanks to the difference in data quantity, collocated datasets contain more diverse metocean situations, decreasing the risk of overfitting. It also makes it possible to study the behavior of the models under several conditions, which is only possible if enough samples are available for a specific metocean situation.

However, to ensure the representativeness of the dataset, one has to be careful to correctly preprocess the data and especially avoid data contamination and balance the training, validation, and test sets with care.

Chapter 5

SEA ICE CONCENTRATION ESTIMATION USING METNO/SENTINEL-1 COLLOCATIONS

This section is the reproduction of [54], in which the author of this manuscript was involved as second author. His contributions were directed at enhancing the model to take into account the specifics of the sea ice concentration estimation problem and, in particular, the relation between the class-oriented segmentation, justified by the groundtruth's format, and its translation as a continuous value of sea ice concentration. This led to the formulation of both a regression and categorization loss and two different output layers in the model architecture.

5.1 Introduction

In the context of climate change, the development of automated algorithms for ocean and sea ice monitoring has become a main activity for the scientific community. Given the continuation of many Earth Observation (EO) programs as well as the proliferation of sensors with increased capabilities, EO satellites are widely used to derive sea ice maps. Several types of sensors are used to estimate sea ice information. First of all, visible and infrared sensors such as the Moderate Resolution Imaging Spectroradiometer (MODIS) are often used because of their ability to image details on sea ice and its albedo, which is important for the study of energy exchanges between the ocean and the atmosphere [62]. However, sunlight and meteorological conditions over the poles make the generation of continuous observations from optical sensors very complicated. Thus, microwave sensors (both passive and active measurements) are used. Passive MicroWave radiometers (PMW) such as the Special Sensor Microwave Imager / Sounder (SSMIS) or the Advanced Microwave Scanning Radiometer (AMSR) have been the most extensive sources of information on sea ice since 1972 [30]. Therefore, it is generally integrated by sea ice services for automatic sea ice chart generation [194]. Concerning active sensors, scatterometers like QuikSCAT or ASCAT complement the data coming from passive microwave sensors. Indeed, scatterometers enable better precision on melting sea ice during spring and summer seasons than passive microwave sensors [160]. However, the main drawback of radiometers and scatterometers is their coarse spatial resolutions, in the order of tens of kilometers. Synthetic Aperture Radar (SAR) sensors appear to be effective because of their high spatial resolution (as of 2020, a few tens of meters for swaths of a few hundred kilometers), their sensitivity to sea ice development stages, and finally their independence from cloud cover or light conditions.

SAR is now commonly used for sea ice classification, being a primary information source for operational sea ice services. Sea ice charts made by human analysts are the results of the integration of SAR images combined with multiple information sources at different scales. Because of the time required to manually analyze these images, the profusion of images available from different sources, and the short-term prospect of constellations of micro-satellites, many methodologies have been proposed to automate the production of sea ice charts from SAR images. Most of them are based on the consensus that dual-polarized HH/HV data provide the best classification capability [149, 74, 88].

Before the advent of Deep Learning (DL), SAR-based sea ice classifications were generally based on a two-step method: First, a feature extraction is made, and then a regression or classification model is applied to retrieve sea ice information. The feature extraction can be based on texture analysis with different methods such as Markov Random Fields (MRFs) [143], Gabor filters, and Gray-Level Co-occurrence Matrices (GLCMs). From [25, 24], the discrimination between different ice types seems to be better achieved with GLCMs, then with Gabor filters and finally with MRFs. However, MRFs could also be used as a model for image segmentation [37]. From these extracted features, various methods could be employed to retrieve information on sea ice: Support Vector Machine (SVM) [107, 221], Multi-Layer Perceptron (MLP) [88, 87], Bayesian techniques, and regression techniques are examples of such methods, among others [120]. These two-step methods take a non-negligible time to develop as it requires "hand-crafted" features which are hard to select and sometimes not optimal when the methodology to select them uses different statistical properties than the classifier itself.

In the field of EO, the first use of DL with EO imagery dates back to 2012-2013 [201, 218], with a strong acceleration over the last years [235]. In [236], an exhaustive review of DL for SAR imagery is given with more than 200 references. Convolutional Neural Network (CNN) is, for now, the most common DL architecture used in computer vision and EO remote sensing

CNN methodologies can learn low- and high-level features from "raw" images. CNNs consist of a stacking of convolution and pooling layers. Convolution layers are composed of filters that move through the image to obtain convolved images. Each filter extracts different features in the image. The number and size of filters are different depending on the architecture. This convolution stage is an affine transformation. It is followed by an activation function (also called detector stage) to introduce non-linearity into the model, which enables the model to learn more complex functions [59]. Then, pooling layers aim to down-sample convolved images in order to reduce redundancy and the resolution of extracted features. Proceeding in the deepest part of the model, the feature extraction is done at different scales and resolutions thanks to pooling layers. This alternation of convolution and pooling layers performs as a feature extractor. This constitutes the so-called "encoder phase". A fully connected layer is then used to assemble and classify these different extracted features (so-called latent space).

Regarding sea ice applications, a set of studies have shown promising results using deep CNNs over dual-polarization SAR data. In [210] and [209], Sea Ice Concentrations (SIC) are estimated during the melting season using a CNN trained over 25 Radarsat-2 dual-pol (HH/HV) images. A patch size corresponding to 18 x 18 km ground distance is used. In [14], a CNN methodology is applied to Sentinel-1 dual-polarized HH/HV data for the classification of sea ice types in four classes (ice-free, young ice, first-year ice, and old ice). Two CNNs are trained over two subsets of 44 and 255 SAR scenes acquired in January-March 2018 and January-February 2020, respectively. The ground truth is extracted from the sea ice charts provided by the National Ice Center (NIC). These studies use an "encoder+classifier" approach, providing a 2-dimensional segmentation by locally applying the classification to the entire SAR image with a given stride. It is shown that a 2 x 2 km patch size provides better results than a 1 x 1 km patch.

In a Fully Convolution Network (FCN), the classical encoder step is followed by a decoder phase that up-samples the image to restore the spatial properties of the predictions until it recovers the original input image size. These up-sample operations are done by layers called deconvolutions. FCNs seem to surpass other methods for SAR image segmentation for multiple applications [141, 68, 214]. A comparative study for the classification of river ice [174] shows that FCN-like methods (DenseNet, U-net, DeepLab) provide very good results. U-Net [163] gives the best compromise between classification performance and generalization capability. The use of more advanced FCN architectures (e.g. Unet++ [234] and DeepLab v3 [19]) is certainly of interest but is out of the scope of this study.

In Table 5.1, a comparative evaluation of recent studies about automatic sea ice mapping from SAR images is provided. In this study, dual-pol Sentinel-1 images are used to estimate sea ice concentration with the use of an FCN technique. An adequate configuration of the U-Net architecture and a specific study on loss functions are the primary advances with respect to the other studies. We aim to provide a single but robust classifier able to provide reliable SIC maps over the whole year thanks to an unprecedented and massive database.

Table 5.1: An overview of recent studies on automatic sea ice mapping using SAR images and Deep Learning framework

Study	Sea ice Product Spatial resolution	ROI	Time	Dataset	Methodology	Accuracy	Major contribution (+) Major limitations (-)	Perspectives
Wang et al. 2017 [209]	SIC regression 400 m	Gulf of St. Lawrence Canada	17 Jan. to 10 Feb. 2014.	CIS sea ice charts Dual-pol. RS2 (25 scenes)	CNN input 18×18 km^2 400 m pixel input	E_{L1} 0.13 E_{rmse} 0.22	First complete DL methodology (+) Limited database/ROI/time (-)	Add windspeed information
Kruk et al. 2020 [96]	Sea Ice Type classification 5 km	Hudson Bay Canadian Territory	June-Dec. 2018	CIS sea ice charts Dual-pol. RS2 (350 scenes)	DenseNet-121 U-Net (encoder only) input 5×5 km^2	91.8% OA 3 classes - ice free, New Ice, FYI	No conclusion on summer melt period (-)	Up to 14 classes, expand the ROI and time of year
Cooke et al. 2019 [31]	SIC regression 3.21 km	Canadian Arctic Archipelago	Jan. 2010 to Sept. 2011	SIC from AMSR-E Dual-pol. RS2 (42 scenes)	CNN DenseNet input 3.21×3.21 km^2 with no overlap	E_{L1} from 0.2 to 0.3	Test scores over a limited number of scenes (-)	Larger patch size for better incorporation of spatial context
Malmgren-Hansen et al. 2020 [128]	SIC classification 2 km	Greenland areas	2014-2017 over all months	DMI sea ice chart Dual-pol. EW S-1 (912) AMSR2	CNN input 12×12 km^2 2 km overlap	R^2 0.8 - 0.89 with AMSR2 11 classes - ice free, SIC with 10% increment	Add of AMSR2 (+) Evaluation based on mean prediction within each of the test polygons (-)	Larger patch (higher dilation rate in pooling layer, coarser input data), noise correction
Boulze et al. 2020 [147]	Sea Ice Type classification 2 km	From Canadian Archipelago to Franz Josef Land	Jan. - March 2018 Jan.-Feb. 2020	NIC sea ice chart Dual-pol. EW S-1 (44 in 2018, 255 in 2020)	CNN input 1×1 or 2× km^2 (best)	90.5% OA in 2018, 91.6% OA in 2018 4 classes - ice free, Young Ice, FYI, Old Ice	Sensitivity experiments (+) Improved noise correction [148] (+) Limited dataset Jan. - Mar.	
This study	SIC classification 200 m resolution	from Greenland Sea to Kara Sea	June 2016 to May 2019 over all months	MET Norway sea ice chart Dual-pol. EW S-1 (1320) Wind speed	FCN input 89.6 × 89.6 km^2 200 m pixel input	78.2% OA 6 classes SIC	New auxiliary loss large dataset assessment for freezing or melting periods high resolution SIC products (+) Heterogeneous dataset with noise (-)	Improved noise correction homogeneous IPF for S-1 data Test advanced architectures [19] [234]

Section 5.2 presents the area of interest, the SAR database and the collocated sea ice charts. In Section 5.3, the employed methodology is described. A comprehensive evaluation of the performances is shown in Section 5.4. The Conclusion appears in Section 5.5.

5.2 Study Area and Data

The study area covers some parts of the Arctic Ocean, extending from the Greenland Sea to the Kara Sea ($80°W$-$80°E$; $60°N$-$85°N$). According to [153], these areas have a well-defined annual cycle with a sinusoidal temporal evolution. The Barents and Greenland Seas have an important ice cover dynamic between summer and winter, with potentially a completely open sea in summer. For the Kara Sea, freeze-up begins rapidly in late fall and early winter for the entire region, with melting in early summer. North of Svalbard or Franz Joseph islands, the area is almost completely frozen through the year. While the Kara and Barents Seas are not covered with multi-year ice, the western part of the Greenland Sea up to Fram Strait from $70°N$ is covered with this ice type as early as October (see products generated from [219]).

Our study is based on the regional sea ice products generated by the MET Norway and made available as part of CMEMS Sea Ice Thematic Assembly Centre (SI TAC). As mentioned in [39][82], sea ice concentration charts are generated from manual interpretation of EO data, with since 2014 mainly Sentinel-1 SAR data co-analyzed with visible and InfraRed (IR) data from NOAA VIIRS, MODIS and AVHRR. The Sentinel-1 mission provides C-band dual-polarized (HH and HV) SAR images. For the sake of coverage optimization, all the scenes are acquired in the Extra Wide mode with a 410 km swath width. As referred to in the previous Section, both channels are used by the analysts. Sea ice charts are produced every weekday with the latest available EO scenes acquired in the morning acquisition passes with a delivery in the early afternoon. These data actually contain 6 concentration classes with the following intervals: fast ice 100%, very close drift ice: 90 - 100%, close drift ice: 70-80%, open drift ice: 40-60%, very open drift Ice: 10-30%, open water: $<$ 10%. The full open water with SIC equal to 0 is also provided. For this study, the first 2 classes are merged because 1) the objective is to provide the classification of SIC and not of the sea ice type, and 2) the occurrence of the fast ice class is very limited. This last point is discussed later. These ice charts are considered as ground truth for the supervised learning in this study.

For the validation of the results, our estimates are compared to other data coming from the Ocean and Sea Ice Satellite Application Facility (OSISAF) from the European Organisation for the Exploitation of Meteorological Satellites (EUMETSAT). The product series has been operational since 2005 (version OSI-401-b since 26/05/2016). The OSI-401-b sea ice concentration is retrieved from the SSMIS sensor and provided daily on a 10 km polar stereographic grid. It is based on the combination of two algorithms [192] based on brightness temperature at the 19 and 37 GHz frequency channels ($T_{37v} - T_{19v}$ for the Bootstrap algorithm [29]) and $T_{37h} - T_{37v} - T_{19v}$ for the Bristol algorithm [175]. The validation of OSI-401-b products against operational sea ice charts from the Danish Meteorological Institute (DMI) for the Northern Hemisphere shows a total standard deviation of 15% (up to 20% during the summer season in June) and a negative bias of 5%

(up to 10-20% from May to July) [193]. OSISAF ice concentrations is systematically lower than the DMI ice charts especially at intermediate concentrations. The use of ARTIST Sea Ice (ASI) sea ice concentration algorithm [177] with AMSR2 data [131] may represent an alternative source of validation with higher resolution (3.125 km).

This study aims to obtain the most robust classification model for different ice conditions from S-1 dual-polarized images. Indeed, the ice condition varies throughout the year and its C-band radar backscattering ice may have a large variability from the ice growth process up to the melting season. In this study, 1320 S-1 EW scenes are collected from the Copernicus data hub over the area of interest from June 2016 to May 2019. These data are evenly distributed over the 12 months of the year (110 images per month) and cover at least some sea ice areas (based on OSISAF-based analysis). Among these images, 70 images per month are kept for the training phase. The remaining (40x12) is used for the final assessment. The Ground Range Detected Medium resolution (GRDM) product type is used: the pixel spacing of these products is 40 meters. For the EW acquisition, the Terrain Observation by Progressive Scan (TOPS) mode operates with 5 sub-swaths which have distinct properties in terms of radiometric calibration and thermal noise contamination (measured by the Noise Equivalent Sigma Zero - NESZ). The incidence angle ranges from 18.9^o to 47^o.

5.3 Methodology

The general methodology used in this experiment is presented in Fig. 5.1.

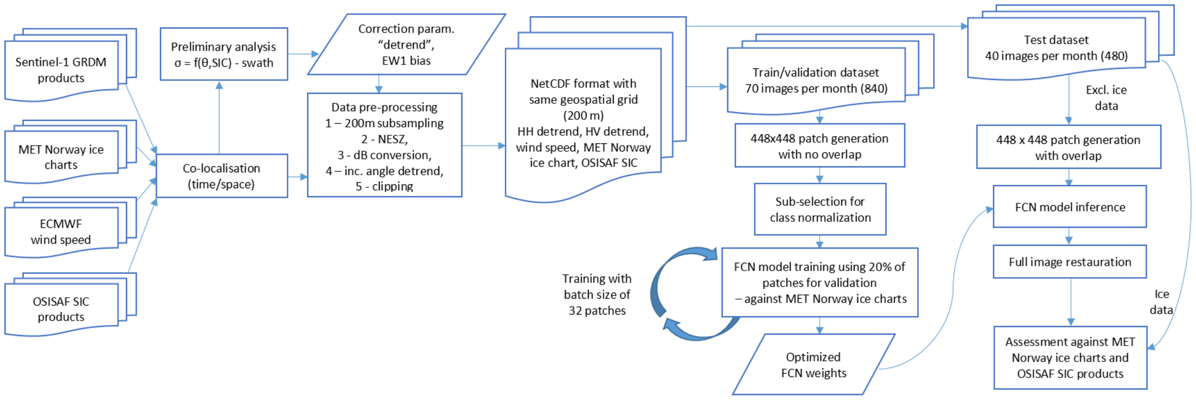


Fig. 5.1: General methodology for this study

5.3.1 Data pre-processing

On the co- and cross-polarized SAR data With the ambition to mimic the work of an analyst who does not necessarily use the highest possible resolution but speckle-noise polluted information, the SAR products are degraded to a resolution of 200 meters. In addition, FCN training is limited by available Graphics Processing Units (GPU) memory, which depends on the batch size (with impact on the convergence of the training process) and patch size. Reducing the resolution helps to maintain a sufficient batch size with patch size representing a large area on the ground (here 448 × 448 pixels representing $89.6 \times 89.6\ \text{km}^2$). Without this additional sub-sampling, the network would have a reduced patch size of $19.2 \times 19.2\ \text{km}^2$. The general context in the order of several tens of kilometers can be then used for a single patch. The training and validation sets are composed of 840 EW observations divided in 38644 patches of 448 by 448 pixels. The patches are later selected so that the percentage of data for each class would be more or less even, leaving 15495 samples in the training set, and 3851 in the validation set. The test set is composed of 480 EW observations.

A land mask based on the Global Self-consistent, Hierarchical, High-resolution Geography (GSHHG) database is also applied to select only ocean pixels. A noise correction is applied to the calibrated Normalized Radar Cross Section (NRCS) by simply subtracting the NESZ. The data for the co- and cross-polarized channels are then converted into decibels. As mentioned in Section 5.2, the methodology is built on a database of Level-1 EW products acquired from 2016 to 2019, thus processed with various Instrument Processing Facility (IPF) versions. Before the IPF version v2.9 [available from 16/01/2018], we note the NESZ annotation is only dependent on the range position.

There is a noteworthy dependence of co-polarized NRCS (and to a lesser extent cross-polarized NRCS) with respect to incidence angle for sea ice and open ocean (see Fig. 5.2). At least three approaches remain possible:

(1) The FCN network could eventually manage to handle this dependence by itself solely with radiometric information.

(2) The incidence angle could be ingested as an input layer with bi-dimensional information of the same size as the input patches. This approach does not seem fully optimal because the incidence angle does not contain any textural information potentially extracted by the encoding phase. Hence, the incidence angle could be concatenated between the encoding and decoding phases, thus at lower resolution.

(3) Each NRCS input layer could be normalized (or “detrended”) with respect to its averaged dependence of the incidence angle.

In this study, option (3) has been considered. This dependence is known to vary according to sea ice condition (*inter alia* SIC or First/Multi Year Ice [127]). The corresponding variation of $\sigma^0 = f(\theta, 50\%)$ with θ the incidence angle are thus calculated for both HH and HV channels and for the reference/neutral SIC value of 50%, approximated using a polynomial function (2nd order), and then subtracted from the data (see Fig. 5.2). For the HV channel contaminated by residual effects at sub-swath borders (due to the errors of NESZ estimation), only the pixels around the center of each sub-swath-center (±1°) are kept. Note that for the cross-polarized HV data, a calibration bias for the first subswath (hereafter denoted by EW1) has been observed (see Fig. 5.2 top right - incidence angle up to about 28°): an histogram matching process with interpolation splines is applied to correct this effect beforehand.

For each channel, a fixed clipping at 3-97 % determined over the “detrended” HV and HH database is finally applied in order to scale the values between 0 and 1. This overall approach helps to reduce the dynamic of the input data (see Fig. 5.2). As a consequence, the quantification step size between 0 and 1 is increased, maximizing the extraction of textural information. The impact of this pre-processing step has been assessed and found essential (+7.3% for the Overall Accuracy with respect to option (1) above). It is expected that option (2) would give similar overall performances to option (3). This analysis is left to another study.

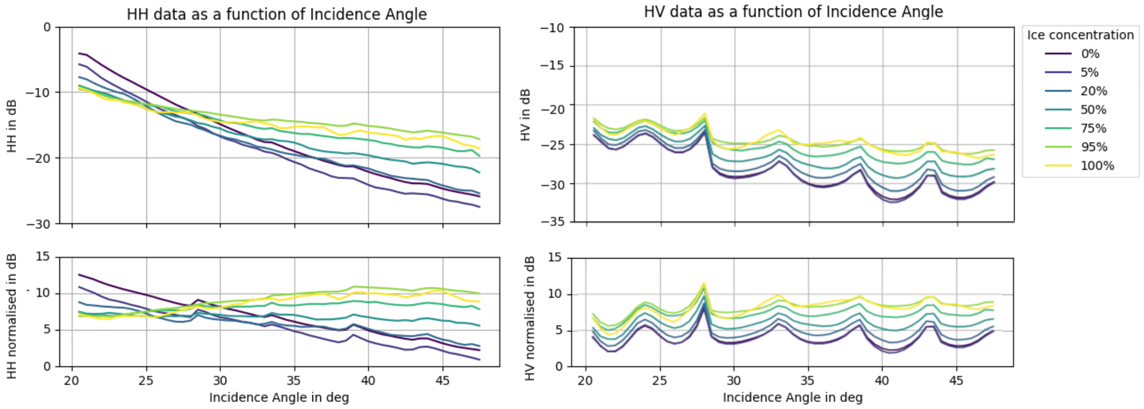


Fig. 5.2: Average NRCS values from the dataset for various SIC values: before (top) and after (bottom) data normalization for HH (Left) and HV channels (Right)

On the ancillary wind and sea ice data Over the open ocean, the NRCS primarily depends on the incidence angle but also on the wind speed and its relative direction with respect to the emitting electromagnetic waves (as modelled by the so-called Geophysical Model Function (GMF) or CMOD (C-band MODel) when applied to C-band scatterometer (e.g. CMOD 4 [179], CMOD5n [70]). In the Marginal Ice Zone (MIZ), the discrimination of ice and water can be made difficult depending on wind conditions [74, 56]. Some sea ice types such as frost flowers or pancake ice have NRCS values similar to the wind roughened sea surface. In the ice pack, the detection of open leads may remain challenging for wind roughened ones. As a consequence, the SIC estimates may be erroneous if the wind effect is not properly considered. In addition to the HH and HV “detrended" NRCS, the wind speed is extracted from a Numerical Weather Prediction (NWP) model, for instance from the model of the European Centre for Medium-Range Weather Forecasts (ECMWF). This information is projected over the same 200 meter spatial grid as the SAR data. The benefit of having this third layer is found negligible over the entire test database (+0.5%). This solution is however kept as it presumably brings some added-value for specific cases with radiometric ambiguities between sea ice and open water (even if it is not statistically representative).

Finally, the ECWMF wind speed and the co- and cross-pol detrended data are collocated on the 200 m spatial grid with OSISAF SIC products and MET Norway sea ice charts.

5.3.2 FCN architecture

General information The chosen FCN architecture is inspired by the U-Net architecture [163]. This model has a particular architecture in the shape of a U (see Fig. 5.3). Composed of an encoder and a decoder, its main characteristic is the so-called skip connections between the encoder and the decoder part. The depth of the model is determined by the number of max pooling stages. The detector stage, that is to say the activation phase to introduce non-linearity to the system is realized with a rectified linear activation unit (ReLU). Within the framework of this study, there are no in-depth studies on the hyperparameters of the model [14]. Instead, a set of tests has been conducted to properly adapt the FCN architecture to the problem. The depth of the network, the filter size and the loss function are specified in the following Subsections 5.3.2 and 5.3.2.

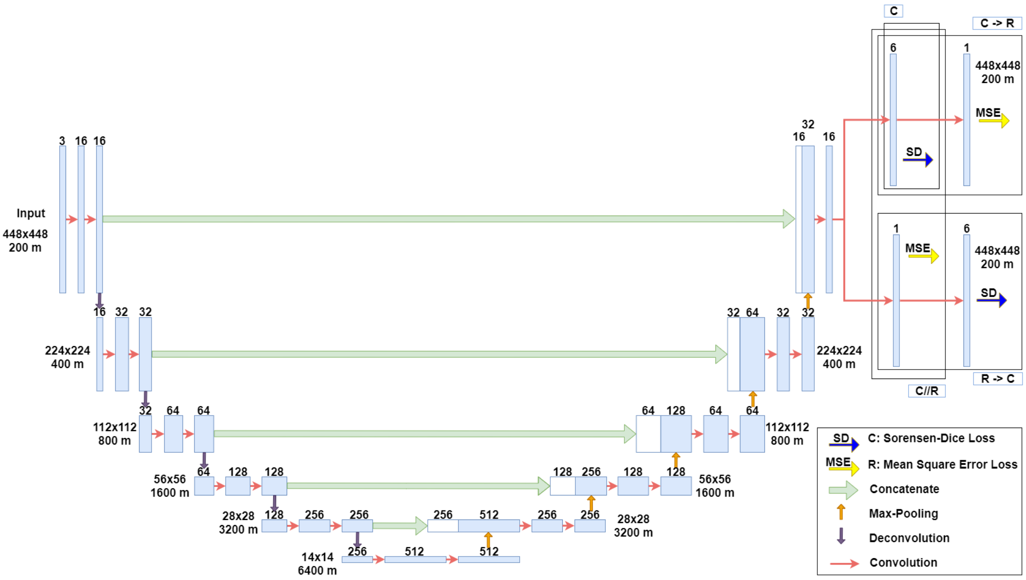


Fig. 5.3: Proposed architecture based on U-Net. The number indicated above blocks represents the number of filters or layer in each block. The numbers indicated beside blocks indicate the patch and pixel sizes

On the architecture depth A depth of five for the architecture has been chosen. This order of depth linked to the patch size has been decided to again try to mimic the human interpretation by sea ice analyst. At the coarsest scale, the chosen depth and the pooling factor of 2 induce a pixel spacing of $200 * 2^5 = 6400$ meters. This order of magnitude seems relevant for the problem of SIC estimates, defined as the ratio of sea ice coverage for a given area [238]. The size of convolution filters is set at 3×3, but some tests have been carried out with a size of 5×5 for filters. No benefit to the performance has been noticed.

On the loss functions As mentioned in Section 5.2, this study uses MET Norway ice charts considered as categorical information. The loss function can be set as categorical cross-entropy, which is a general configuration for multi-class classification. Here the Sorensen-Dice loss function, i.e. the F_1 score, which is the harmonic mean of the precision and recall, is used to further increase the robustness to class imbalance [182] (hereafter denoted by $\mathcal{C}$). Denoting y the ground-truth and $\hat{y}$ the prediction, this loss function is defined as:

$$1 - DICE(y, \hat{y}) = 1 - 2\frac{y \cdot \hat{y}}{y + \hat{y}} \qquad \text{(Eq. 5.1)}$$

For the final SIC prediction, the class with the highest pseudo-probability is chosen. However, this approach assumes that errors between all class pairs are equivalent. The semantic distance between classes should be considered. To integrate this prior information, the SIC estimation could be considered as a regression task [21] trained using the Mean Square Error as the loss function (hereafter denoted by $\mathcal{R}$), defined as:

$$MSE(y, \hat{y}) = (y - \hat{y})^2 \qquad \text{(Eq. 5.2)}$$

In this case, the output of the model is a single float value (between 0 and 1) rather than a vector of six elements. To keep the categorization framework while taking the semantic into account, a regression auxiliary loss is thus used. Auxiliary loss [228, 137] are additional losses that push the network to solve secondary tasks. Here, the addition of a convolution layer (either parallel [$\mathcal{C}$ // $\mathcal{R}$], meaning taking the same inputs, or subsequent, taking the categorization as input [$\mathcal{C}$ ->$\mathcal{R}$] or vice versa [$\mathcal{R}$ ->$\mathcal{C}$]) incentivizes the network to learn the classification task. In inference, the regression output is discarded and only the categorization output is kept. In the [$\mathcal{R}$ ->$\mathcal{C}$] variant, the categorization layer is expected to threshold the output of the regression layer (both with a minimum and maximum value). As such the chosen activation function is the sine function to ensure that it is non-monotonic and also has non-null derivatives almost everywhere. For the other variants, the activation function is a sigmoid for regression layers and softmax for categorization layers.

5.4 Results

From the data pre-processing mentioned in the previous Section (including pre-processing on the incidence angle normalization and bias correction for EW1, with wind speed as a third layer), the FCN architecture is then trained based on the training dataset, resulting in a set of optimized weights for the various configurations (see Fig. 5.3 with different loss settings). The optimizer is the Rectified Adam [117] with the default parameters $\beta_1 = 0.9$ and $\beta_2 = 0.999$. With the exception of the output layers, all convolutional layers use ReLu activation. No weight decay nor dropout is used. The learning rate is $\gamma = 10^{-4}$. The training phase is done by mini-batch Stochastic Gradient Descent (SGD). Each batch contains 16 training patches. The number of epoch is defined by an early stopping mechanism that stop the training when the network begins to overfit and only keep the model with the lowest loss on the validation set. The training uses two Tesla M10

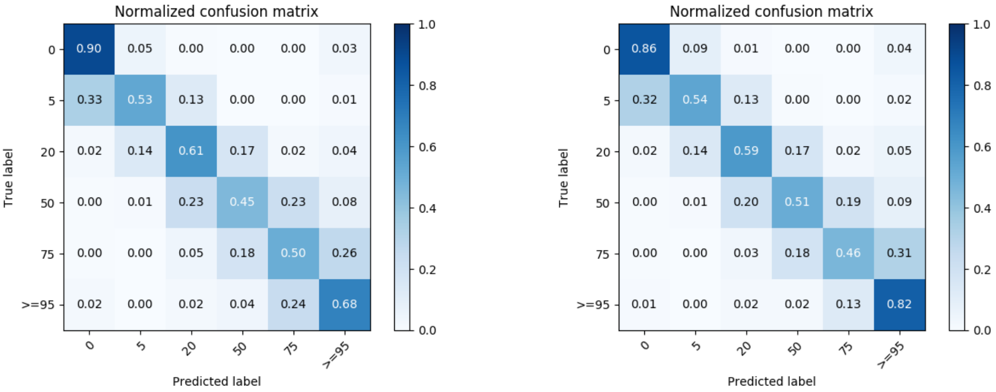


Fig. 5.4: Normalized confusion matrix against MET Norway data for the $\mathcal{C}$ (left) and $\mathcal{R} \rightarrow \mathcal{C}$ (right) configuration

GPU of 8 GB of RAM each. The analyses presented hereafter are based on the testing dataset (480 independent SAR images).

5.4.1 On the optimal loss function

The following table shows the classification results for the four aforementioned configurations.

Table 5.2: Overall Accuracy (AO) depending on the chosen loss functions

	$\mathcal{C}$	$\mathcal{C} \,//\, \mathcal{R}$	$\mathcal{C} \rightarrow \mathcal{R}$	$\mathcal{R} \rightarrow \mathcal{C}$
OA (in %)	73.0	74.0	69.7	78.2

With the $\mathcal{C}$ configuration alone, the Overall Accuracy (OA) is 73.0% on the test database. Adding the regression auxiliary loss function achieves a better performance, reaching an OA of 78.2% when the regression is followed by the categorization. When the layers are placed in parallel, the improvements are marginal (+1%). The OA is 69.7% when the regression is placed after the categorization. The over-performance of the $\mathcal{R} \rightarrow \mathcal{C}$ configuration over the classical $\mathcal{C}$ configuration is highlighted by the following confusion matrices (see Fig. 5.4). Interestingly, this configuration helps to improve the performance for high SIC values (from 68% with $\mathcal{C}$ to 82% with $\mathcal{R} \rightarrow \mathcal{C}$). Overall, the errors for the $\mathcal{C}$ configuration are still concentrated between adjacent classes, meaning that the semantic distances between classes are generally learned without the need for an auxiliary loss.

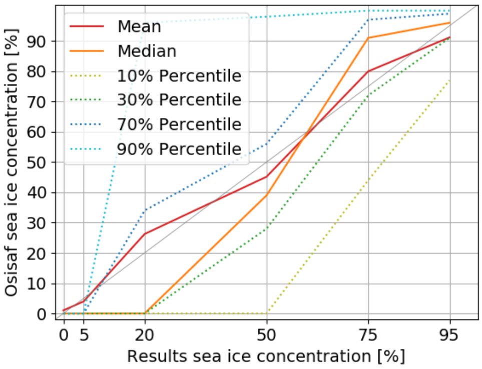


Fig. 5.5: Comparison between $\mathcal{R} \to \mathcal{C}$ classification and OSISAF SIC products which provide continuous estimates

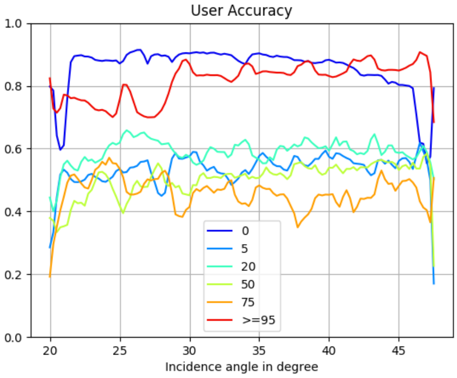


Fig. 5.6: User accuracy of $\mathcal{R} \to \mathcal{C}$ classification w.r.t MET Norway sea ice charts as a function of incidence angle depending on SIC classes

5.4.2 Classification performance and robustness

Fig. 5.5 shows the performance of the $\mathcal{R} \to \mathcal{C}$ classification with respect to OSISAF products. The estimates remain broadly similar between OSISAF and the model with a bias or median close to the unit line. The model seems to slightly overestimate SIC values for Very Open Drift Ice (20%), whereas this trend is reversed for higher concentrations ($> 50\%$). As indicated in [46], the observed trend is consistent with an overestimation of MET Norway ice chart compared to OSISAF SIC. In the absence of massive *in situ* data, this comparison does not prejudge the absolute quality of all these products [89]. Furthermore, the inherent variability of the manual annotation by sea ice experts at MET Norway is mitigated by the capacity of neural networks to work with noisy data [161].

With the proposed methodology integrating the noise correction, the bias correction for EW1 and the normalisation of the NRCS values w.r.t. the incidence angle, the classification performance is expected to be independent of that variable. In order to study more precisely the robustness of the model, Fig. 5.6 shows the dependency of the classification as a function of the incidence angle. The incidence angle does not seem to have a clear impact on the classification performance. The pre-processing carried out in the proposed methodology seems to be sufficient. However, there is a drop in performance at SIC of 95% for sub-swath EW1. The performance also fluctuates at the edges of each subswath indicating some remaining impacts due to the noise correction. Note that a homogenisation of the input database with enhanced noise annotation (IPF versions newer than v2.9 [available from 16/01/2018] with both range/azimuth noise dependency) would be required, in addition to a proper calibration gain at Level-1 for subswath EW1. The thermal noise removal as proposed in [148] is another alternative.

The radar backscattering from sea ice has a certain seasonal variability. A monthly analysis has been performed to analyze the behavior of the FCN model throughout the year. Fig. 5.7 represents the user accuracy for each SIC class and for each month. The same pattern is observed for SIC between 20 and 75 %: sea ice from May to July exhibits a slightly lower classification performance. For SIC of 5% and 95 %, the five first months of the year have degraded performance. Fig. 5.7 also presents the averaged difference between the model prediction and MET Norway ice chart. The overall bias is below 10% with a general trend towards overestimation for medium SIC values. The bias is negligible for the period from June to September. For sea ice concentration equal to 95% and above, the bias is negative due to the bounded and discretized values of SIC. The bias is higher during the summer season.

5.4.3 Case studies

Fig. 5.8 to Fig. 5.13 present some examples of the SIC estimates made by the FCN model. The general sea ice conditions are globally well transcribed even in harder cases like in Fig. 5.8, except for the last case in Fig. 5.13. The main differences with the MET Norway ice chart are for medium-to-high SIC values. The FCN model tends to predict a SIC of 75 % whereas the reference ice chart indicates a compact sea ice pack (see also Fig. 5.10). On the other hand, when there are no predominant sea ice fracture zones (see Fig. 5.9), the algorithm is able to estimate sea ice at a concentration of 95 %. On the right side of the study case of Fig. 5.10, the FCN models locally estimate SIC values of 20 % whereas the reference products (OSISAF and MET Norway) estimate a higher

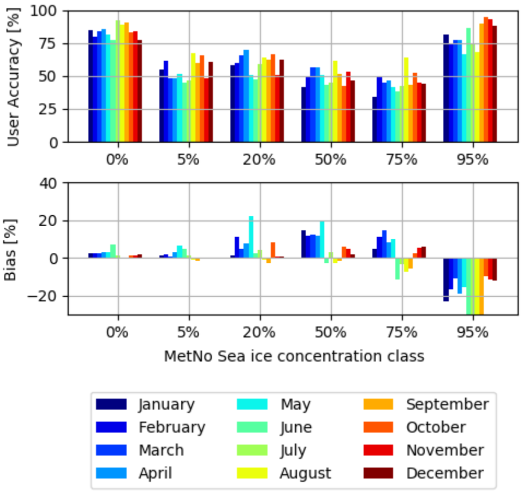


Fig. 5.7: User accuracy (Top) and bias between the $\mathcal{R} \rightarrow \mathcal{C}$ FCN configuration and MET Norway ice charts (Bottom) depending on Sea Ice Concentration classes, and for each month of the year

SIC value. In the corresponding SAR images, these areas show indeed low backscattering but spatially textured values in both co- and cross polarization channels. It certainly corresponds to refrozen leads. The $\mathcal{C}$ configuration seems to be largely influenced by the radiometric information of the input images. The $\mathcal{R} \rightarrow \mathcal{C}$ configuration provides a better performance, penalizing the large distance between classes with low and high SIC values, and hence being likely more sensitive to textural information. In the MIZ, the transition between free water and ice with high concentration is generally well performed by the FCN models.

From Fig. 5.11 to Fig. 5.13, some limitations of the current methodologies are shown. In Fig. 5.11, even though the scalloping and the inter-swath banding effects in the EW1 sub-swath is correctly handled by the FCN models, additional Radio-Frequency Interferences (RFI) create some erroneous estimates of SIC over the open sea surface. This case highlights also the need for a very accurate land mask and/or SAR image georeferencing. Some small islands in the Arctic regions may not well delineated (see land mask in black overlaid with the SAR images with a small island appearing as bright). This artefact results in erroneous estimates of SIC. The use of extra-large buffer to the GSHHG information may be a solution (in the current version only 1 km buffer). However, this will imply the impossibility to retrieve any sea ice information close to the coast. In Fig. 5.11, the $\mathcal{R} \rightarrow \mathcal{C}$ configuration provides a better performance for very compact ice pack

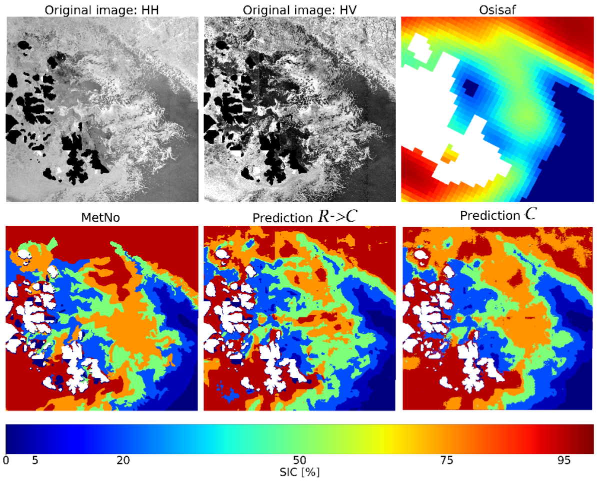


Fig. 5.8: Case study from S-1B EW acquired on 2018/12/05 at 04:43 UTC (Top from left to right) HH data, HV data, OSISAF SIC; (Bottom from left to right) MET Norway ice chart, and results from the $\mathcal{R} \rightarrow \mathcal{C}$ and $\mathcal{C}$ FCN configurations. Land and not valid data appear in white on sea ice charts.

compared to the $\mathcal{C}$ configuration. However, the $\mathcal{R} \rightarrow \mathcal{C}$ configuration fails in the area with the inter-swath banding effects. A full reprocessing of the database with new IPF version and enhanced noise correction would probably solve this issue. The case study in Fig. 5.13 shows Micro Convective Cells (MCC) over open sea surface at the right of the land area. This effect is mis-estimated by the FCN models as compact sea ice. As a fully supervised technique, FCN model may provide inaccurate results for outlier cases. The use of weakly-supervised or self-supervised techniques may provide interesting results. In addition, the joint use of specific CNN models for open sea surface may be adequate. Based on the initial work in [207, 205], the semantic segmentation of SAR scenes acquired over open and coastal ocean surface is indeed a actual prospect.

For all these cases, the benefit of SAR-based SIC estimates with respect to OSISAF products generated from coarse PMW data is clear. The time difference between averaged PMW track data and snapshot of SAR image may also explain some discrepancies, especially in the MIZ prone to rapid change (see Fig. 5.9 and Fig. 5.11).

5.5 Conclusion

This study uses an FCN inspired by the U-Net architecture to generate sea ice concentration maps from SAR images. Dual-polarized HH and HV Sentinel-1 data in EW Swath

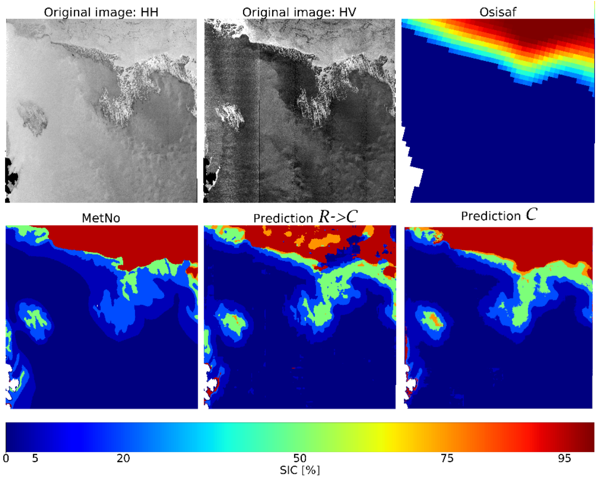


Fig. 5.9: Case study from S-1B EW acquired on 2019/02/04 at 03:46 UTC (Top from left to right) HH data, HV data, OSISAF SIC; (Bottom from left to right) MET Norway ice chart, and results from the $\mathcal{R} \to \mathcal{C}$ and $\mathcal{C}$ FCN configurations. Land and not valid data appear in white on sea ice charts.

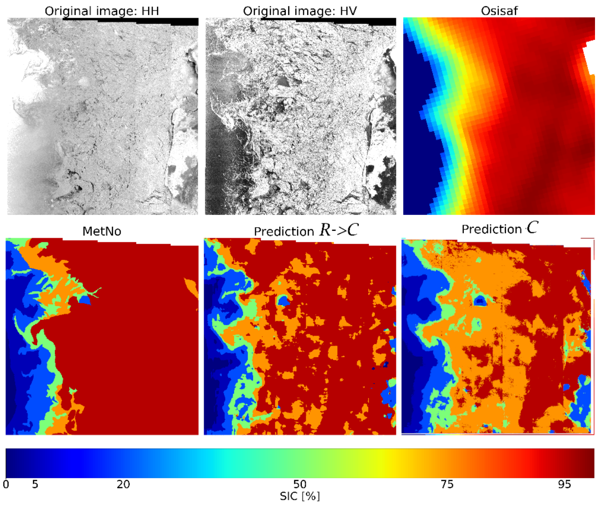


Fig. 5.10: Case study from S-1A EW acquired on 2019/04/05 at 07:53 UTC (Top from left to right) HH data, HV data, OSISAF SIC; (Bottom from left to right) MET Norway ice chart, and results from the $\mathcal{R} \to \mathcal{C}$ and $\mathcal{C}$ FCN configurations. Land and not valid data appear in white on sea ice charts.

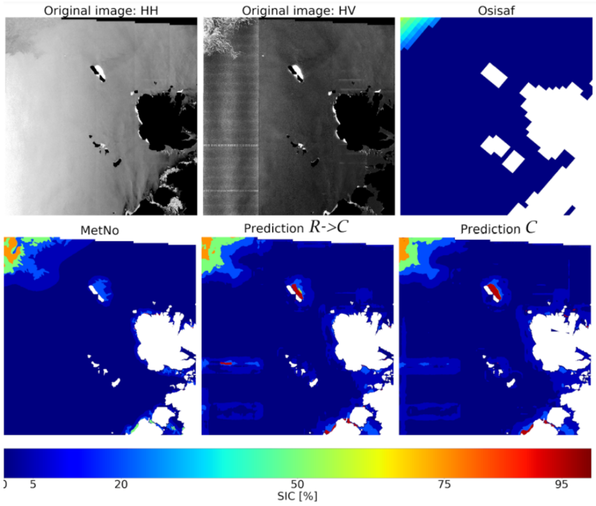


Fig. 5.11: Case study from S-1A EW acquired on 2018/12/05 at 05:33 UTC (Top from left to right) HH data, HV data, OSISAF SIC; (Bottom from left to right) MET Norway ice chart, and results from the $\mathcal{R} \rightarrow \mathcal{C}$ and $\mathcal{C}$ FCN configurations. Land and not valid data appear in white on sea ice charts.

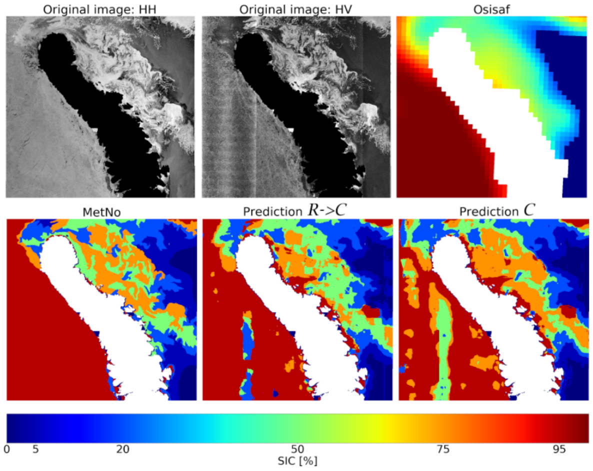


Fig. 5.12: Case study from S-1A EW acquired on 2019/04/05 at 02:58 UTC (Top from left to right) HH data, HV data, OSISAF SIC; (Bottom from left to right) MET Norway ice chart, and results from the $\mathcal{R} \rightarrow \mathcal{C}$ and $\mathcal{C}$ FCN configurations. Land and not valid data appear in white on sea ice charts.

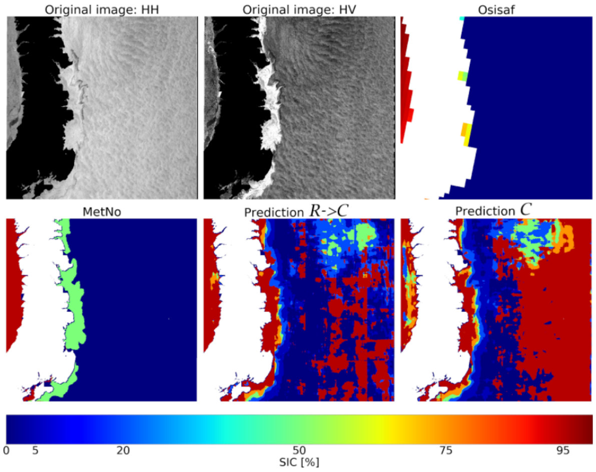


Fig. 5.13: Case study from S-1B EW acquired on 2019/02/18 at 03:31 UTC (Top from left to right) HH data, HV data, OSISAF SIC; (Bottom from left to right) MET Norway ice chart, and results from the $\mathcal{R} \rightarrow \mathcal{C}$ and $\mathcal{C}$ FCN configurations. Land and not valid data appear in white on sea ice charts.

mode are used to train the FCN model. The objective of this study is to produce sea ice chart all over the year, even in freezing or melting periods. Several configurations of the FCN architecture are tested. Our approach differs from other current studies: a FCN is able to integrate a large field-of-view of 89.6 × 89.6 km^2 and to produce SIC maps at 200 m resolution. The best results are obtained with pre-processed NRCS values corrected for 1) the bias induced by sub-swath mis-calibration and 2) the impact of incidence angle on NRCS. The wind speed information is also added as a third input layer. Its contribution in the global statistic assessment is not major although it certainly helps in specific cases.

The results are quite encouraging, with the principal confusions appearing between neighbouring classes. SIC charts are generally in good agreement with reference data. Thanks to the integration of an auxiliary regression loss, the performance of the FCN model is improved and seems to be equally sensitive to backscattering levels and textural information.

A comprehensive benchmark would be required to inter-compare all recently proposed methodologies in line with this study. It is all the more required that each study encompasses a specific set of sea ice variable, region of interest, temporal period, evaluation metric, and reference sea ice product prone to subjective analysis and uncertainties [20], and likely hidden bias in the training/test dataset. A comprehensive validation against various SIC products (e.g. ASI AMSR2 SIC products at 3.125) would represent an interesting perspective.

- END OF THE ARTICLE -

5.6 Afterword

To solve the sea ice concentration estimation, it was found beneficial to have both the categorization and the regression loss. We believe that the reason the regression loss is insufficient when used alone lies in the annotation process. As it divides the observation in segmentation areas rather than associating values in a continuous interval [0,1], it is intrinsically noisy. However, it is beneficial to the model as the categorization loss attribute the same error if open ocean is categorized as 100% sea ice concentration and 25% of sea ice.

Another factor of concern is the distribution of the sea ice concentration as a bimodal distribution with modes at 0% and 100%. In the following section, we consider the case of the rainfall estimation. This is also an initially continuous variable, though it only has one mode (the probability of rainfall decreases with its intensity). However, uncertainty with the ground truth led the direct regression difficult. Thus, similarly to the Sea Ice Concentration problem, we opted for segmenting the rainfall based on thresholds of rain rate rather than a direct regression.

Chapter 6

RAIN RATE ESTIMATION THROUGH SAR-TO-NEXRAD TRANSLATION USING CONVOLUTIONAL NEURAL NETWORKS

The following chapter is the reproduction of the preprint published in [28], in which the author of the manuscript was the lead author. At the time of this manuscript publication, the article is still being peer-reviewed.

6.1 Introduction

Precipitation monitoring and forecasting are major operational and scientific challenges. In the context of climate change in particular, extreme precipitation events and floods are expected to become more frequent [41]. The effect of these events can be mitigated by warning systems if early detection of precipitation is available.

Ground-based weather radars provide high-resolution rainfall measurements that are limited to coastal areas due to their range that spans only a few hundred kilometers. Their spatial resolution decreases with distance from the station with an azimuthal resolution of 1°. Such radars measure the reflectivity of the air column at different inclinations. When the beam encounters precipitation, part of the emitted signal is reflected back to the sensor with an intensity that depends on both the size and density of the water droplets. A direct relationship can then be derived between weather radar reflectivity and rainfall rate [129]. In previous work, data assimilation schemes have been investigated in order to reconstruct precipitation patterns from weather radar observation data [134].

Rain events can also be observed further from the coast using satellite instruments, though at lower spatial resolution. For example, brightness temperatures measured by microwave radiometers [116], such as SSMI/S, can provide rain rate measurements. These sensors are deployed on low-earth orbit satellites, thus providing coverage of the entire globe at an extended temporal resolution (revisit period of several days). SSMIS' along-track and cross-track resolution is respectively 14 and 13 km/px. Rain events can also be observed indirectly through potential associated lightning activity, as measured by near-infrared optical transient sensors such as the Geostationary Lightning Mapper (GLM) found aboard the Geostationary Operational Environmental Satellite (GOES-16).These geostationary on-board sensors provide continuous observations over a large area, covering almost an entire hemisphere in the case of the GLM. They do, however, lack in spatial resolution (8 to 14 km).

Space-based synthetic aperture radar (SAR) observations measure the backscattered radar signal at high resolution, typically 10 to 25 m for Sentinel-1, and thus provide image of a wide variety of meteorological and atmospheric phenomena [80]. Among these, rain signatures are often visible, appearing as light and/or dark spots. Studied for a long time now, these signatures can be a combination of different contributions from the roughness of the sea surface (increased or decreased surface scattering) or from the atmosphere (volume scattering or attenuation by hydrometeors). Their impact on radar backscatter varies as a function of many parameters such as incidence angle, wind conditions, signal polarization and frequency, or precipitation rate; [132, 26, 113, 130]. The preparation and analysis of well collocated Next Generation Weather Radar (NEXRAD) and Sentinel-1 measurements, both at high resolution, provide a unique opportunity to better characterize and detect rain signatures in SAR acquisitions.

Sentinel-1 satellites can operate in different acquisition modes. One such mode is the Interferometric Wide Swath (IW) mode. It covers several hundred kilometers in range and azimuth directions and extends over incidence angles between 29° and 46°. These products are generally used to study coastal areas. In contrast, the open ocean is mainly observed by the Wave (WV) acquisition mode, producing images of about 20x20 kilometers at two possible incidence angles: 23° (WV1) and 37° (WV2). Although different studies have addressed the categorization and segmentation of rain cells [205, 27], the calibration of SAR-derived rainfall product remains a challenge. The lack of a SAR dataset with ground reference rainfall data has certainly been a critical limitation. Earth observation systems, such as the Sentinel-1 satellites, now allow for large-scale datasets of SAR observations combined with rainfall data provided by weather radars, specifically NOAA's NEXRAD sensors [232].

In this study, we examine how the availability of such a large-scale SAR-NEXRAD dataset combined with deep learning approaches can lead to a breakthrough in SAR-derived rainfall estimation. We focus specifically on vertical-vertical (VV) polarization, available for both IW and WV modes. A key contribution lies in registration-based pre-processing steps to implement state-of-the-art supervised deep learning schemes. We show that a U-Net architecture far outperforms the filtering-based schemes previously suggested [232]. Thus, this study opens new avenues towards SAR-derived global precipitation products.

In Section 2, we present our enhanced SAR-NEXRAD dataset and preprocessing methodology. In Section 3, we describe the proposed deep learning approaches. In Section 4, we present a quantitative and qualitative evaluation of the SAR-derived precipitation estimate and in Section 5, we discuss the main aspects and contributions of this study in more detail.

6.2 Dataset

The Sentinel-1 mission consists of two satellites, Sentinel-1A and Sentinel-1B, whose synthetic aperture radars (SAR) regularly acquire data at 5.4 GHz (C-band). In this study, we used the IW acquisition mode. IW Ground Range Detected High Resolution (GRDH) products were obtained with a pixel spacing of 10 x 10 meters and a spatial resolution of approximately 20 x 22 meters. These products extended over a few hundred kilometers in range (250 km) and in azimuth. Weather radar reflectivity was obtained from NEXRAD, a network of 160 Doppler weather radars operating between 2.7 and 3 GHz. We used the basic reflectivity with a resolution of 1 km in range and 1° in azimuth.

6.2.1 Sea surface wind fields

As explained above, in the absence of rainfall wind speed is the primary parameter governing variations in sea surface roughness. These variables, however, can be decorrelated under heavy precipitation [207]. As such, it is necessary to collocate SAR observations with the most reliable wind speed information.

Two sources of wind speed were considered: European Centre for Medium-Rande Weather Forecast (ECMWF) model estimates and wind inversions from SAR observations.

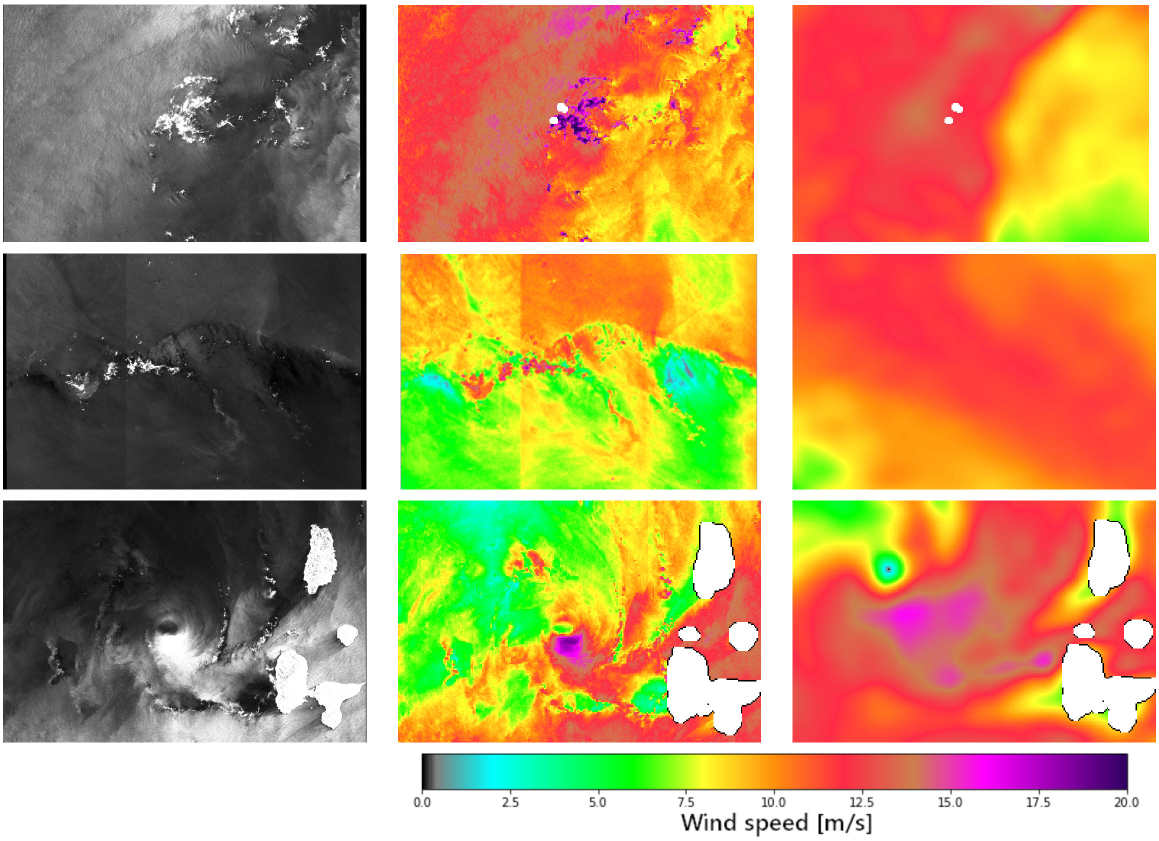


Fig. 6.1: Example of the discrepancies between ECMWF wind speed estimates and SAR observation. Each line corresponds to a product acquired (from top to bottom) on October 7th 2020 at 00:16:15, on May 17th 2018 at 23:05:21, on August 27th 2019 at 22:19:26 (when hurricane Dorian was near Guadeloupe). The first column shows the SAR observation, the second shows the SAR inversion, the third shows the wind speed obtained from the ECMWF in m/s. The SAR observations are shown in radar geometry, with the low incidence angles on the left.

Atmospheric model data wind field Model estimates were obtained from sea surface wind at a height of 10 m though the analysis of the ECMWF's global model: the Integrated Forecasting System (IFS); Were analyzed either 3-hourly or hourly forecasts operating at resolutions of 0.125° or 0.1°, before or after August 2019 (respectively).

SAR-derived wind field The SAR-derived sea surface wind field is estimated using CMOD5.N (a C-band geophysical model function) and auxiliary wind information from the ECMWF in a Bayesian wind inversion [140].

Wind field comparisons Significant differences between the SAR-derived and model-derived wind fields can occur if the model is not well phased with respect to the actual situation or if the difference between the analysis and observation times is too large. Other significant differences can occur when the sea surface roughness is related not only to the sea surface wind. Four physical processes contribute to the radar signature of rainfall events [3]: 1) scattering of the radar signal from the sea surface, the roughness of which is altered by both ring wave generation and wave damping due to turbulence caused by raindrops hitting the sea surface, 2) increased sea surface roughness due to downdraft winds often associated with rain cells, 3) scattering from splash products, i.e. craters, stalks, crowns and rain drops bouncing upwards, and 4) scattering and attenuation of the radar pulse by raindrops (hydrometeors) in the atmosphere (volume scattering and attenuation) which can become non-negligible at very high rain rates. Direct interpretation of the sea surface roughness as being a result of sea surface winds would lead to significant errors. An incorrect a priori wind direction model can also lead to incorrect estimates of SAR-derived wind speed and direction. This is generally the case for fast moving phenomena such as hurricanes, or local structures not seen by modeled winds such as cold pools, that are nonetheless related to the scattering mechanism.

Fig. 6.1 illustrates some discrepancies between the SAR inversion and the ECMWF wind speed. On the first line, acquired on October 7th at 00:16:15, heavy precipitation has produced a bright area that leads the SAR inversion to overestimate the wind while the other parts of the IW show good concordance with the ECMWF product. The second line, observed on May 17th at 23:05:21 shows an area of low sea surface roughness in the lower half. The SAR inversion contradicts the ECMWF results by more than 5 m/s over most of the IW. This discrepancy is related to a time difference of 55 minutes. The last line, in which there is a time difference of 79 minutes (observations acquired on August 27th at 22:19:26), indicates a misalignment between the ECMWF wind speed and the SAR inversion, as illustrated by the position of the eye of the cyclone.

In the following sections, analyses using wind speed were performed using the ECMWF wind speed to ensure independence with the presence or absence of rain, despite occasional errors.

6.2.2 SAR preprocessing

Unlike Stripmap, in which the SAR sensor observes the entire swath simultaneously, TOPSAR [222] divides its observations into three subswaths for the IW mode. The subswaths are themselves divided into several bursts in the azimuth direction. This technique, used in both the IW and Extra Wide Swath (EW) modes, is intended to provide a uniform signal-to-noise ratio over the entire observation. The calibration of each subswath (and burst) is performed by calculating the theoretical gain of the SAR antenna.

In addition to the data calibration and noise correction, the radar signal strength is also corrected for its dependence on the incidence angle. This reduces the signal dynamics resulting from high or low signal reflectivity at low or high incidence angles, respectively. The CMOD5.N geophysical model function (GMF) [71] is used assuming a wind speed of 10 m/s and a direction of 45° with the radar azimuth, providing the same SAR preprocessing as applied by previous studies [205].

Residual variation (of about 20%) can still be observed in the data set. We chose not to further normalize the sea surface roughness (e.g., by averaging over each incidence angle or sub-band), as such normalization would be affected by the presence of meteo-ocean or land phenomena and would vary from one SAR product to another.

6.2.3 Enhanced collocated Sentinel-1/NEXRAD dataset

Several imporvement steps, described below, were applied to the initial collocations from [232] before their use in deep learning.

Reflectivity measured from weather radars can be affected by a multitude of factors that degrade their quality. For example, since the minimum elevation angle of NEXRAD is not zero (but 5°), the further the distance from the ground station, the higher the altitude of the observed volume. Since SAR imagery observes the ocean surface, the concordance between the meteorological-ocean (metoceanic) situation of the two observations decreases as the distance from the ground station increases. This concordance is even more difficult to qualify automatically since the altitude observed by the ground station depends not only on the distance but also on the refractive index of the air, which is related to the atmospheric temperature gradient. The bright rain signatures may also be due to strong reflectivity from the melting layer (4th physical process aforementioned) [3], located typically at 3-4 km in altitude. Again, varying altitude levels seen by the ground Radar may lead to slight geographical shifts with respect to the SAR signatures at this fixed altitude. Other phenomena can hinder the radar's ability to provide meaningful information. Topography, for example, can mask rain signatures located behind an obstacle. Other factors include moisture on the radom enclosing the weather radar [215], which affects the observed reflectivity and is not corrected. These sources of discrepancy are difficult to quantify automatically and require manual verification.

The maximum time difference between the SAR and NEXRAD observations was limited to twenty minutes, in order to reduce the displacement and evolution of the rain cell between the observation times of the two instruments. Collocations with imperfect agreement between NEXRAD and Sentinel-1 were also removed manually. In addition, to ensure similar SAR processing for all observations, the SAR data is limited to observations after March 2018. This corresponds to Sentinel-1 Instrument Processing Facility version 2.90, which has improved noise correction and the associated signal-to-noise ratio (SNR). In total, these three constraints reduced the number of available wide swath SAR products from 1064 to 53.

Each of these IWs was divided into patches of approximately 25 x 25 kilometers and 256 x 256 pixels. Patches were extracted from the swaths with a step size of half their width. This extraction ensures that a metocean situation occurring at the edge of one patch is found in the center of the next patch. It also implies that each pixel is present in four patches. Data leakage between the training, validation, and test sets is resolved by splitting these subsets at the IW level rather than the patch level.

Patches were rejected if more than fifty percent of the area was covered by land, as this could lead to land masking. In addition, patches were rejected if the maximum reflectivity of NEXRAD was less than 25 dBZ (less than 1 mm/h according to the Z-R relationship [215]). Rejecting patches without measured reflectivity ensures that the area is neither out of radar range nor hidden behind an obstacle. Rejecting patches without rain also mitigates the imbalance in the dataset caused by the overrepresentation of low reflectivity measurements. Finally, to maximize the overlap of SAR and weather radar signatures, the patches were aligned manually. This was carried out independently for each patch using a constant translation of the NEXRAD measurement to overlap the SAR signatures. An example of this operation is shown in Fig. 6.2. This geographic repositioning corrects for remaining collocation problems that may be related to the displacement or evolution of rain cells between NEXRAD and SAR observation times or to the different altitudes at which the phenomena were observed.

Patches showing a low correlation between NEXRAD and Sentinel-1 observations were rejected by visual inspection. After these operations, from the 9574 patches extracted from the 53 IWs, 1570 remained. Patch locations, alongside the NEXRAD stations, are shown in Fig. 6.3.

To determine why the collocations do not overlap perfectly, statistics were computed

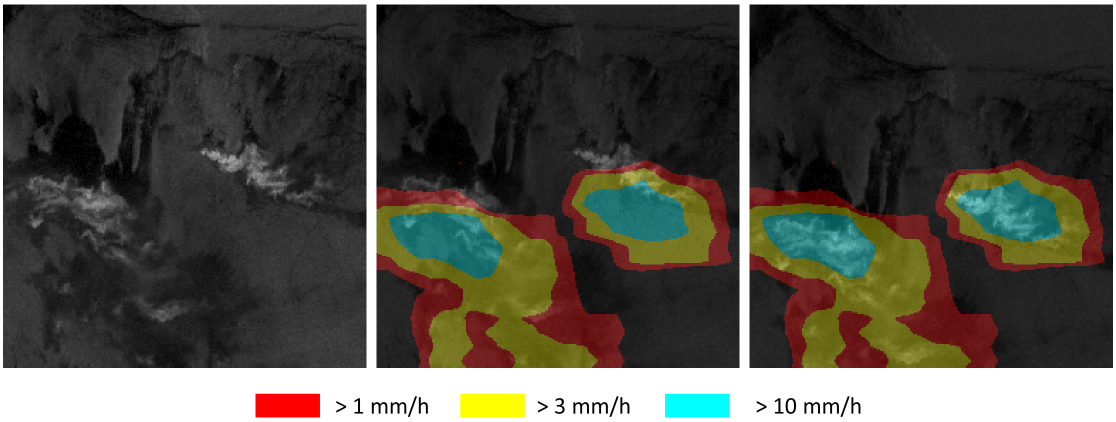


Fig. 6.2: Registration example. On the left, a patch (20 x 20 km) observed on May 5th 2018 at 23:05:20. In the center, the corresponding NEXRAD measurement. The cyan-colored area does not overlap perfectly with the SAR signature (center), therefore, we perform a manual registration (right).

on the realignment vectors. These showed a correlation between the distance to the nearest ground station and the error (correlation coefficient $R^2 = 40.4\%$). However, no such correlation was found for the direction of displacement ($R^2 = 3.4\%$), nor for the ECMWF wind speed at 10 meters ($R^2 = 2.7\%$) which would have indicated displacement of the rain cell by the wind.

After extraction, the patches were divided into training, validation and test sets. Special attention was paid to their distribution. As explained in the previous section, patch extraction leads to the co-occurrence of the same pixels in two adjacent patches. Even without overlap between patches, two adjacent patches would be affected by the same phenomenon, both from a meteorological and a sensor point of view, as we can assume that they would share characteristics in measurement, collocation, etc. To avoid any kind of leakage between the different sets, the division between the three subsets of data was performed at the swath level. They were equalized on both the NEXRAD reflectivity, to ensure a similar distribution of rainfall in each data subset, and on wind speed, which presumably has an impact on rain rate prediction capabilities. This assumption is driven by the known increase in sea surface roughness under the impact of rain and wind, as illustrated in Fig. 6.4. This figure also indicates that the NEXRAD reflectivity and Sentinel1 backscatter are decorrelated for reflectivities below 30 dBZ.

The resulting distributions for each subset of the data are detailed in Table 6.1. As noted, the dataset suffers a lack of data at higher reflectivity and wind speeds. Indeed, as

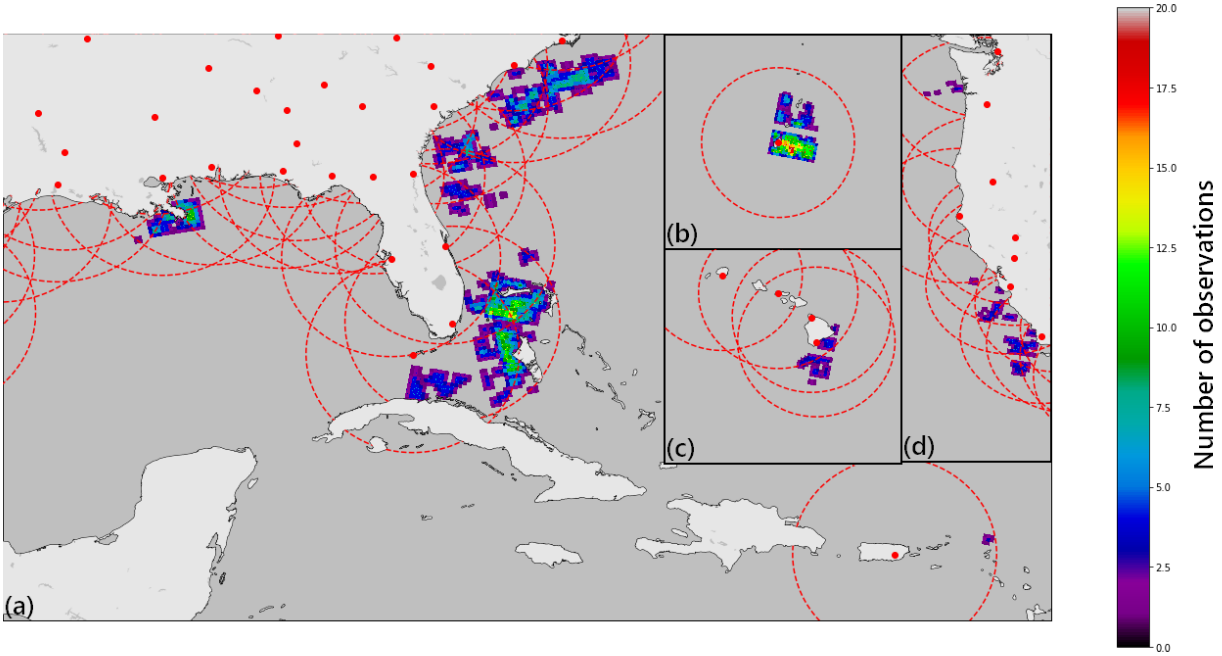


Fig. 6.3: Geographic distribution of patches in the dataset centered around Florida (a), Guam (b), Hawaii (c), and California (d). NEXRAD stations are represented by a red dot. Circles correspond to a 350 km range. Values are the number of observations over the corresponding latitude/longitude.

an example, only two wide-swaths contain wind speeds above 12 m/s, and only one has wind speeds above 16 m/s. It can also be noted that the standard deviation of the SAR surface roughness is higher when wind speed is higher than 12 m/s.

The reflectivity is divided into four intervals: [0, 24.7], [24.7, 31.5], [31.5, 38.8] and [38.8, $+\infty$] dBZ. With the general NEXRAD radar formula [215], these intervals can be approximated in terms of rainfall by [0, 1], [1, 3], [3, 10], and [10, $+\infty$] mm/h. Therefore, three segmentation masks are generated by thresholding at these values. The objective of the precipitation estimation thus shifts from a continuous regression to the segmentation of the following overlapping precipitation classes: $\geq$ 24.7 dBZ, $\geq$ 31.5 dBZ, and $\geq$ 38.8 dBZ. Reflectivities below 24.7 dBZ are considered to be rain-free. A direct approach, to predict continuous reflectivity, proved difficult due to the rarity of the strongest rain events and the poor correlation between the reflectivity value and the SAR signature. The non-uniform distribution of NEXRAD reflectivity is also problematic because low reflectivities are over-represented, as shown in Fig. 6.4.b. The 1570 patches, divided in training, validation and test sets are available on kaggle[1].

---

[1] www.kaggle.com/rignak/sentinel1-nexrad

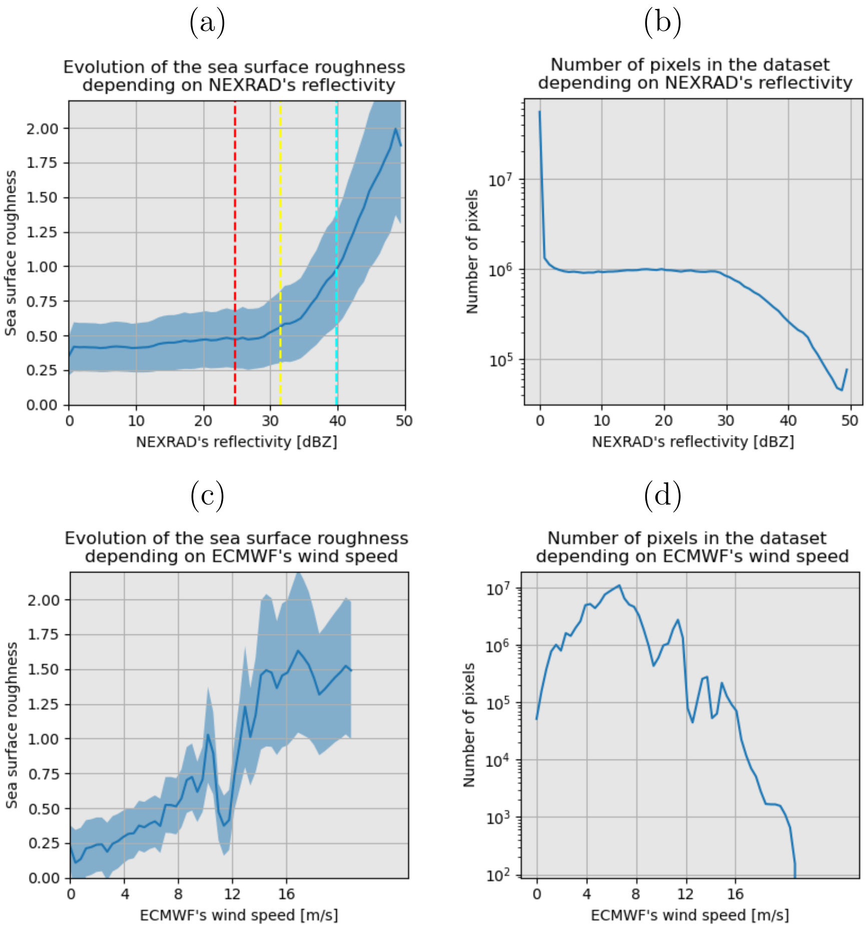

(a) (b) (c) (d)

Fig. 6.4: Evolution of sea surface roughness (and associated number of pixels) as a function of NEXRAD reflectivity (a) or ECMWF wind speed (c) and their associated number of pixels (b, d). The red, yellow and cyan vertical lines are the threshold values used to separate the four precipitation classes. The blue area represents the standard deviation around the mean at each point. The decrease in sea surface roughness at 11 m/s is due to a single IW (taken on May 17th 2018 at 23:05:21). Comparisons between SAR-derived and the model wind speed and rain free regions confirms the model is overestimating the wind speed for this particular case.

GOES-16 Geostationary Lightning Mapper As illustrated in Table 6.1, the available collocated data are limited to wind speeds below 12 m/s. To investigate the ability of the model to generalize to higher wind speeds, a collocated lightning data set was constructed. It has been demonstrated that precipitation can be approximated by a linear function of the lightning frequency [176]. The quality of this approximation can be very variable depending on the type of precipitation (convective/stratiform), the loca-

Dataset		Train (39 IW)	Validation (7 IW)	Test (7 IW)	% of the total
Reflectivity & Rain Rate (NEXRAD)	$[0, 24.7[$ dBZ $[0, 1[$ mm/h	79.5 %	9.6 %	10.9 %	85.1 %
	$[24.7, 31.5[$ dBZ $[1, 3[$ mm/h	79.9 %	9.6 %	10.5 %	7.7 %
	$[31.5, 38.8[$ dBZ $[3, 10[$ mm/h	79.3 %	9.7 %	11.0 %	5.4 %
	$\geq$ 38.8 dBZ $\geq$ 10 mm/h	79.0 %	9.8 %	11.2 %	1.8 %
Wind Speed (ECMWF)	$[0, 4[$ m/s	79.3 %	9.7 %	11.0 %	11.7 %
	$[4, 8[$ m/s	79.1 %	9.7 %	11.1 %	69.7 %
	$[8, 12[$ m/s	79.1 %	9.5 %	11.3 %	17.1 %
	$[12, 16[$ m/s	100 %	0.0 %	0.0 %	1.5 %
	$\geq$ 16 m/s	100 %	0.0 %	0.0 %	0.1 %

Table 6.1: Per-pixel distribution of NEXRAD reflectivity and ECMWF wind speed for the patches contained in each set.

tion (land/ocean) or the climate of the studied area (humid/arid). Still, the collocation between weather radars and lightning activity was also explored in a data assimilation framework by [101], described satisfactory concordance between higher reflectivity and lightning events. The use of this lightning proxy to compare with bright rain signatures on SAR images is further motivated by their relation with graupel and hail, typically met in cumulonimbus clouds with strong lightning activity.

Thus, in this study, we used a binary segmentation mask of lightning presence or absence as a proxy for a binary rain map. These collocations were obtained with the GOES-16 GLM [61]. The GLM is a camera embedded on a satellite in geostationary orbit. It records lightning at a spatial resolution between 8 km/px and 14 km/px (depending on the latitude). With a temporal resolution of 20 seconds and coverage extending to almost an entire hemisphere, it is possible to collocate a large amount of data between the GLM and Sentinel-1. This allows us to search for extreme events (i.e. high wind speed, high number of lightning events). Lightning events are defined as a single intensity peak occurring in an image but are aggregated into clusters, defined as pixel-wise adjacent events, and flashes (groups with a pair of events less than 330 ms and 16.5 km apart). In this study, only collocations less than 20 minutes from the SAR acquisition were used. The objective was to study the concordance between our precipitation classification and this proxy for different Sentinel-1 incidence angles and ECMWF wind speeds.

6.3 Proposed Deep learning framework

This section introduces the proposed deep learning schemes. Within a supervised training framework, we explore two neural architectures: the first one derived from the state-of-the-art processing based on Koch filters [93], the second one based on a U-Net architecture [163].

6.3.1 Koch-filter-based architecture

Koch filters [93] are four different high-pass filters that each detect different patterns thus allowing the detection of heterogeneous areas of ocean surface roughness. Their original use was to identify areas where backscatter is caused by non-wind phenomena (ships, rain, interference, tidal currents...), as this would exclude these areas from a wind speed/direction estimate. Koch filters can be optimized to produce binary rainfall maps, as precipitation is a major source of heterogeneity [232].

Specifically, Koch filters are defined as four different high-pass filters scaled by a linear function and clipped to maintain the result between 0 and 1. The output of the filters is the root mean square of these clippings. [232] estimated thresholds in order to derive binary rain maps from this final value, depending on the resolution and polarization of the input. We extended this framework to multi-threshold segmentation by rewriting the Koch filters as a Convolutional Neural Network (CNN) defining the scaling function parameters. The four high-pass filters were used on the input and left side as in the original version. To guarantee a non-zero gradient, the clipping is replaced by the sigmoid function $\sigma(x) = [1 + \exp(-a(x + b))]^{-1}$. We set $a = 4$ and $b = 0.5$ so that the inflection point is at x=0.5, $\sigma(0.5) = 0.5$ and $\frac{d\sigma}{dx}(0.5) = 1$. The change in activation affects the filter result, but the relative difference from the original Koch filter is only 0.8% when initialized with the same parameters. Fig. 6.5 illustrates this Koch filter incorporated into the CNN. This formulation allows for segmentations for different rain thresholds, unlike the original rain detection [232].

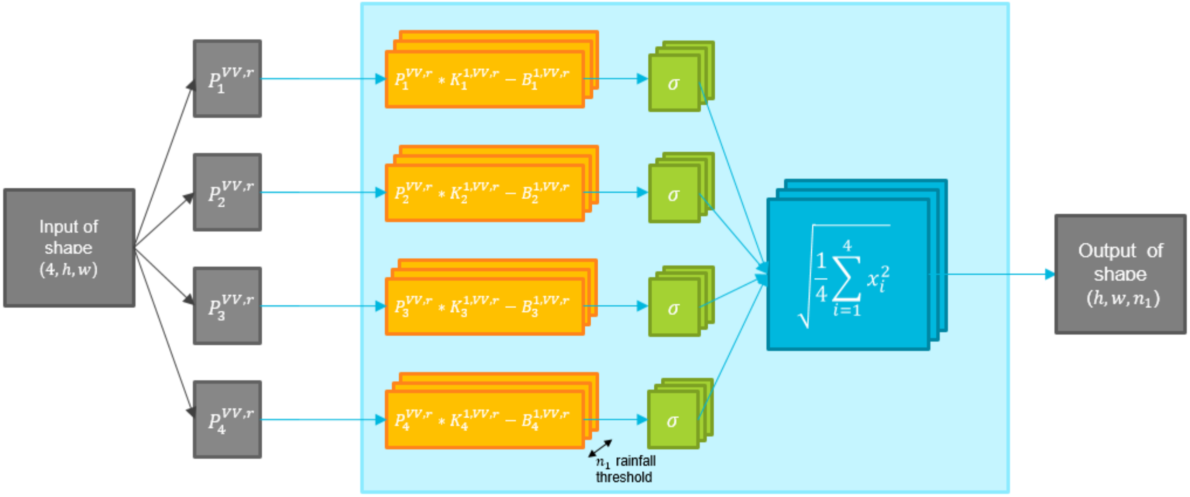


Fig. 6.5: Architecture of the multi-threshold Koch filter as a convolutionnal neural network. $P_i^{VV,r}$ is the output of the high-pass filter i at the resolution r, for the VV polarization. σ is the sigmoid function defined as $\sigma(x) = [1 + \exp(-4(x + 0.5))]^{-1}$. $K_j^{i,VV,r}$ and $B_j^{i,VV,r}$ are the scaling function parameters for each resolution r, polarization VV, filter j and precipitation regime i. The results are fused along the filters by a quadratic mean.

The model is trained to minimize the mean square error with the Adam optimizer, a learning rate of 10^{-3} over 200 epochs with a batch size of 32. As previously mentioned, the convolution kernels were initialized following the original Koch filter formulation [93].

6.3.2 U-Net architecture

Among the variety of state-of-the-art neural architectures for image segmentation and image-to-image translation problems, we consider here a U-Net architecture [163]. U-Net is an auto-encoder model with "skip connections" that allow information to be propagated horizontally from the encoder to the decoder, bypassing the central part of the network. This architecture is well established and has already been used in SAR imagery for sea ice concentration estimation [54] and semantic segmentation [27].

The specific model used is shown in Fig. 6.6. Compared to the original U-Net model, it has one less stage to reduce the receptive field and ensure that, when applied to full IW observations divided into overlapping tiles, the output mosaic has continuity between adjacent tiles. The width of the theoretical receptive field is 140 pixels, but the effective receptive field, which is smaller due to the contribution of neighboring pixels that decreases with distance [125], is small enough to ensure continuity. The number of weights, independent of input size and spatial resolution, was 3,117,731.

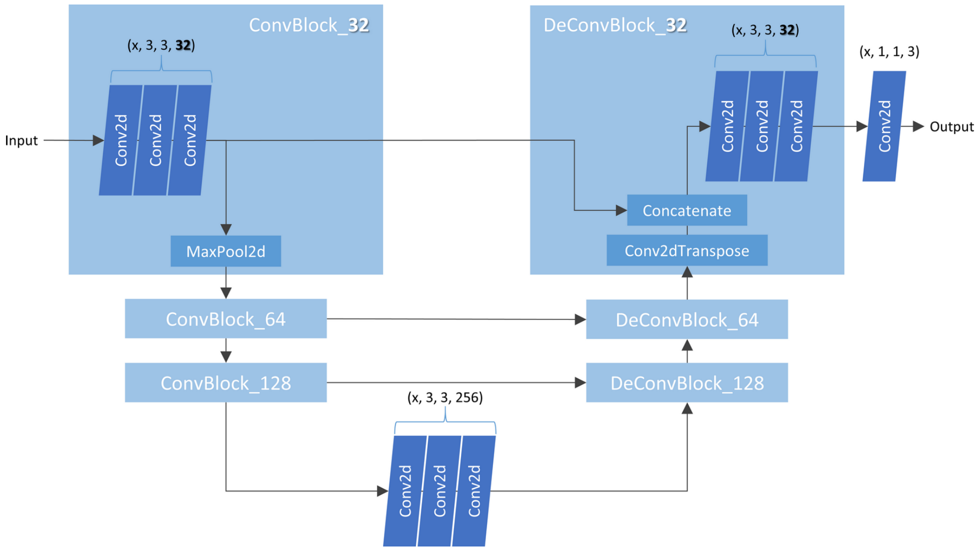


Fig. 6.6: Architecture of the U-Net model used to classify rainfall amount (from light to heavy) using Sentinel-1 ocean surface roughness.

The model was trained to minimize the mean square error, using Adaptive Momentum with a learning rate of 10^{-5}, for 500 epochs. In all experiments, 500 epochs were sufficient to achieve loss convergence. The batch size was 32, except at 100 m/px where GPU memory constraints led to reducing the batch size to 16. The code used to train the model is accessible on GitHub[2].

6.4 Evaluation framework

Existing Koch filters are designed to produce binary maps of rain presence or absence. To compare this framework with our multi-threshold models (i.e., the fine-tuned Koch filter and U-Net), we computed the F1 score on the binary segmentation problem for each threshold (1 mm/h, 3 mm/h, and 10 mm/h). The F1 score is defined as the harmonic mean of recall and precision. Recall is the average diagonal value of the row-normalized confusion matrix. It is also known as the producer's accuracy. Precision is the diagonal mean value of the column-normalized confusion matrix. It is also known as the user's accuracy. When evaluated on a binary segmentation problem (through a 2 x 2 confusion matrix), we call it the "Binary F1-score". This value indicates the ability to separate rain-

---

[2] https://github.com/CIA-Oceanix/SAR-Segmentation/tree/oceanix

free observations from patches with rain. The F1 score is also used to evaluate the ability to distinguish between different rain rates. In this case, the shape of the confusion matrix is (4,4). This F1-score is indicated as the "Multiclass F1-score".

We therefore compare the *Binary Koch filter*, which is the baseline and state of the art in rain detection, to the *Fine-tuned Koch filter* (the CNN-embedded multi-label Koch filter), and the *U-Net* architecture. For the latter, results using a dataset without the manual registration are also provided to justify the need for this particular operation. To evaluate the performance of the segmentation as a function of incidence angle, wind speed, and distance from the coast, the binary F1 score was also calculated by varying these parameters. In this case, the groundtruth was not provided by the NEXRAD weather radars, but by collocated lightning groups from GLM. These lightning clusters were used as a proxy for a binary precipitation segmentation. The U-Net models were trained and tested at resolutions of 100 m/px to 800 m/px. Because the receptive field of the Koch filters is smaller, they were only used down to 200 m/px, in accordance previous works [232]. Since some parts of the methodology were stochastic, such as the order of the images provided to the network or the initialization of its weights, the results are given as the mean and standard deviation over five training runs to decrease the chanceof a lucky initialization [11].

6.5 Results and discussion

Table 6.2 compares the binary Koch filters, the fine-tuned multi-label Koch filters, and the U-Net architectures for binary segmentation (for different precipitation thresholds) and multi-class segmentation. The binary Koch filter performs worse on the binary F1 score at each precipitation threshold than do both the fine-tuned Koch filter and the U-Net architecture. The best binary segmentation is obtained at 200 m/px for each method, and the U-Net architecture outperforms both variants of the Koch filter. Great variability is observed in the results and can be explained by the difference in the number of parameters (24 for the finely tuned Koch filter and over 3 000 000 for the U-Net model).

As the multi-class F1 score is not only influenced by its ability to detect precipitation but also by its ability to distinguish the severity of precipitation, it indicates higher performance at 400 m/px. Interestingly, this is also the best resolution obtained by [232], although the results were computed on a different data set. This leads one to believe that the increase in resolution, while giving more accurate information, is counterbalanced by

Model	**Input resolution**	**Binary F1-score (> 1 mm/h)**	**Binary F1-score (> 3 mm/h)**	**Binary F1-score (> 10 mm/h)**	**Multiclass F1-score**
Binary Koch's filter	200 m/px	44.3%	34.7%	22.8%	N/A
	400 m/px	37.3%	26.5%	15.1%	N/A
	800 m/px	32.9%	22.2%	11.1%	N/A
Fine-tuned Koch's filter	200 m/px	45.9% (0.04%)	41.6%(0.06%)	38.7% (2.09%)	34.8% (0.2%)
	400 m/px	43.2% (0.15%)	40.9% (0.14%)	37.9% (0.58%)	35.9% (0.3%)
	800 m/px	38.3% (0.05%)	37.2% (0.18%)	32.3% (1.65%)	35.2% (0%)
U-Net	100 m/px	53.7% (2.36%)	52.5% (2.03%)	55.6% (2.30%)	47.2% (1.9%)
	200 m/px	50.5% (1.69%)	47.5% (1.72%)	48.0% (1.87%)	46.0% (3.0%)
	400 m/px	51.2% (1.72%)	46.8% (1.75%)	47.2% (2.14%)	**50.5% (2.8%)**
	800 m/px	45.4%(0.93%)	40.4%(1.26%)	40.2%(1.56%)	47.1% (0.9%)
U-Net without registration	100 m/px	51.6% (0.56%)	50.2% (0.42%)	25.8 % (14.76%)	41.0% (1.2%)
	200 m/px	50.1% (0.54%)	48.2% (1.32%)	42.7 % (9.24%)	36.1% (2.4%)
	400 m/px	49.1% (0.93%)	47.6% (0.87%)	48.5 % (1.35%)	41.0% (1.2%)
	800 m/px	44.2% (2.57%)	42.5% (3.06%)	43.8 % (2.41%)	41.5% (1.0%)

Table 6.2: Evaluation of the binary and fine-tuned Koch filters and U-Net model on the test subset. Results are provided as a mean with standard deviation over five runs.

the decrease in contextual information. Since the architecture of the network does not change, the receptive field is the same if measured in pixels, but is reduced if we consider the area covered in km^2 . The Koch filters are less affected by the change of context because their effective field is defined by the low pass filters they use as input.

The confusion matrices, shown in Fig. 6.7, indicate that the U-Net architecture (right) is more accurate than the fine-tuned Koch filters (left) for each threshold (11.9% vs. 31.0%, 10.6% vs. 26.2%, and 29.6% vs. 70.5% for the 1, 3, and 10 mm/h thresholds, respectively). However, 31% of the [3, 10] mm/h class ramains unrecognized by the model as being rain. A large standard deviation in the confusion matrix indicates instability in the training. The U-Net performs particularly well in detecting heavy rainfall, as 93.7% of rainfall above 10 mm/h was predicted to be above 3 mm/h. The refined Koch filter only achieved 64.4%. Without the manual registration (center), the deep learning model was still able to differentiate rainfall from rain-free areas, but failed to assign a rainfall class.

Fig. 6.8 shows some examples of rainfall predictions using SAR data and either the U-Net architecture or the fine-tuned Koch filters. Overall, the SAR rainfall predictions appear to concord with the NEXRAD acquisitions over the ocean, with different sensitivities. In the first case (line 1), the fine-tuned filter well detects the rainy regions but indicate less or no rain within the largest rainy patches. This is mainly due by the direct use of high-pass filters while the U-Net architecture is more general. Also, the fine-tuned filter tends to detects smaller rain patches, not detected in NEXRAD. Three neighbouring

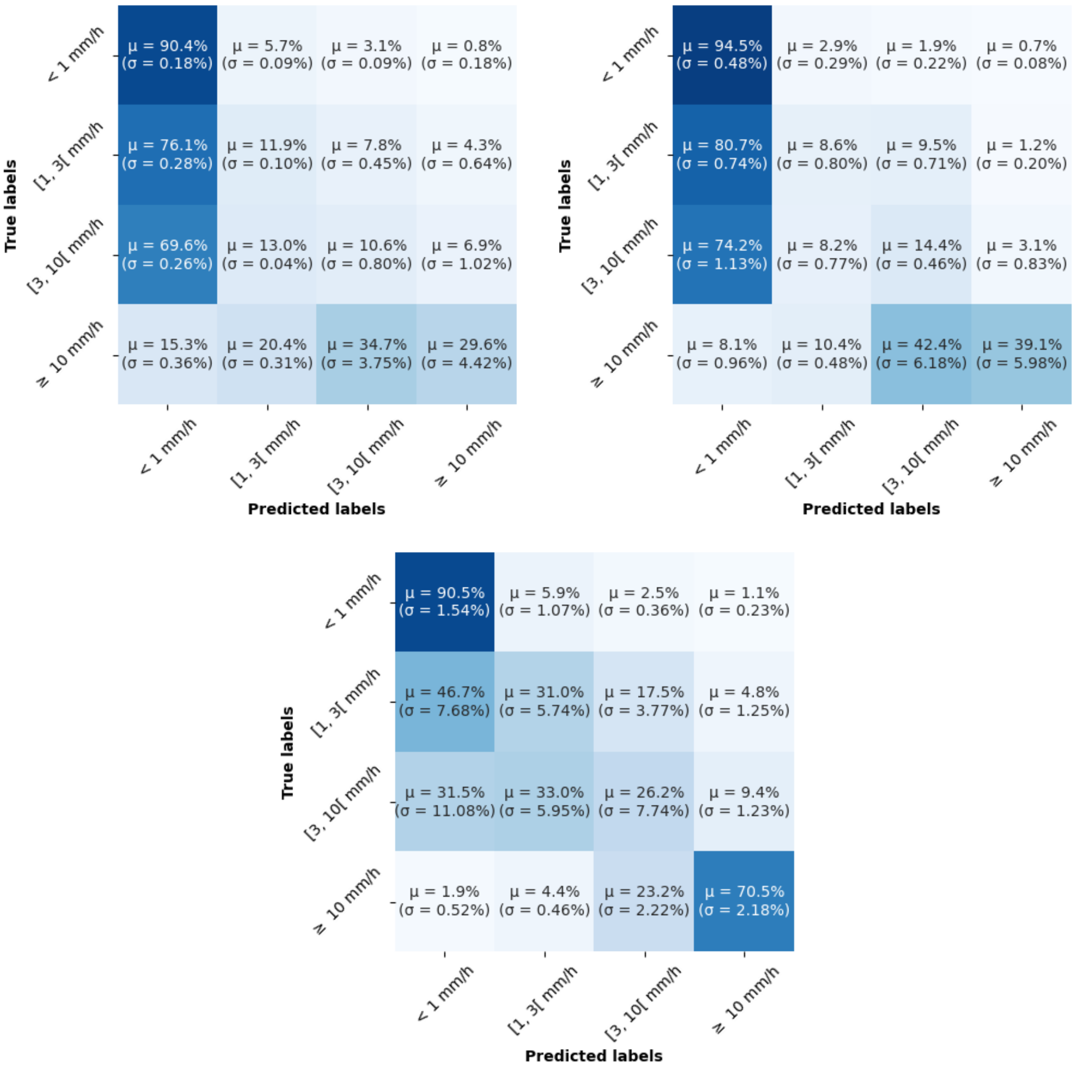


Fig. 6.7: Normalized confusion matrices of the fine-tuned multi-class Koch filter (top-left), the U-Net model without (top-right) and with the manually-registred dataset (bottom), each at 400 m/px.

dots located on the right-hand side actually correspond to 3 ships. In the second case (line 2), the fine-tuned filter wrongly interprets gusts fronts as strong rain due to their strong discontinuity with respect to the background radar signal. In the third case (line 3), we illustrate limitations of the NEXRAD system as it is unable to detect the rain patches located on top and on right-hand side, possibly due to masking by the topography.

Since the test set is quite small, due to both a lack of data and the requirement to balance it over each wind regime, we use the GLM data as an auxiliary data source to

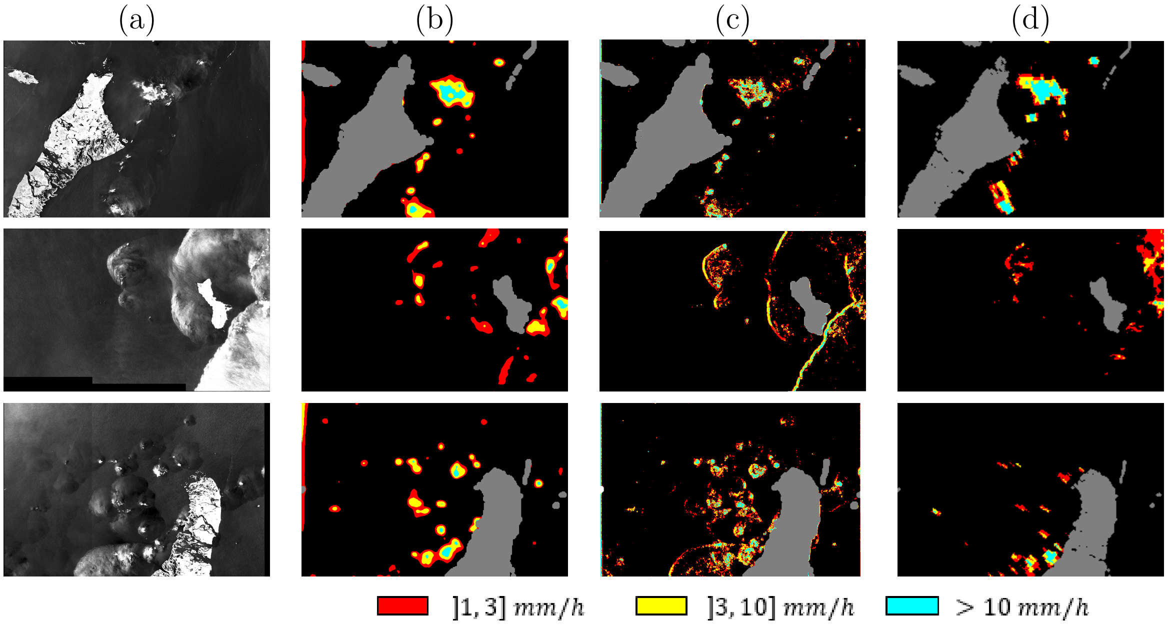


Fig. 6.8: Examples of SAR-derived rain rate estimation: Sentinel-1 SAR observations (a) with U-Net-based estimation (b), the estimation using the trained Koch-based architecture to 400 m/px (c), and the thresholded and collocated NEXRAD reflectivity. From top to bottom, the observations were acquired on April 24th 2018 at 11:10:12, August 05th 2018 at 20:07:39, and August 18th 2018 at 23:19:09.

estimate the sensitivity to incidence angle, wind speed, and distance from shore. The result of this analysis is compiled in Fig. 6.9. Figures Fig. 6.9.a and Fig. 6.9.b indicate a negative impact of wind speed. This can be explained by both the low representation of stronger wind regimes and their similar impact on the SAR observations, as wind speed and precipitation can both increase the sea surface roughness and thus the rain signature contrast. However, the models do not overestimate rainfall at stronger wind regimes, except for the 1 mm/h threshold (Fig. 6.9.a). Thus, the decrease in performance is thought to be due to the disappearance of the rain signature under high wind speed. In contrast, the Koch filters show an overestimation (and thus lower performance) of the rain rate at low wind speed. The incidence angle (Fig. 6.9.c) shows that the F1 scores with the pseudo rain map are stable for the 1 mm/h and 3 mm/h thresholds, but that the prediction of rainfall over 10 mm/h is negatively impacted by the low incidence angle. The U-Net model is also less impacted by the coast (Fig. 6.9.d) and can thus be used up to two kilometers off the coast while the Koch filter is affected up to six kilometers off the coast.

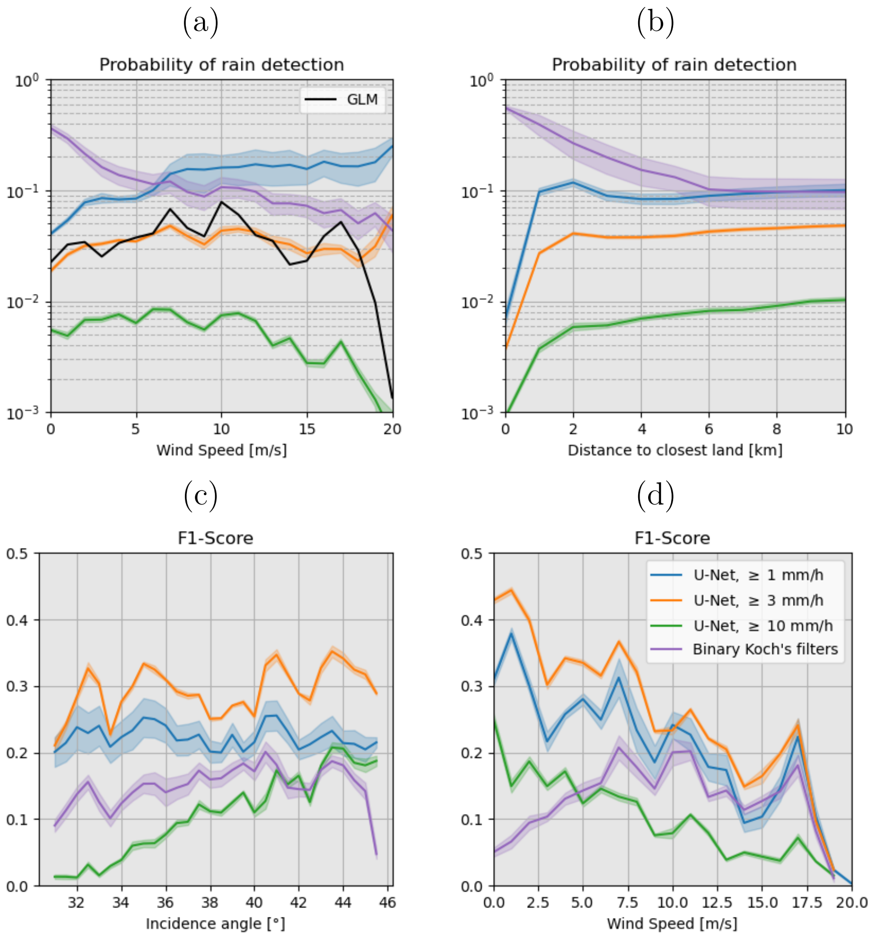


Fig. 6.9: Probability of rain detection as a function of wind speed (a) and distance to the coast (b). F1-score as a function of incidence angle (c) and wind speed (d), using GLM as a rain proxy. The blue, orange and green curve respectively correspond to rainfalls higher than 1 mm/h, 3 mm/h and 10 mm/h. The purple curve correspond to the fine-tuned Binary Koch's filters.

Fig. 6.10 shows examples of collocations between the GLM and Sentinel-1 observations, especially at high wind speeds. The first observation, made September 4th 2019 at 11:09:34 indicates a very close correlation between lightning detection and rainfall estimation from the U-Net architecture, despite wind speeds above 15 m/s in most images. The second observation, made on August 27th 2020 at 00:09:33 shows an even higher wind speed, as it was acquired over Hurricane Laura. A smaller proportion of pixels are predicted to indicate precipitation greater than 10 mm/h. In the lower left corner, where the wind speed is highest, a series of 3 mm/h rainfall areas are detected although they were not recorded on the lightning map.

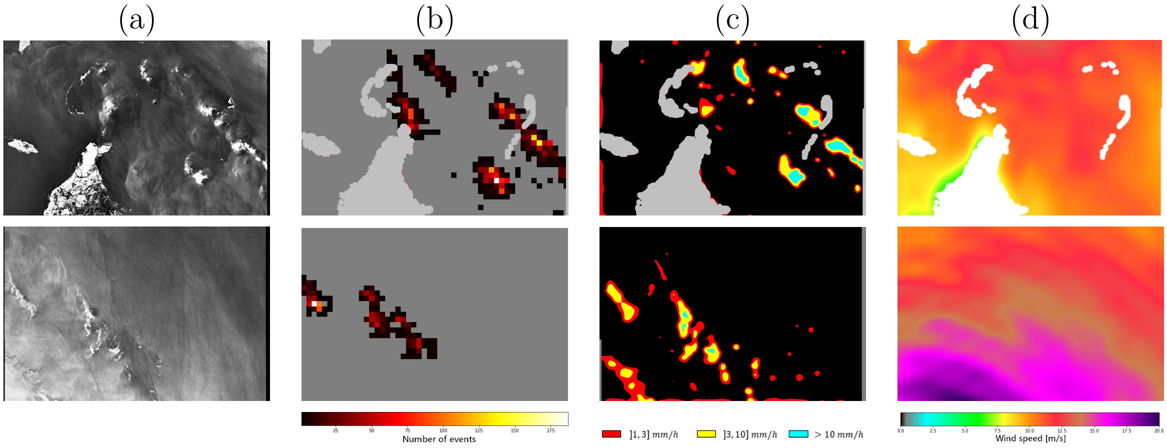


Fig. 6.10: Collocation of Sentinel-1 SAR observations (a) with GOES-16 GLM lightning clusters (b), U-Net-based NEXRAD emulation (c) and ECMWF wind speed (d) in meters per second. The first line shows an observation from September 4th 2019. The second line shows Hurricane Laura on August 27th at 00:09:33.

6.6 Conclusion

The monitoring of rain over the oceans is a key challenge for weather modeling and forecasting. This is particularly important for flood mitigation in coastal areas. While land-based sensors cannot monitor the open ocean, the satellite-derived retrieval of rain rate remains a challenge, especially at high resolution, despite the variety of rain-impacted and rain-measuring spaceborne instruments. In this respect, the effect of precipitation on satellite SAR observations of the sea surface has been widely documented.

This study demonstrates that deep learning opens new avenues for the estimation of sea surface rain rate at high resolution from satellite SAR observations. We exploit a state-of-the-art image-to-image translation architecture, namely a U-Net. The training scheme relies on a collocated dataset of NEXRAD weather radar data and Sentinel-1 SAR observations. We report an accurate segmentation of rainy areas at sea surface and satisfactory ability to discriminate rain between 1 mm/h and 3 mm/h, 3 mm/h and 10 mm/h and above 10 mm/h. The proposed approach outperforms previous work based on Koch filters and points out the importance of a registration-based preprocessing of the training dataset.

We further assess the relevance of the retrieved rate rain estimation at sea surface with respect to another proxy of local precipitation, namely the frequency of lightning events detected by the GOES-16 GLM.

This analysis indicates an impact of incidence angle and wind speed, with performance decreasing at low incidence angles and/or high wind speeds. Future work could therefore benefit from the generalisation of the proposed approach to other SAR modes such as WV modes which may involve other incidence angle ranges. It also supports a joint retrieval of wind speed and rain rate at sea surface from Sentinel-1 SAR observations.

- END OF THE ARTICLE -

CHAPTER 7

IMPROVING SAR-DERIVED WIND SPEED ESTIMATION IN HEAVY RAIN CONDITIONS WITH DEEP LEARNING

7.1 Introduction

Chapter 6, the negative impact of precipitation and wind on segmenting rainfall patterns in Synthetic Aperture Radar (SAR) observations was addressed. It was shown that segmentation was hindered at high wind speeds due to the disappearance of the rain signature caused by saturation in the VV channel, and the lack of data for meteorological processes that include both heavy rainfall and high winds.

Geophysical Model Functions (GMF) have been developed to estimate wind speeds from SAR observations, including CMOD3 [119], CMOD4 [179], CMOD5 [69], CMOD5.N [70], CMOD6 [47], CMOD7 [180], C_SARMOD2 [123]. Typically, GMFs use the copolarization channel (usually the VV channel on Sentinel-1 IW) and additional information such as the incidence angle and a prior on the wind speed direction relative to the satellite heading. This prior is typically provided by atmospheric models, but recent research has shown the possibility of estimating it directly from SAR observations [223, 44]. While GMFs have proven effective in estimating wind speeds in most conditions, they work at the pixel level and are unable to capture contextual information. They also assume that most of the backscattering is caused by wind, which leads to overestimating wind speeds when hydrometeors impact the ocean surface and increase surface roughness. On the other hand, Convolutional Neural Networks (CNN) have a non-null receptive field, as explained in Section 1.3.2, and have been demonstrated to detect rainfall in Chapter 6. Thus, they are able to use contextual information to estimate wind speeds despite the contamination of rain signatures.

As discussed in earlier chapters, one of the difficulties of using deep learning models is the need for large datasets. This is particularly true for the current problem, as it would require observations of various rain events under different wind speed conditions. However, it is possible to collocate Synthetic Aperture Radar (SAR) observations with atmospheric models, which are spatially available globally and independent of the SAR system. Using the rain detector from Chapter 6, we demonstrate that it is therefore possible to build a dataset of both rain and rainless situations associated with modeled wind speed. The deep learning model proved to be closer to the ground truth (i.e., the modeled wind) both at high and low rainfall. These results are confirmed by in-situ measurements from buoys, which are used to measure the performance of the wind estimator in the real world.

7.2 Dataset

The SAR measurements used in the chapter originate from 19978 IW acquired globally between the March 03th 2018 and the February 23th 2022. Each of these observations covers approximately 44 000 km^2. The resolution is set at 100 m/px, dowscaled from the GRDH L1 products at 10 m/px.

All of these observations are segmented by the rainfall estimator built in 6. A previous study conducted using a global Sentinel-1 dataset found only 2304 partial collocations with the GPM-DPR out of 182153 IW. 'Partial collocations' refers to instances where at least 20x20 km of a swath is observed by the spaceborne weather radar 20 minutes before or after the SAR observation. Acquiring global information on rain in large quantities is therefore difficult. Additionally, misalignment that could be manually resolved on small-scale datasets - such as the Sentinel-1/NEXRAD collocations considered in chapter 6 - is difficult to correct on tens of thousands of patches. Therefore, SAR-based rain estimation is preferred to maximize the number of available observations and simplify the collocation process.

Ancillary information, such as the incidence angle, the wind speed derived from the GMF and atmospheric model are retrieved from Sentinel-1 L2 products. In particular, the atmospheric model contains both the wind speed used as groundtruth and the wind direction, used as a prior. It has a spatial resolution of 0.25x0.25 degrees and a temporal resolution of 3 hours.

7.3 Methodology

The aim of this section is to present the methodology for building the rain-invariant wind speed estimator. First, we will describe the deep learning architecture of the model. Next, we will discuss the creation of the dataset, specifically designed to have a large number of rain examples. Finally, we will describe the evaluation procedure.

7.3.1 Deep Learning Model

The architecture used in this chapter is the U-Net architecture [234] , which has already been studied in previous chapters and is depicted in Fig. 2.3. U-Net is an auto-encoder architecture that is fully convolutional, meaning it maintains the property of translation equivariance (the translation of the input results in the translation of the output). Additionally, skip connections between the encoder and the decoder facilitate training, particularly by reducing the [150]. The output of the model always contains a single convolution kernel, which is activated by the ReLU function to ensure that the prediction is in the interval $[0, +\infty[$. All the convolution kernels in the hidden layers are also activated by ReLU functions. The model is set to take input of 256x256 pixels during training, but since the weights only describe convolution kernels, it is possible to use the model in inference on images of any shape if the input resolution is 100 m/px.

Variants of the model are trained with different numbers of input channels. The modifications to the architecture are limited to the first convolution kernel, which is defined as a kernel of size $(3, 3, n)$ with n being the number of input channels.

7.3.2 Training procedure

Particular attention is given to the creation of the dataset. In particular, we strive to satisfy the following assumptions:

1. Accuracy: the output -in our case, the wind speed from the atmospheric model- must be close to the real-world wind speed, which is unknown.
2. Representativeness: the patch distribution in the dataset must reflect the real-world distribution.
3. The subset coherence: the training, validation and test sets follow the same distribution.

4. The balancing: the low representation of rain patches in the real-world distribution should be corrected to ensure the model works regardless of the rain rate.

5. The isolation: no information is leaked between the training, validation and test set.

Each of these points is presented in the following subsections. It is important to note that the definition of representativity (Assumption 2) is somewhat ambiguous, as the parameters that determine its distribution are not explicit. Furthermore, this concept contradicts the idea of balancing (Assumption 4) as the latter implies an over-representation of rain patches. In the following, the representativeness is only considered for the distribution of wind speeds.

Restriction to *a priori* Accurate Model Wind Speeds

Despite its limitations in resolution (0.25x0.25 degree spatially, 3h temporally) and can be unable to accurately depict the real wind field, the atmospheric model is available for every SAR observation. We are therefore able to restrict the dataset to samples where we estimate the atmospheric model to be accurate (Assumption 1).

The number of pixels contained in $\mathcal{A}^-$ is denoted as $\#_{\mathcal{A}^-}$, and the discrepancy between the geophysical model function and the atmospheric model on rain-free pixels is denoted as $\Delta_{\mathcal{A}^-}$.

$$\mathcal{A}^+ = \{x : Rainfall(x) >= 3mm/h\} \tag{Eq. 7.1}$$

$$\mathcal{A}^- = \{x : Rainfall(x) < 3mm/h\} \tag{Eq. 7.2}$$

$\#_{\mathcal{A}^-}$ is the number of pixels contained in $\mathcal{A}^-$. We then compute $\Delta_{\mathcal{A}^-}$ the discrepancy between the geophysical model function and the atmospheric model on rain-free pixels.

$$\Delta_{\mathcal{A}^-} = \#_{\mathcal{A}^-}^{-1} \sum_{(x,y)\in\mathcal{A}^-} (ECMWF(x,y) - CMOD5N(x,y))^2 \tag{Eq. 7.3}$$

In our experiments, the threshold for $\Delta_{\mathcal{A}_-}$ was set at less than 1 m/s. All patches that displayed a higher discrepancy between the two wind speed sources were discarded.

Forced Over-representation of Rain Signatures

Since the dataset will be used to train a rain-invariant model, rain patches have to be fed to the model during the training phase. However, rain only appears in a minority of the patches, as shown in Fig. 7.1. To balance the dataset (Assumption 4), a second condition, on the ratio r_+ between the cardinals of $\mathcal{A}^+$ and $\mathcal{A}$ is added. A sample is selected and added to the dataset if it contains significant rain signature. In the following, the terms "rain patch" will be used if $r_+ > 5\%$, else the sample will be called "rainless patch". Under this threshold, 105k rain patches are extracted from the wide swaths (out of 4347k patches).

By construction, the dataset lacks diversity. Indeed, only the patches with predicted rainfall stronger than 3 mm/h are put in it. Thus, the dataset is dependent on hypothetical bias that could affect the rainfall source. Because of this selection bias, the dataset lacks diversity, especially for phenomena that do not often occur with rainfall: biological slicks, low wind areas, etc. To ensure as many meteocean scenes appear, the dataset is filled with rainless patches. We denote:

- n_+ the number of patches containing rain.
- n_- the number of patches without rain.
- P^+ the wind speed distribution on n_+.
- P^- the wind speed distribution on n_-.
- P the wind speed distribution on $n_- \cup n_+$.

Dataset-Level and Bin-Level Balancing

Fig. 7.1 indicates, using collocation between GPM-DPR precipitation measurement and ECMWF atmospheric model, the evolution of the wind speed distribution depending on the rainfall intensity, demonstrating a low impact of the former. At very high rain rates, the wind speeds appear to be usually lower, but this difference is marginal until 10 mm/h. Following Assumption 2 implies that P^- and P^+ should be similar distributions. The performance of the rainfall estimator, on the other hand, is not wind-invariant. The assumption that P^- and P^+ are close does not hold. To avoid bias to be introduced in the model, especially toward the resulting distribution of wind speed, it is necessary

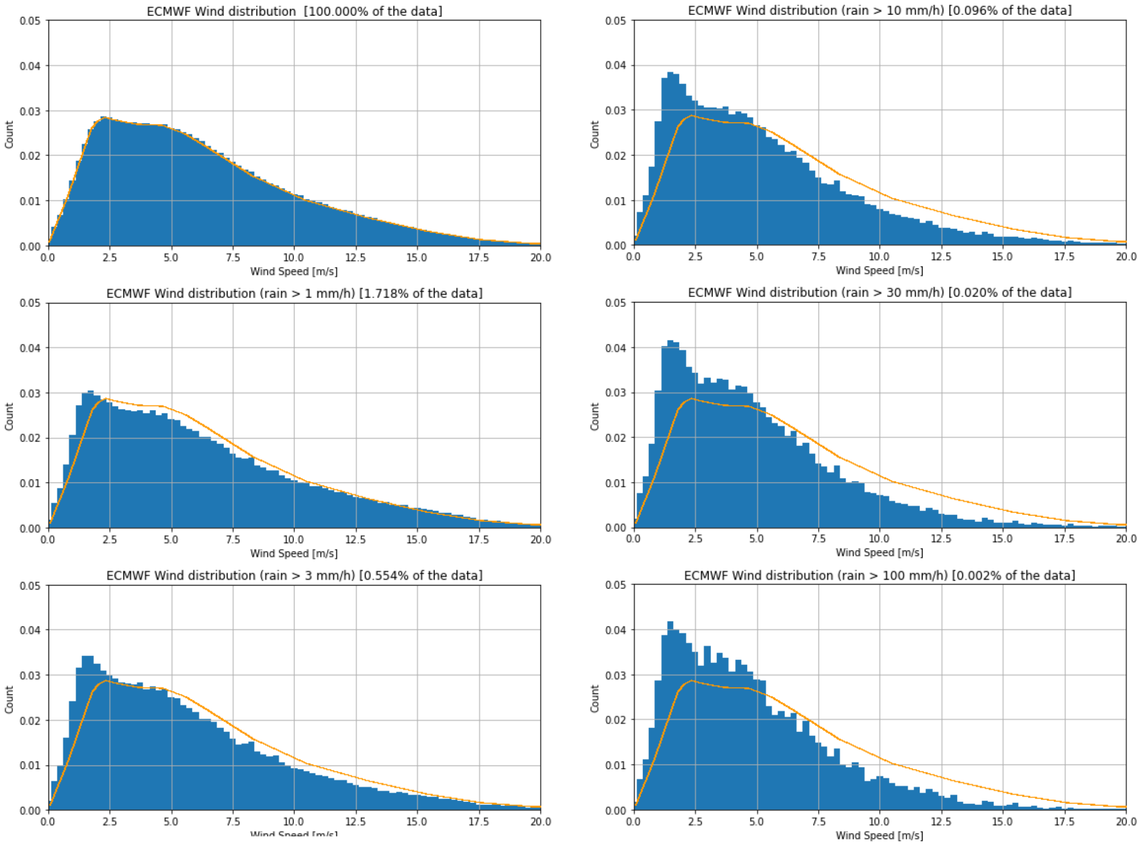


Fig. 7.1: Distribution of ECMWF ERA interim wind speed collocated with surface rain rate from GPM-DPR. The orange curve is the wind distribution regardless of the observed rainfall and therefore correspond to the top-left histogram.

to balance the dataset to a distribution close to P. Considering the variable x and n_+ (resp. n_-) the amount of rain (resp. rain-free patches), two datasets balancing methods are considered:

- The **Dataset-level balancing** ensures that:

$$\forall x, P(x) = \frac{n_+ P^+(x) + n_- P^-(x)}{n_+ + n_-} \qquad \text{(Eq. 7.4)}$$

- The **Bin-level balancing** ensures that:

$$\forall x, P(x) = P^+(x) = P^-(x) \qquad \text{(Eq. 7.5)}$$

For simplicity, we set $n_+ = n_-$.

Dataset-level balancing

Dataset-level balancing aims to select all rain samples and pad the dataset with rainless patches to achieve the original distribution P.

By choosing $n_+ = n_-$, the equation Eq. 7.4 simplifies to $P^-(x) = 2P(x) - P^+(x)$. However, this choice can result in $P^-(x)$ being negative when $P^+(x) > 2P(x)$. To address this, P^- is rescaled through normalization in Eq. 7.6.

$$\widetilde{P}^- = \frac{max(0, P^-)}{\int max(0, P^-)dx} \tag{Eq. 7.6}$$

To build the rainless dataset, random samples are extracted from all the rainless patches. These random samples is selected in the dataset with a probability of acceptance described in Eq. 7.7.

$$p_{accept}(x) = (1 - P(x)) \cdot \widetilde{P}^-(x) \tag{Eq. 7.7}$$

As a result of this operation, the dataset contains 210k patches. The distribution of these patches is shown in Fig. 7.2.a using the wind speed as the balancing parameter x. A discrepancy between $\frac{P^+ + \widetilde{P}^-}{2}$ and P is shown in Fig. 7.2.b. In this balancing scheme, the MAE between $\frac{\widetilde{P}^- + P^+}{2}$ and P is equal to 8.8%. In Fig. 7.2.c shows that, since the distribution P^+ and $\widetilde{P}^-$ are different, for a given wind speed, knowledge of the rain signature contained in the patch gives information on the wind speed. As we saw in Fig. 7.1, this behavior is undesirable, as it arises from bias in the dataset rather than an actual metocean process.

Bin-Level Balancing

The bin-level balancing consists in ensuring $\forall w : P^+(w) = P^-(w) = P(w)$, whereas the previous dataset-level balancing only states that $\frac{P^+(w) + P^-(w)}{2} = P(w)$. Since the number of samples that can be extracted to build P^- is one magnitude higher than P^+, we assume than only rain patches are limited.

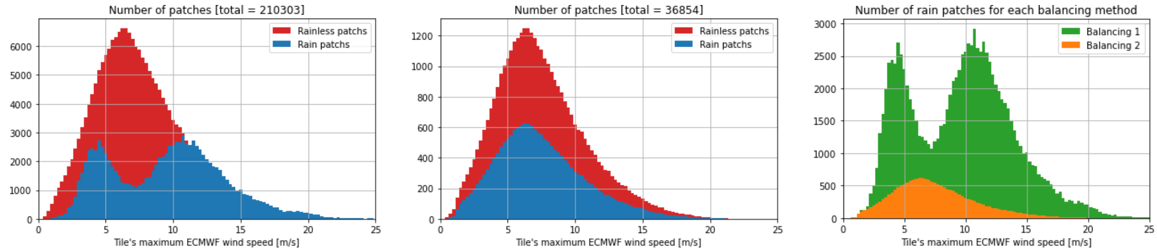
The number $n^*(w)$ of samples of P^+ at wind speed w to be extracted must verify that (1) it is lower than the amount of rain patches at this wind speed and (2) it preserves the distribution P. These two conditions translate as:

$$n^*(w) = P(w) \cdot \alpha < P^+(w) \cdot n_+ \qquad \text{(Eq. 7.8)}$$

where α is:

$$\alpha = n_+ \cdot \min_w \frac{P^+(w)}{P(w)} \qquad \text{(Eq. 7.9)}$$

α denote the ratio of selected data. In our experiment, $\alpha = 17.5\%$, meaning than five out of six samples were removed.



(a) (b) (c)

Fig. 7.2: Distribution of the wind speed depending on the balancing method. In (a), dataset is constraint to have equal proportion of rainless and rain patches at wind speed level. In (b), constraint is applied at the dataset level. (c) indicates the trimming resulting of the wind speed level constraint.

Theoretically, the balancing should be performed for any ancillary parameters, including the incidence angle, the wind direction, the latitude, the longitude... This can be easily done by denoting the joint variable $x = (w, INC, WDIR, LAT, LON)$ and applying the previous formula to x rather than w. However, it can be seen that the bin-level balancing on joint variables would severely decrease the number of samples since 82.5% of the data were removed for balancing the wind speed only. Under the dataset-level balancing, however, the difficulty arises when $\frac{n_+}{n_-}P^+(X) > \frac{n_+ + n_-}{n_-}P(X)$ since it constrains all the patches at the ancillary parameter combination X to contain rain. It also increases the difference between $\frac{\widetilde{P^-} + P^+}{2}$ and P. In the case of wind speed balancing, the MAE between these distributions was 8.8%.

Training, Validation and Test Set Division

Similarly to the previous segmentation tasks, training, validation and test sets are created. The training set is used for the gradient descent. The validation set to prevent overfitting and ensure the model is able to generalize out of the training set. The test set, not accessible during the training, is only used for the performance measurement. Therefore, the test and validation set must avoid any kind of data leak between the patches that would hide generalization issues, which is the Assumption 5. We consider two sources of data leaks:

- During the initial patch extraction, the stride was smaller than the patch size. A smaller stride has the advantage of ensuring that a rain signature in the corner of one patch, which would otherwise be useless because of its proximity to a border, would also appear in the centre of another patch. However, it means that two adjacent patches have overlapping areas and should appear in the same subset.
- Non-adjacent patches from the same swath could contain similar content as the process they observed would be caused by the same metocean phenomenon.

Because of both of these potential data leaks, the training, validation and test sets are divided at the swath-level. This means that if a patch from one swath is in one subset, all patches from the swath must belong to the same subset.

The second property that has to be verified at this step is the conservation of the underlying metocean distribution, which derives from Assumption 3. Denoting P_{train}, P_{val} and P_{test} as the distribution of each subset, the respective errors e_{train}, e_{val} and e_{test}, defined as $e_A = MAE(P, P_A)$, must be minimized. This minimization problem is similar to the 3-partition problem. To solve this optimization problem quickly, we implement the stochastic brute forcing method describes in algorithm 1. Being of time complexity linear with the number of swaths, it quickly selects candidate distribution for both the validation and the test sets and minimize the harmonic mean of both MAE.

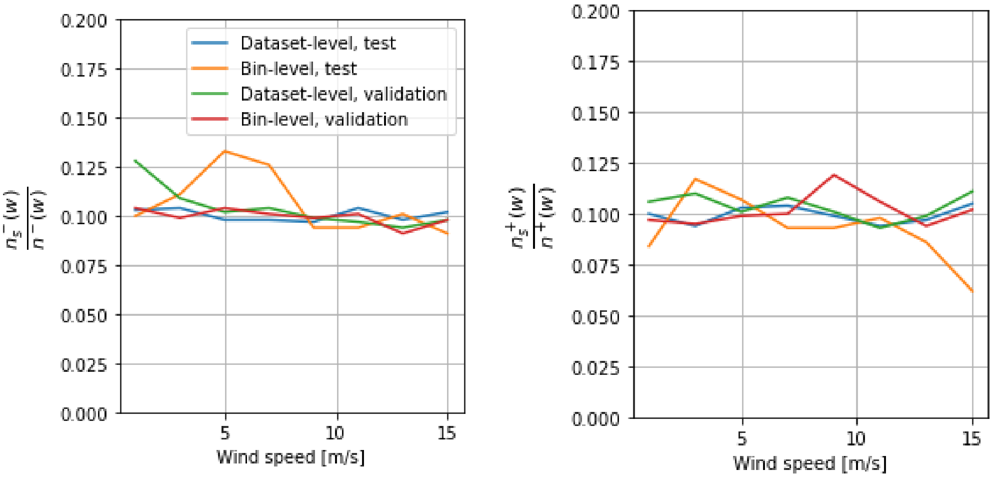
Fig. 7.3 indicates the result of Algorithm 1. Both the validation and the test sets contain around 10% of the data. The remaining 80% are contained in the training set. The test set produced for the bin-level balancing slightly diverges from the original distribution, with less high wind speeds rain patches and more rainless patches at 5 m/s.

At the end of the building, the dataset-level balancing scheme return respectively 168349, 20944 and 21010 patches for respectively the train, test and validation subsets for 14169, 1763 and 1763 IW. Under the bin-level balancing scheme, the train, test and validation subsets contain respectively 30422, 3150 and 3282 patches from 10531, 1319 and 1319 IW.

Algorithm 1: Stochastic brute forcing

```
input  : The list $L$ of all swath $i$ and their patch distribution $P_i$.
output: $L_{val}$ and $L_{test}$, the list of the swaths contained in the validation and test
        sets.

$e_{min} = +\infty$;
$n_1 = \text{len(L)}/10$;
$n_2 = \text{len(L)}/3$;
for $i \leftarrow 0$ to $1000000$ by $1$ do
    $n = \text{random.integer(min=}n_1\text{, max=}n_2\text{)}$;
    $candidate = \text{random.choice(}L\text{, size=}n\text{, replace=False)}$;
    $c_{val} = candidate[: n/2]$;
    $c_{test} = candidate[n/2 :]$;
    $e_{val} = MAE(P^+, P^-_{c_{val}}) + MAE(P^-, P^-_{c_{val}})$;
    $e_{test} = MAE(P^+, P^+_{c_{test}}) + MAE(P^-, P^-_{c_{val}})$;
    $e = \frac{2 \cdot e_{val} \cdot e_{e_{test}}}{e_{val} + e_{test}}$;
    if $e < e_{min}$ then
        $L_{val} = c_{val}$;
        $L_{test} = c_{test}$;
        $e_{min} = e$;
    end
end
```



(a) (b)

Fig. 7.3: The percentage of the whole dataset contained in the validation and test subsets, for rainless patches (a) and rain patches (b). The train subset is not represented as it is the complement of the validation and test subsets.

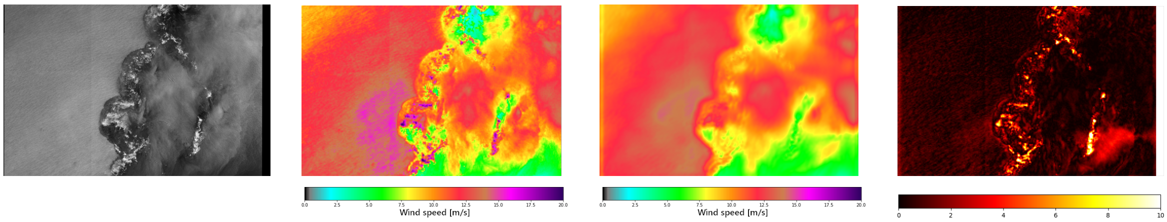


Fig. 7.4: Observation from the November 11$^{\text{th}}$ 2018 at 04:56:47. From left to right: Copolarization SAR observation, GMF output, deep learning output and absolute difference between the two.

7.3.3 Evaluation procedure

To assess the impact of each input channels, various variants of the model are trained:

1. I uses only the VV channel.
2. II uses the same inputs as the GMF: the VV channel, the incidence angle and the *a priori* wind speed direction.
3. III uses the same inputs as II and the VH channel.
4. IV uses the same inputs as III and the wind speed prior.
5. V uses only the wind speed prior. We note that this variant implicitly use the VV channel, though at a resolution of only 1 km/px.

Each architecture is trained five times to reduce the impact of a lucky initialization on the result of the evaluation. The results are therefore recorded as the mean and the standard deviation over these five independent training.

The results are compared using the Root Mean Square Error (RMSE) and the Pearson correlation coefficient (PCC). The latter has the advantage of being a normalized measurement of the covariance and not being affected by the bias. Results are computed both the groundtruths from the atmospheric model, which has the advantage of providing a large test set, and for collocations with buoys, which have a good temporal resolution and are *in-situ* measurements.

7.4 Results

7.4.1 Benchmarking experiments

The performances of the model against ECMWF are computed on the test subset for input variant and baseline the GMF. Results of this analysis are contained in Table 7.1. It seems that the most important input is the GMF itself, as both IV and V have better results than the other variants. I, II and III aren't able to obtain better results than the GMF except under strong rainfall, even though II and III have access to all the channels used by the GMF.

The choice of the balencing scheme appears to be secondary as the performances over both datasets are similar, despite the difference of test set over which they were computed.

MODEL AND CHANNELS	RAIN RATE	RMSE		PCC	
		Dataset-level balancing	Bin-level balancing	Dataset-level balancing	Bin-level balancing
I [VV]	$[0, 1[mm/h$	1.38 [0.016]	1.38 [0.008]	89.9% [0.19%]	89.9% [0.17%]
	$[1, 3[mm/h$	1.64 [0.046]	1.29 [0.034]	92.7% [0.39%]	94.4% [0.28%]
	$[3, 10[mm/h$	1.59 [0.038]	1.43 [0.077]	♢ 92.7% [0.38%]	♢ 94.7% [0.52%]
	$\geq 10mm/h$	♢ 2.12 [0.052]	♢ 2.01 [0.092]	♢ 81.7% [1.03%]	♢ 84.6% [1.24%]
II [VV, INC, GMF]	$[0, 1[mm/h$	0.87 [0.009]	0.87 [0.005	96.2% [0.06%]]	96.1% [0.04%]
	$[1, 3[mm/h$	1.04 [0.086]	1.05 [0.093]	97.2% [0.47%]	96.2% [0.70%]
	$[3, 10[mm/h$	♢ 1.19 [0.053]	♢ 1.24 [0.096]	96.2% [0.37%]	♢ 95.8% [0.64%]
	$\geq 10mm/h$	♢ 2.27 [0.151]	♢ 1.74 [0.129]	♢ 81.7% [1.89%]	♢ 88.6% [1.27%]
III [VV, VH, INC, WDIR]	$[0, 1[mm/h$	0.83 [0.002]	0.86 [0.013]	96.5% [0.02%]	96.2% [0.09%]
	$[1, 3[mm/h$	0.93 [0.020]	0.92 [0.018]	97.7% [0.07%]	97.2% [0.09%]
	$[3, 10[mm/h$	♢ 1.09 [0.022]	♢ 1.06 [0.033]	♢ 96.7% [0.07%]	♢ 97.0% [0.21%]
	$\geq 10mm/h$	♢ 2.13 [0.050]	♢ 1.70 [0.106]	♢ 83.9% [0.66%]	♢ 88.9% [1.34%]
IV [VV, VH, INC, WDIR, GMF]	$[0, 1[mm/h$	★ 0.64 [0.007]	★ 0.65 [0.006]	★ 97.9% [0.04%]	★ 97.9% [0.02%]
	$[1, 3[mm/h$	★ 0.63 [0.015]	★ 0.65 [0.007]	★ 98.9% [0.03%]	★ 98.6% [0.03%]
	$[3, 10[mm/h$	★ 0.78 [0.040]	★ 0.77 [0.005]	★ 98.4% [0.08%]	★ 98.4% [0.03%]
	$\geq 10mm/h$	★ 1.63 [0.162]	★ 1.37 [0.078]	★ 90.9% [1.17%]	★ 93.7% [0.44%]
V [GMF]	$[0, 1[mm/h$	♢ 0.67 [0.003]	♢ 0.69 [0.003]	♢ 97.7% [0.02%]	♢ 97.6% [0.01%]
	$[1, 3[mm/h$	♢ 0.68 [0.005]	♢ 0.71 [0.005]	♢ 98.8% [0.01%]	♢ 98.3% [0.01%]
	$[3, 10[mm/h$	♢ 0.88 [0.021]	♢ 0.86 [0.009]	♢ 98.0% [0.03%]	♢ 98.0% [0.03%]
	$\geq 10mm/h$	♢ 1.94 [0.087]	♢ 1.56 [0.075]	♢ 87.9% [0.66%]	♢ 91.5% [0.51%]
GMF	$[0, 1[mm/h$	0.77	0.78	97.0%	96.9%
	$[1, 3[mm/h$	0.84	0.91	98.1%	97.3%
	$[3, 10[mm/h$	1.25	1.43	96.5%	95.2%
	$\geq 10mm/h$	4.65	5.32	52.5%	55.8%

Table 7.1: Comparison of the five variants of the model and the two datasets. RMSE and PCC are computed on the respective test set and for five training with random initialization. Results are given with mean and standard deviation. The best result for each metric is indicated by ★. Results better than the baseline are indicated by ♢.

7.4.2 Application to SAR observation with groundtruthed *in situ* data

However, ECMWF wind speeds are reanalysis data and not in situ data. Those can be obtained with anemometers aboard buoys. Using the dataset build in [136], 6030 IW were collocated with 42 NDBC buoys. These data were collected and made freely available by NOAA/NDBC. It resulted in 3614 collocation points. The rain prediction model predicted 3548 to be rainless, 66 to have been recorded during a rainfall higher than 1 mm/h and 13 at more than 3 mm/h. As the height of these measures varies (most are around between 3.8 m and 4.0 m above the ocean level). Following [136], the SAR inversion and the deep learning prediction are both normalized to the altitude of the corresponding in situ measurement using an exponential law [155]:

$$w(h) = \left(\frac{10}{h}\right)^{0.11} \qquad \text{(Eq. 7.10)}$$

Table 7.2 indicates that, on rainless patches the performances of the IV model are similar to that of the GMF itself. For the model trained on the dataset A, the RMSE decrease from 1.46 m/s to 1.38 m/s, and the PCC increases from 93.9% to 94.0% both improvements are considered negligible. One important difference is contained in the bias of the linear regression associated with the distribution: this bias indicates that the GMF overestimates the wind speed by 0.31 m/s and reach 0.44 m/s for Ψ_A. This is believed to be caused by both the low number of samples at wind speed lower than 1 m/s and by the bias from ECMWF data. According to Table 7.2, the ECMWF atmospheric model overestimate the wind speed given by the buoys by 1.17 m/s. However, for points with rainfall higher than 1 mm/h, the RMSE and the bias increase dramatically for the GMF-baseline but are mitigated with the deep learning models. Training the model on the dataset A (dataset-level balancing) appear to be more efficient as Ψ_A have slightly better metrics than Ψ_B (trained under the bin-level balancing scheme), especially on the PCC.

Here, we observe two cases where rainfall was detected on the buoy position at the time of the observation.

2017-01-08 01:58:19 at NDBC 46054

The observation from 2017-01-08 01:58:19 covers the north of the Californian Channel Islands (Fig. 7.5.a). Several meteorological buoys are dispersed over the channel, notably

		Number of points	IV$_A$	IV$_B$	GMF	ECMWF
Bias	$< 1mm/h$	3548	0.44	0.47	★ 0.31	1.17
	$\geq 1mm/h$	66	★ 0.97	♢ 0.98	1.18	♢ 1.10
	$\geq 3mm/h$	13	♢ 0.49	★ 0.47	2.78	♢ 1.26
RMSE	$< 1mm/h$	3548	★ 1.39	1.53	1.46	2.13
	$\geq 1mm/h$	66	♢ 1.98	★ 1.97	2.56	♢ 2.14
	$\geq 3mm/h$	13	★ 1.78	♢ 1.89	3.42	♢ 2.23
PCC	$< 1mm/h$	3548	★ 94.1%	93.6 %	94.0%	86.8%
	$\geq 1mm/h$	66	★ 95.9%	♢ 95.7%	92.8%	92.3%
	$\geq 3mm/h$	13	★ 95.6%	♢ 93.0%	89.6%	87.5%

Table 7.2: RMSE, PCC and bias of IV, the GMF and ECMWF against the buoys, for each rain prediction. The best result for each metric is indicated by ★. Results better than the baseline is indicated by ♢. IV$_A$ indicates the dataset-level balancing and IV$_B$ the bin-level balancing.

NDBC 46054 and NDBC 46053, which are indicated as red dots. The wind speed over the area is mostly around 6 m/s but a squall line appears at the position of NDBC 46054 and span over dozen kilometres. Rain signature appears clearly on the southern -or upper, as the image is in sensor-geometry- half of the front, which is detected by the rain detector build in chapter 6 (Fig. 7.5.c). In the northern half, the backscattering is still high, but the rain signature is difficult to read. The GMF indicates a very high wind speed, higher than 20 m/s (Fig. 7.5.d), while the deep learning model attenuates this value to between 6 m/s and 8 m/s (Fig. 7.5.e). The southern half of the front, where rain signatures are visible, is the most attenuated.

For NDBC 46054, only one measurement of wind speed and direction per hour is available. It recorded a wind speed of 6.3 m/s eight minutes before the SAR observation. The GMF and the deep learning estimated the wind at respectively 15.1 m/s and 5.9 m/s. Though the temporal resolution of NDBC 46054, NDBC 46053 is recording data every ten minutes. Moreover, the gust front appears to be progressing toward the right part of the observation. It appears on the time serie in Fig. 7.5 as a large variation of the wind direction between 02:40:00 and 03:00:00. The variation of the wind speed seems to precede the variation of the direction by first increasing then decreasing to a lower wind regime. On NDBC 46053, the GMF and the deep learning model are in agreement at 4.5 m/s, which is slightly lower than the in situ data at 5.4 m/s. Since the distance between NDBC 46054 and NDBC 46053 is approximately 60 km, the progression of the gust front can be estimated to be around 90 km/h. Its width being around 5 or 6 km, the whole system is passing the buoys in three minutes. This means that even NDBC 46053 may not have been able to correctly estimate the wind speed due to the lack of temporal

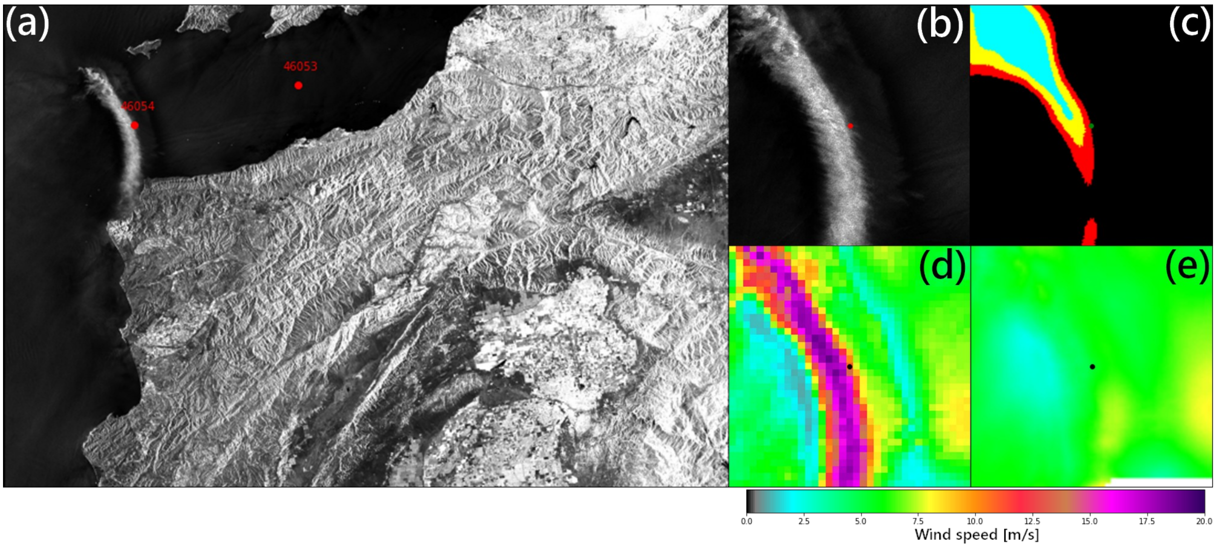


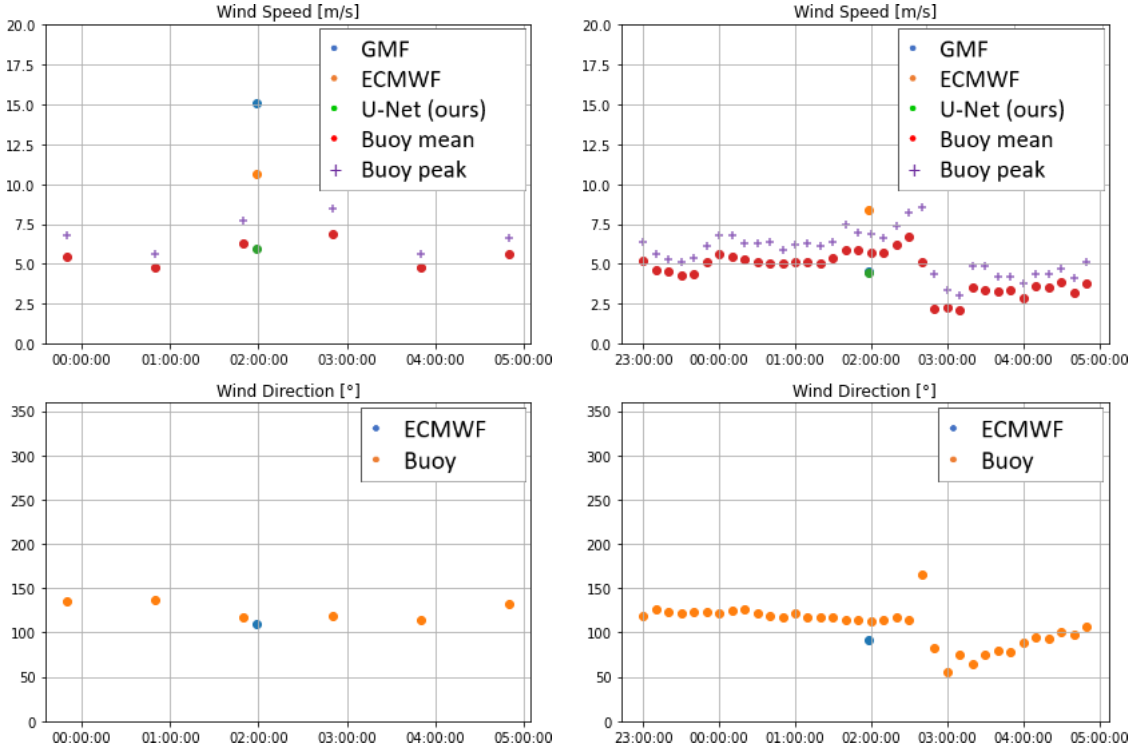
Fig. 7.5: Observation from the January 08th 2017 at 01:58:19 in VV channel (a), zoom on an area of 35x35 km centred on the buoy NDBC 45054 (b), segmentation of the rain rate (c), wind speed given by the GMF (d) and by the Ψ_A (e).

resolution. However, it has to be noted that even the gust speed at NDBC 46054, defined as the maximum wind speed over height seconds, does not record a speed higher than 9 m/s.

SAR-20191006T232853 NDBC-41009

The observation from 2019-10-06 23:28:53 was recorded on the east-coast of Florida. Though most of the swath cover the marshes around Orlando and Cap Canaveral rather than the ocean, convective precipitation can be observed in the right part of the image (Fig. 7.7.a). The cells are progressing downward (north-north-east) as indicated by the stronger gradient of the convective front. As the wind from the convection opposed the underlaying wind regime, an area of lower wind speed appears as an area of lower backscatter. Then, rain signature is clearly visible south of NDBC 41009 (Fig. 7.7.b). The GMF is impacted by this rain signature and estimate a local very high wind speed (Fig. 7.7.c). The deep learning model is less affected by the rain signature, but also appear to blur the low wind speed area (Fig. 7.7.d).

The time serie from NDBC 41009 in Fig. 7.8 indicates that the lower backscattering was indeed caused by a drop of the wind speed rather than a direction change, as the latter does not significantly change over the passage of the convective cell (possibly because the



(a) (b)

Fig. 7.6: Time series of the NDBC buoy wind measurements around January 08th 2017 01:58:19 for NDBC 46054 (a) and NDBC 46053 (b), and the estimation from the GMF (indicated as owi), IV$_A$ (U-Net) and the atmospheric model (indicated as owiEcmwf).

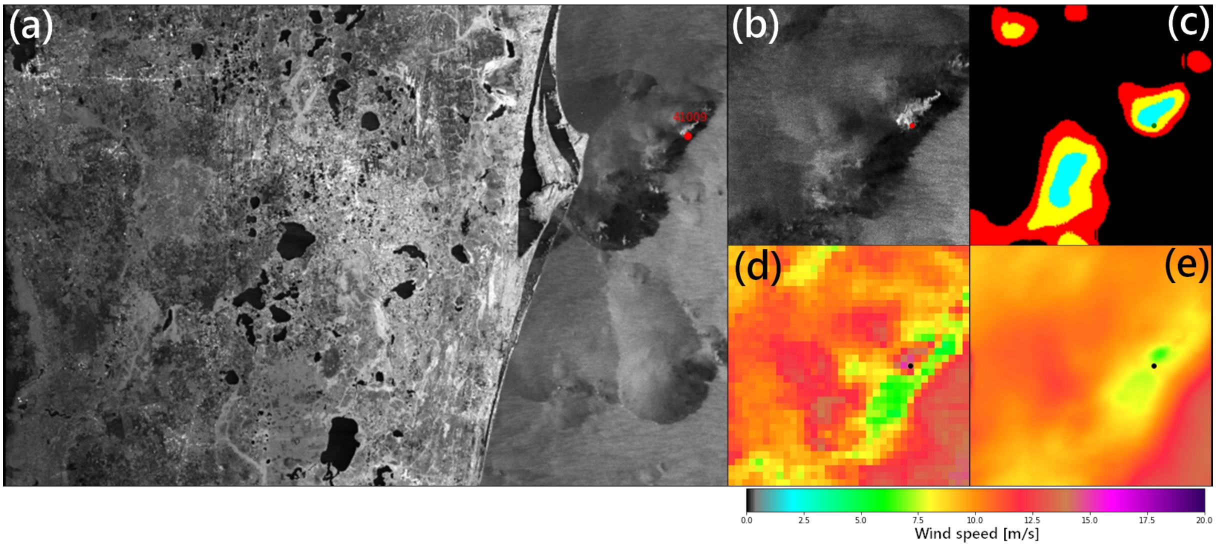


Fig. 7.7: Observation from October 06th 2019 at 23:28:53 in VV channel (a), zoom on an area of 35x35 km around the buoy NDBC 41009 (b), segmentation of the rain rate (c), wind speed given by the GMF (d) and by Ψ_A (e).

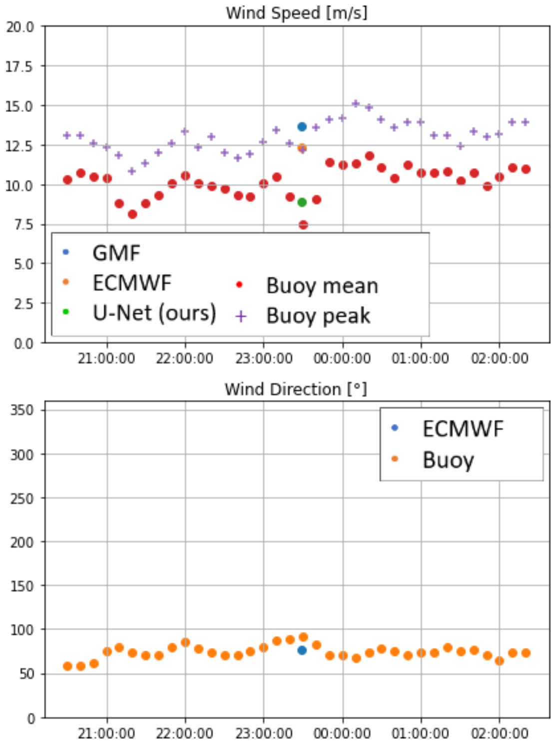


Fig. 7.8: Time series of the NDBC buoy wind measurements around October 06th 2019 at 23:28:53 for NDBC 41009, and the estimation from the GMF (indicated as owi), IV_A (U-Net) and the atmospheric model (indicated as owiEcmwf).

underlying wind regime is rather strong). It does record a sudden drop of the wind speed to 7.5 m/s one minute after the SAR observation, whereas the GMF and the deep learning were estimating respectively 13.7 m/s and 8.9 m/s.

7.5 Conclusion

The study in Chapter 6 highlighted that precise rain signatures could be extracted automatically from SAR observations. Heterogeneity masking was also conducted in [232]. However, it is possible to go further and estimate the wind speed hidden behind the SAR signature, which is akin to in-painting. This can be done by relying on an architecture similar to the other problems studied in this document. In particular, it does not rely on a secondary input channel providing rain information as the ability to detect rain-induced overestimation of the wind speed comes from the dataset rather than the architecture. Once trained, it is therefore possible to use this wind speed estimator as a standalone model.

Collocations with buoy *in-situ* measurement indicate that the model outperforms the current Geophysical Model Function even on rainless areas, decreases the Root Mean Square Error of 4.8%. The performance difference increases with the rain rate, where the GMF is known to overestimate the wind speed. The RMSE decreases by 23% and 48% for rain respectively higher than 1 mm/h and 3 mm/h. However, since the buoys have a time resolution of ten minutes, some quick sub-mesoscale processes, such as gust fronts, are difficult to register. The limited spatial range of the buoys is also an issue to retrieve observations of rare phenomena. Future work should address these concerns.

CONCLUSION

1 Contributions

The segmentation of meteorological and oceanic processes is a crucial topic to understand the meteorological and oceanic conditions. The diversity in their nature, scale, appearance, and the specifics of the sensors capable of detecting them require consideration of their unique qualities when designing a machine learning method. Although the manuscript focuses on the segmentation of meteorological and oceanic processes, Synthetic Aperture Radar (SAR) and Deep Learning can be applied to a wide range of tasks, such as categorization [207], despeckling [103, 50], and super resolution [51, 162, 102]. By limiting ourselves to segmentation problems, we observed that, from a statistical perspective, the estimation of sea ice concentration, rainfall, and wind speed pose different challenges, with respectively a bimodal distribution with modes at the two limits of a $[0, 1]$ interval (in Chapter 5) and unimodal distributions on $[0, +\infty]$ with either a null (in Chapter 6) or non-null mode (in Chapter 7).

Each of these tasks therefore needs for a dedicated solution. The semantic segmentation using the ten classes from TenGeoP-SARwv in Chapter 3 was subject to class unbalance, which led to the formulation of a balanced loss. In the case of the segmentation of *Convective Cells* in Chapter 4, a dynamic loss varying during the training was designed to be able to segment convective cells as well as their fronts. The segmenter architecture was also enhanced to take into account a multiscale aspect and increase the receptive field of the model. In Chapter 5, the regression of a single value, the Sea Ice Concentration, is simplified as a multiclass segmentation, though it uses an additional regression layer embedded in the model to constrain the subsequent categorization layer. Chapter 6 reuses this simplification of regression into a segmentation task to facilitate the balancing of stronger and scarcer rainfall. The Chapter 7 shifts the balancing considerations by considering them at the dataset creation rather than in the architecture of the model or the loss to minimize.

Another common focus is the quality of the annotations. Chapter 3 explores the use of weakly supervised frameworks for the segmentation of metocean processes and concludes that a small number of high-quality annotations is better than a large number of imprecise annotations. This conclusion drove the construction of manually annotated datasets used in Chapter 2 and 4. However, the annotation process for metocean conditions is time-consuming, making it difficult to extend the process to capture the full range of metocean variations. In Chapter 5, the dataset was obtained *via* collocations with manual annotations from the Norwegian Meteorological Institute. The objective of high-quality annotations also appears in Chapter 6 since the existing NEXRAD collocations were manually checked and realigned against the rain signatures on the SAR observations. In Chapter 7, the dataset is built to specifically select samples where the SAR-based wind speed prior was close to the atmospheric model, though this condition is only enforced on rainless pixels.

Each of these segmentation tasks has similar difficulties, but the considerations on the model architectures, loss functions and datasets have to be carefully studied to fine-tune the solutions. This is especially true when it comes to validating the model performances. In data-driven frameworks, a part of the dataset, namely the test set, is reserved for measuring the performances on data independent of those on which the model was trained. However, because of the aforementioned difficulty in retrieving data over diverse metocean condition, it is often necessary to build a secondary dataset for a second validation. In Chapter 2, the validation of the *Biological Slicks* segmentation was performed using geologic and oceanographic priors -namely the presence of cold seeps or of gyres- as well as comparison with measurements of ocean colour from Sentinel-3. Chapter 4 compares the segmentation of the *Convective Cells* to the measurements from infrared sensors aboard GOES-16 and GOES-17. Chapter 6 aims at measuring the performances on winds too scarce in the training set by collocating GLM, using its lightning detection capability as a proxy of rain information. Finally, Chapter 7 used *in situ* wind speed measurements to validate that learning from an atmospheric model was possible. This validation process helps to understand the inner behaviour of the models and detect the situations that would raise incorrect predictions.

Considering these points, we can answer, at least partially, to the questions raised in the foreword.

It is possible to segment metocean processes on Synthetic Aperture Radar observations?

The segmentation of several processes appears to be possible. In particular, segmentation of Sea Ice, biological slicks, convective processes, rainfall, and the estimation of the wind speed all obtain satisfying results and prove the possibility to use supervised learning and convolutional neural networks to obtain the segmentations. The use of the U-Net architecture for most segmentation tasks led to believe that the segmentation of additional metocean processes could be possible with a similar architecture if an adequate dataset is provided.

Is it possible to separate processes appearing on the same pixels?

The rain-induced wind speed estimation is an example of two metoceans processes appearing simultaneously on the same pixels. Chapter 7 proved that it was possible to mitigate this overestimate, though only partially. The main difficulty was to retrieve accurate groundtruth for the wind speed. This issue was solved by comparing the SAR-based wind speed and the atmospheric model on rainless pixels. This methodology is therefore only possible when the observation contains non-overlapping areas. The extension to other co-occurring phenomena would therefore be more difficult than the simpler segmentation of Question 1 that only considers non-overlapping processes.

Is segmentation achievable using current datasets?

Though some methods exist to obtain segmentation maps from the image-level information -and were applied on the TenGeoP dataset in Chapter 3- it proved to be difficult to obtain high accuracy from weakly-supervised methods. Chapter 3 was not an exhaustive review of weakly-supervised methods, but it highlighted the importance of high quality groundtruth. This point was confirmed by the rainfall estimation (Chapter 6) and the wind speed estimation (Chapter 7). It convinced us that fully-supervised methods have advantages that counterbalance the difficulty in obtaining segmentation datasets. Furthermore, groundtruth can be obtained from collocation with non-SAR sensors rather than manual annotation. It decreases the time needed to build a dataset, despite a cleaning step required to assure the quality of the automatically retrieved groundtruth. Other frameworks, such as semi-supervised [16], could also help by providing means to learn from both annotated and unannotated data but are left for further studies.

2 Perspectives

SAR images are richer than what is used in our models and additional channels, unused in this manuscript, are available to enhance the quality of the segmentations. The cross polarization channel is less subject to saturation at high wind speed. The thermal noise annotation, available in GRD products, is thought to be beneficial to select which pixels of the observation are the most informative. Topographic information such as the bathymetry/elevation and the distance to coast can be used to understand the context of the acquisition and enhanced to capacity of the models to work in various meteocean contexts. Internal variables such as the incidence angle the heading are also of interest as the aspect of the process change depending on the geometry of the observation. Finally, meteorological information, such as the wind direction or its speed, can provide prior beneficial to the models. However, the integration of additional channels is difficult due to the increase of the input vector space dimension. Especially when working on datasets of limited size, it has been experienced that supplementary channels tend to overfit the model and, as such, did not systematically lead to enhanced capabilities. Further work should study the possibility to use these ancillary channels in addition to the single VV polarization. The use of the additional channels could be achieved on the rainfall estimation problem. Current work is directed at increasing the dataset using the Operational Programme for the Exchange of weather Radar Information (OPERA) that provides pan-European radar composite. Since we now have access to a rainfall estimator, it is easier to retrieve the metocean condition likely to over- or under-estimate the precipitation and compare with weather radar observations. This would allow to extend the capabilities of the rainfall estimator to stronger wind regimes. This is of particular interest since extreme events such as hurricanes are among the phenomena with both high rainfall and wind speed.

The convective process segmenters can be used to perform systematic collocations with secondary sensors. Geostationary satellites, such as the GOES, Himawari, and MSG series, are well-suited for collocation because they provide large observations at high time resolution. These satellites all have passive imaging radiometers on board with IR band observations that provide information on water vapor and indirectly, convective processes. The 11.2 µm band (GOES & Himawari) or the 10.8 µm band (MSG) can be used as priors to estimate convective cells. However, it is unknown if using this supplementary channel will increase overfitting and require new annotations.

The study on biological slicks found a correlation between the probability of their appearance and ocean color, which was expected as they are both byproducts of photosynthesis. However, the correlation is weaker at large scales. Further studies, including collocation of additional ocean properties, such as salinity, sea surface temperature, and surface currents, are needed to better understand the differences between the two observation processes.

With the increasing availability of data and sensors, it is believed that the number of models dedicated to geophysical processes will continue to rise and improve their capabilities. There are still many metocean phenomena that have not been studied, such as orographic waves, upwelling, internal waves, eddies, and extreme events like hurricanes. These subjects are of high interest in the oceanographic community, but acquiring annotated datasets can be difficult. Nevertheless, it is expected that in the near future, the number of SAR-based segmentation maps will increase and cover an even wider range of phenomena.

BIBLIOGRAPHY

[1] W. Alpers, Heng Wang-Chen, and Lim Hock, “Observation of internal waves in the Andaman Sea by ERS SAR”, *in*: *IGARSS'97. 1997 IEEE International Geoscience and Remote Sensing Symposium Proceedings. Remote Sensing - A Scientific Vision for Sustainable Development*, IEEE, DOI: `10.1109/igarss.1997.608926`, URL: `https://doi.org/10.1109/igarss.1997.608926`.

[2] Werner Alpers, “Theory of radar imaging of internal waves”, *in*: *Nature* 314.*6008* (Mar. 1985), pp. 245–247, DOI: `10.1038/314245a0`, URL: `https://doi.org/10.1038/314245a0`.

[3] Werner Alpers et al., “Rain footprints on C-band synthetic aperture radar images of the ocean - Revisited”, *in*: *Remote Sensing of Environment* 187 (Dec. 2016), pp. 169–185, DOI: `10.1016/j.rse.2016.10.015`, URL: `https://doi.org/10.1016/j.rse.2016.10.015`.

[4] Werner R. Alpers, Duncan B. Ross, and Clifford L. Rufenach, “On the detectability of ocean surface waves by real and synthetic aperture radar”, *in*: *Journal of Geophysical Research* 86.*C7* (1981), p. 6481, DOI: `10.1029/jc086ic07p06481`, URL: `https://doi.org/10.1029/jc086ic07p06481`.

[5] Arabi Mohammed El Amin, Qingjie Liu, and Yunhong Wang, “Convolutional neural network features based change detection in satellite images”, *in*: *SPIE Proceedings*, ed. by Xudong Jiang et al., SPIE, July 2016, DOI: `10.1117/12.2243798`, URL: `https://doi.org/10.1117/12.2243798`.

[6] Emna Amri et al., “Automatic offshore oil slick detection based on deep learning using SAR data and contextual information”, *in*: *Remote Sensing of the Ocean, Sea Ice, Coastal Waters, and Large Water Regions 2021*, ed. by Charles R. Bostater and Xavier Neyt, SPIE, Sept. 2021, DOI: `10.1117/12.2598032`, URL: `https://doi.org/10.1117/12.2598032`.

[7] Dongyang Ao et al., “Dialectical GAN for SAR Image Translation: From Sentinel-1 to TerraSAR-X”, *in*: *Remote Sensing* 10.*10* (Oct. 2018), p. 1597, DOI: `10.3390/rs10101597`, URL: `https://doi.org/10.3390/rs10101597`.

[8] Fabrice Ardhuin et al., “Measuring ocean waves in sea ice using SAR imagery: A quasi-deterministic approach evaluated with Sentinel-1 and in situ data”, *in*: *Remote Sensing of Environment* 189 (Feb. 2017), pp. 211–222, DOI: 10.1016/j.rse.2016.11.024, URL: https://doi.org/10.1016/j.rse.2016.11.024.

[9] Evert Attema, Yves-Louis Desnos, and Guy Duchossois, “Synthetic aperture radar in europe: ERS, Envisat, and beyond”, *in*: *Johns Hopkins Apl Technical Digest* 21 (2000), pp. 155–161.

[10] Evert Attema et al., “Sentinel-1 mission overview”, *in*: *2009 IEEE International Geoscience and Remote Sensing Symposium*, IEEE, 2009, DOI: 10.1109/igarss.2009.5416921, URL: https://doi.org/10.1109/igarss.2009.5416921.

[11] Yoshua Bengio, “Practical Recommendations for Gradient-Based Training of Deep Architectures”, *in*: *Lecture Notes in Computer Science*, Springer Berlin Heidelberg, 2012, pp. 437–478, DOI: 10.1007/978-3-642-35289-8_26, URL: https://doi.org/10.1007/978-3-642-35289-8_26.

[12] T. I. Bern et al., “Oil spill detection using satellite based SAR - Experience from a field experiment”, *in*: *Photogrammetric Engineering and Remote Sensing; (United States)* 59:3 (Mar. 1993), ISSN: 0099-1112, URL: https://www.osti.gov/biblio/6311687.

[13] E. Bosc, A. Bricaud, and D. Antoine, “Seasonal and interannual variability in algal biomass and primary production in the Mediterranean Sea, as derived from 4 years of SeaWiFS observations”, *in*: *Global Biogeochemical Cycles* 18.*1* (2004), n/a–n/a, DOI: 10.1029/2003gb002034, URL: https://doi.org/10.1029/2003gb002034.

[14] Hugo Boulze, Anton Korosov, and Julien Brajard, “Classification of Sea Ice Types in Sentinel-1 SAR Data Using Convolutional Neural Networks”, *in*: *Remote Sensing* 12.*13* (July 2020), p. 2165, DOI: 10.3390/rs12132165, URL: https://doi.org/10.3390/rs12132165.

[15] Diego Cantorna et al., “Oil spill segmentation in SAR images using convolutional neural networks. A comparative analysis with clustering and logistic regression algorithms”, *in*: *Applied Soft Computing* 84 (Nov. 2019), p. 105716, DOI: 10.1016/j.asoc.2019.105716, URL: https://doi.org/10.1016/j.asoc.2019.105716.

[16] Javiera Castillo-Navarro et al., "Semi-supervised semantic segmentation in Earth Observation: the MiniFrance suite, dataset analysis and multi-task network study", *in*: *Machine Learning* 111.*9* (Apr. 2021), pp. 3125–3160, DOI: 10.1007/s10994-020-05943-y, URL: https://doi.org/10.1007/s10994-020-05943-y.

[17] Yang-Lang Chang et al., "Ship Detection Based on YOLOv2 for SAR Imagery", *in*: *Remote Sensing* 11.*7* (Apr. 2019), p. 786, DOI: 10.3390/rs11070786, URL: https://doi.org/10.3390/rs11070786.

[18] Chen Chen et al., "A Deep Neural Network Based on an Attention Mechanism for SAR Ship Detection in Multiscale and Complex Scenarios", *in*: *IEEE Access* 7 (2019), pp. 104848–104863, DOI: 10.1109/access.2019.2930939, URL: https://doi.org/10.1109/access.2019.2930939.

[19] Liang-Chieh Chen et al., *Rethinking Atrous Convolution for Semantic Image Segmentation*, 2017, DOI: 10.48550/ARXIV.1706.05587, URL: https://arxiv.org/abs/1706.05587.

[20] Angela Cheng et al., "Accuracy and inter-analyst agreement of visually estimated sea ice concentrations in Canadian Ice Service ice charts using single-polarization RADARSAT-2", *in*: *The Cryosphere* 14.*4* (Apr. 2020), pp. 1289–1310, DOI: 10.5194/tc-14-1289-2020, URL: https://doi.org/10.5194/tc-14-1289-2020.

[21] Junhwa Chi and Hyun choel Kim, "Prediction of Arctic Sea Ice Concentration Using a Fully Data Driven Deep Neural Network", *in*: *Remote Sensing* 9.*12* (Dec. 2017), p. 1305, DOI: 10.3390/rs9121305, URL: https://doi.org/10.3390/rs9121305.

[22] K. H. Christensen and E. Terrile, "Drift and deformation of oil slicks due to surface waves", *in*: *Journal of Fluid Mechanics* 620 (Feb. 2009), pp. 313–332, DOI: 10.1017/s0022112008004606, URL: https://doi.org/10.1017s0022112008004606.

[23] Dan C. Ciresan et al., "Flexible, High Performance Convolutional Neural Networks for Image Classification", *in*: *IJCAI*, 2011, DOI: 10.5591/978-1-57735-516-8/IJCAI11-210.

[24] D.A. Clausi and B. Yue, "Comparing Cooccurrence Probabilities and Markov Random Fields for Texture Analysis of SAR Sea Ice Imagery", *in*: *IEEE Transactions on Geoscience and Remote Sensing* 42.*1* (Jan. 2004), pp. 215–228, DOI: 10.1109/tgrs.2003.817218, URL: https://doi.org/10.1109/tgrs.2003.817218.

[25] David A. Clausi, “Comparison and fusion of co-occurrence, Gabor and MRF texture features for classification of SAR sea-ice imagery”, *in*: *Atmosphere-Ocean* 39.*3* (Sept. 2001), pp. 183–194, DOI: `10.1080/07055900.2001.9649675`, URL: `https://doi.org/10.1080/07055900.2001.9649675`.

[26] Pablo Clemente-Colon et al., “Observations of Hurricane Bonnie in spaceborne synthetic aperture radar (SAR) and next-generation Doppler weather radar (NEXRAD)”, *in*: *SPIE Proceedings*, ed. by Jaqueline E. Russell, SPIE, Dec. 1999, DOI: `10.1117/12.373044`, URL: `https://doi.org/10.1117/12.373044`.

[27] Aurélien Colin et al., “Semantic Segmentation of Metoceanic Processes Using SAR Observations and Deep Learning”, *in*: *Remote Sensing* 14.*4* (Feb. 2022), p. 851, DOI: `10.3390/rs14040851`, URL: `https://doi.org/10.3390/rs14040851`.

[28] Aurélien Colin et al., *Rain Rate Estimation with SAR using NEXRAD measurements with Convolutional Neural Networks*, 2022, DOI: `10.48550/arxiv.2207.07333`, URL: `https://arxiv.org/abs/2207.07333`.

[29] J. C. Comiso, “Characteristics of Arctic winter sea ice from satellite multispectral microwave observations”, *in*: *Journal of Geophysical Research* 91.*C1* (1986), p. 975, DOI: `10.1029/jc091ic01p00975`, URL: `https://doi.org/10.1029/jc091ic01p00975`.

[30] J.C. Comiso, D.J. Cavalieri, and T. Markus, “Sea ice concentration, ice temperature, and snow depth using AMSR-E data”, *in*: *IEEE Transactions on Geoscience and Remote Sensing* 41.*2* (Feb. 2003), pp. 243–252, DOI: `10.1109/tgrs.2002.808317`, URL: `https://doi.org/10.1109/tgrs.2002.808317`.

[31] Colin L. V. Cooke and K. Andrea Scott, “Estimating Sea Ice Concentration From SAR: Training Convolutional Neural Networks With Passive Microwave Data”, *in*: *IEEE Transactions on Geoscience and Remote Sensing* 57.*7* (July 2019), pp. 4735–4747, DOI: `10.1109/tgrs.2019.2892723`, URL: `https://doi.org/10.1109/tgrs.2019.2892723`.

[32] *Copernicus Marine Environment Monitoring Service*, URL: `https://marine.copernicus.eu/` (visited on 10/05/2022).

[33] Zongyong Cui et al., “Image Data Augmentation for SAR Sensor via Generative Adversarial Nets”, *in*: *IEEE Access* 7 (2019), pp. 42255–42268, DOI: `10.1109/access.2019.2907728`, URL: `https://doi.org/10.1109/access.2019.2907728`.

[34] G. Cybenko, "Approximation by superpositions of a sigmoidal function", *in*: *Mathematics of Control, Signals, and Systems* 2.*4* (Dec. 1989), pp. 303–314, DOI: 10.1007/bf02551274, URL: https://doi.org/10.1007/bf02551274.

[35] A. Danklmayer et al., "Assessment of Atmospheric Propagation Effects in SAR Images", *in*: *IEEE Transactions on Geoscience and Remote Sensing* 47.*10* (2009), pp. 3507–3518, DOI: 10.1109/tgrs.2009.2022271, URL: https://doi.org/10.1109/tgrs.2009.2022271.

[36] Clément Dechesne et al., "Ship Identification and Characterization in Sentinel-1 SAR Images with Multi-Task Deep Learning", *in*: *Remote Sensing* 11.*24* (Dec. 2019), p. 2997, DOI: 10.3390/rs11242997, URL: https://doi.org/10.3390/rs11242997.

[37] Huawu Deng and D.A. Clausi, "Unsupervised segmentation of synthetic aperture Radar sea ice imagery using a novel Markov random field model", *in*: *IEEE Transactions on Geoscience and Remote Sensing* 43.*3* (Mar. 2005), pp. 528–538, DOI: 10.1109/tgrs.2004.839589, URL: https://doi.org/10.1109/tgrs.2004.839589.

[38] Wolfgang Dierking and Christine Wesche, "C-Band Radar Polarimetry—Useful for Detection of Icebergs in Sea Ice?", *in*: *IEEE Transactions on Geoscience and Remote Sensing* 52.*1* (Jan. 2014), pp. 25–37, DOI: 10.1109/tgrs.2012.2234756, URL: https://doi.org/10.1109/tgrs.2012.2234756.

[39] Frode Dinessen, Bruce Hackett, and Mathilde Brandt Kreiner, *Product User Manual for SI TAC regional Arctic sea ice product 011002*, tech. rep., 2019.

[40] Michael F. Donovan et al., "The Identification and Verification of Hazardous Convective Cells over Oceans Using Visible and Infrared Satellite Observations", *in*: *Journal of Applied Meteorology and Climatology* 47.*1* (Jan. 2008), pp. 164–184, DOI: 10.1175/2007jamc1471.1, URL: https://doi.org/10.1175/2007jamc1471.1.

[41] H. Douville et al., *in*: *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*, Cambridge University Press, 2021.

[42] Aryeh J. Drager and Susan C. van den Heever, "Characterizing convective cold pools", *in*: *Journal of Advances in Modeling Earth Systems* 9.*2* (May 2017), pp. 1091–1115, DOI: 10.1002/2016ms000788, URL: https://doi.org/10.1002/2016ms000788.

[43] Johannes Drönner et al., "Fast Cloud Segmentation Using Convolutional Neural Networks", *in*: *Remote Sensing* 10.*11* (Nov. 2018), p. 1782, DOI: 10.3390/rs10111782, URL: https://doi.org/10.3390/rs10111782.

[44] Yong Du, Paris W Vachon, and John Wolfe, "Wind direction estimation from SAR images of the ocean using wavelet analysis", *in*: *Canadian Journal of Remote Sensing* 28.*3* (2002), pp. 498–509, DOI: 10.5589/m02-029, URL: https://www.tandfonline.com/doi/abs/10.5589/m02-029.

[45] Vincent Dumoulin and Francesco Visin, *A guide to convolution arithmetic for deep learning*, 2016, DOI: 10.48550/ARXIV.1603.07285, URL: https://arxiv.org/abs/1603.07285.

[46] Eastwood et al., *Quality Information Document For OSI TAC Sea Ice products 011-001, -002, -003, -004, -006, -007, -009, -010, -011, -012*, System documentation report for the Copernicus Marine Environment Monitoring Service, 2017.

[47] Anis Elyouncha et al., "Assessment of the corrected CMOD6 GMF using scatterometer data", *in*: *SPIE Proceedings*, ed. by Charles R. Bostater, Stelios P. Mertikas, and Xavier Neyt, SPIE, Oct. 2015, DOI: 10.1117/12.2195727, URL: https://doi.org/10.1117/12.2195727.

[48] N. V. Evtushenko and A. Yu. Ivanov, "Oil seeps in the southeastern Black Sea studied using satellite synthetic aperture radar images", *in*: *Izvestiya, Atmospheric and Oceanic Physics* 49.*9* (Dec. 2013), pp. 913–918, DOI: 10.1134/s0001433813090065, URL: https://doi.org/10.1134/s0001433813090065.

[49] Zhe Feng et al., "Mechanisms of convective cloud organization by cold pools over tropical warm ocean during the AMIE/DYNAMO field campaign", *in*: *Journal of Advances in Modeling Earth Systems* 7.*2* (Apr. 2015), pp. 357–381, DOI: 10.1002/2014ms000384, URL: https://doi.org/10.1002/2014ms000384.

[50] Giulia Fracastoro et al., "Deep Learning Methods For Synthetic Aperture Radar Image Despeckling: An Overview Of Trends And Perspectives", *in*: *IEEE Geoscience and Remote Sensing Magazine* 9.*2* (June 2021), pp. 29–51, DOI: 10.1109/mgrs.2021.3070956, URL: https://doi.org/10.1109/mgrs.2021.3070956.

[51] Massimiliano Gargiulo et al., "Fast Super-Resolution of 20 m Sentinel-2 Bands Using Convolutional Neural Networks", *in*: *Remote Sensing* 11.*22* (Nov. 2019), p. 2635, DOI: `10.3390/rs11222635`, URL: `https://doi.org/10.3390/rs11222635`.

[52] Robert Geirhos et al., *ImageNet-trained CNNs are biased towards texture; increasing shape bias improves accuracy and robustness*, 2018, DOI: `10.48550/ARXIV.1811.12231`, URL: `https://arxiv.org/abs/1811.12231`.

[53] I. de Gélis, S. Lefèvre, and T. Corpetti, "3D urban change detection with point cloud siamese networks", *in*: *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* XLIII-B3-2021 (June 2021), pp. 879–886, DOI: `10.5194/isprs-archives-xliii-b3-2021-879-2021`, URL: `https://doi.org/10.5194/isprs-archives-xliii-b3-2021-879-2021`.

[54] Iris de Gelis, Aurelien Colin, and Nicolas Longepe, "Prediction of Categorized Sea Ice Concentration From Sentinel-1 SAR Images Based on a Fully Convolutional Network", *in*: *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 14 (2021), pp. 5831–5841, DOI: `10.1109/jstars.2021.3074068`, URL: `https://doi.org/10.1109/jstars.2021.3074068`.

[55] Olivia Gérigny et al., "Small-scale variability of the current in the Strait of Bonifacio", *in*: *Ocean Dynamics* 65.*8* (July 2015), pp. 1165–1182, DOI: `10.1007/s10236-015-0863-5`, URL: `https://doi.org/10.1007/s10236-015-0863-5`.

[56] Jagvijay P.S. Gill and John J. Yackel, "Evaluation of C-band SAR polarimetric parameters for discrimination of first-year sea ice types", *in*: *Canadian Journal of Remote Sensing* 38.*3* (2012), pp. 306–323, DOI: `10.5589/m12-025`, eprint: `https://doi.org/10.5589/m12-025`, URL: `https://doi.org/10.5589/m12-025`.

[57] Girard-Ardhuin, Mercier, and Garello, "Oil slick detection by SAR imagery: potential and limitation", *in*: *Oceans 2003. Celebrating the Past ... Teaming Toward the Future (IEEE Cat. No.03CH37492)*, IEEE, 2003, DOI: `10.1109/oceans.2003.178539`, URL: `https://doi.org/10.1109/oceans.2003.178539`.

[58] Cristina Gómez, Joanne C. White, and Michael A. Wulder, "Optical remotely sensed time series data for land cover classification: A review", *in*: *ISPRS Journal of Photogrammetry and Remote Sensing* 116 (June 2016), pp. 55–72, DOI: `10.1016/j.isprsjprs.2016.03.008`, URL: `https://doi.org/10.1016/j.isprsjprs.2016.03.008`.

[59] Ian Goodfellow, Yoshua Bengio, and Aaron Courville, *Deep Learning*, `http://www.deeplearningbook.org`, MIT Press, 2016.

[60] Ian Goodfellow et al., “Generative adversarial nets”, *in*: *Advances in neural information processing systems*, 2014, pp. 2672–2680.

[61] Steven J. Goodman et al., “The GOES-R Geostationary Lightning Mapper (GLM)”, *in*: *Atmospheric Research* 125-126 (May 2013), pp. 34–49, DOI: `10.1016/j.atmosres.2013.01.006`, URL: `https://doi.org/10.1016/j.atmosres.2013.01.006`.

[62] D.K. Hall et al., “Sea ice surface temperature product from MODIS”, *in*: *IEEE Transactions on Geoscience and Remote Sensing* 42.*5* (May 2004), pp. 1076–1087, DOI: `10.1109/tgrs.2004.825587`, URL: `https://doi.org/10.1109/tgrs.2004.825587`.

[63] Zayd Mahmoud Hamdi, Melanie Brandmeier, and Christoph Straub, “Forest Damage Assessment Using Deep Learning on High Resolution Remote Sensing Data”, *in*: *Remote Sensing* 11.*17* (Aug. 2019), p. 1976, DOI: `10.3390/rs11171976`, URL: `https://doi.org/10.3390/rs11171976`.

[64] Junwei Han et al., “Background Prior-Based Salient Object Detection via Deep Reconstruction Residual”, *in*: *IEEE Transactions on Circuits and Systems for Video Technology* 25.*8* (Aug. 2015), pp. 1309–1321, DOI: `10.1109/tcsvt.2014.2381471`, URL: `https://doi.org/10.1109/tcsvt.2014.2381471`.

[65] Kaiming He et al., *Deep Residual Learning for Image Recognition*, 2015, DOI: `10.48550/ARXIV.1512.03385`, URL: `https://arxiv.org/abs/1512.03385`.

[66] Kaiming He et al., “Deep Residual Learning for Image Recognition”, *in*: *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE, June 2016, DOI: `10.1109/cvpr.2016.90`, URL: `https://doi.org/10.1109/cvpr.2016.90`.

[67] Xuming He, R.S. Zemel, and M.A. Carreira-Perpinan, “Multiscale conditional random fields for image labeling”, *in*: *Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2004. CVPR 2004.* IEEE, DOI: `10.1109/cvpr.2004.1315232`, URL: `https://doi.org/10.1109/cvpr.2004.1315232`.

[68] Corentin Henry, Seyed Majid Azimi, and Nina Merkle, “Road Segmentation in SAR Satellite Images With Deep Fully Convolutional Neural Networks”, *in*: *IEEE Geoscience and Remote Sensing Letters* 15.*12* (Dec. 2018), pp. 1867–1871, DOI: `10.1109/lgrs.2018.2864342`, URL: `https://doi.org/10.1109/lgrs.2018.2864342`.

[69] Hans Hersbach, “CMOD5 An improved geophysical model function for ERS C-band scatterometry”, *in*: (2003), DOI: `10.21957/6GHPJQGW3`, URL: `https://www.ecmwf.int/node/9861`.

[70] Hans Hersbach, “CMOD5.N: A C-band geophysical model function for equivalent neutral wind.”, *in*: (2008), DOI: `10.21957/MZCFM6JFL`, URL: `https://www.ecmwf.int/node/9873`.

[71] Hans Hersbach, “Comparison of C-Band Scatterometer CMOD5.N Equivalent Neutral Winds with ECMWF”, *in*: *Journal of Atmospheric and Oceanic Technology* 27.*4* (Apr. 2010), pp. 721–736, DOI: `10.1175/2009jtecho698.1`, URL: `https://doi.org/10.1175/2009jtecho698.1`.

[72] Joel Hestness et al., *Deep Learning Scaling is Predictable, Empirically*, 2017, DOI: `10.48550/ARXIV.1712.00409`, URL: `https://arxiv.org/abs/1712.00409`.

[73] Yaoshiang Ho and Samuel Wookey, “The Real-World-Weight Cross-Entropy Loss Function: Modeling the Costs of Mislabeling”, *in*: *IEEE Access* 8 (2020), pp. 4806–4813, DOI: `10.1109/access.2019.2962617`, URL: `https://doi.org/10.1109/access.2019.2962617`.

[74] Dan-Bee Hong and Chan-Su Yang, “Automatic discrimination approach of sea ice in the Arctic Ocean using Sentinel-1 Extra Wide Swath dual-polarized SAR data”, *in*: *International Journal of Remote Sensing* 39.*13* (Feb. 2018), pp. 4469–4483, DOI: `10.1080/01431161.2017.1415486`, URL: `https://doi.org/10.1080/01431161.2017.1415486`.

[75] Benyamin Hosseiny et al., “WetNet: A Spatial–Temporal Ensemble Deep Learning Model for Wetland Classification Using Sentinel-1 and Sentinel-2”, *in*: *IEEE Transactions on Geoscience and Remote Sensing* 60 (2022), pp. 1–14, DOI: `10.1109/tgrs.2021.3113856`, URL: `https://doi.org/10.1109%2Ftgrs.2021.3113856`.

[76] C. Howell et al., “Iceberg and ship discrimination with ENVISAT multi-polarization ASAR”, *in*: *IEEE International IEEE International IEEE International Geoscience and Remote Sensing Symposium, 2004. IGARSS '04. Proceedings. 2004*, IEEE, DOI: 10.1109/igarss.2004.1368958, URL: https://doi.org/10.1109/igarss.2004.1368958.

[77] K.A. Hunter and P.S. Liss, “Chapter 9 Organic Sea Surface Films”, *in*: *Elsevier Oceanography Series*, Elsevier, 1981, pp. 259–298, DOI: 10.1016/s0422-9894(08)70331-3, URL: https://doi.org/10.1016/s0422-9894(08)70331-3.

[78] Louis J. Ippolito, “Attenuation by Atmospheric Gases”, *in*: *Radiowave Propagation in Satellite Communications*, Springer Netherlands, 1986, pp. 25–37, DOI: 10.1007/978-94-011-7027-7_3, URL: https://doi.org/10.1007/978-94-011-7027-7_3.

[79] Kostiantyn Isaienkov et al., “Deep Learning for Regular Change Detection in Ukrainian Forest Ecosystem With Sentinel-2”, *in*: *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 14 (2021), pp. 364–376, DOI: 10.1109/jstars.2020.3034186, URL: https://doi.org/10.1109/jstars.2020.3034186.

[80] Christopher Jackson, *Synthetic aperture radar : marine user’s manual*, Washington, D.C: U.S. Dept. of Commerce, National Oceanic et al., 2004, ISBN: 9780160732140.

[81] Christopher R Jackson and John R Apel, eds., *Synthetic Aperture Radar Marine User’s Manual*, U.S. Government Printing Office, Apr. 2005, ISBN: 0-16-073214-X.

[82] JCOMM Expert Team on Sea Ice (ETSI), *Sea-Ice Information Services in the World. Edition 2017*, 2017, DOI: 10.25607/OBP-1325, URL: https://repository.oceanbestpractices.org/handle/11329/394.

[83] A. T. Jones et al., “Coral spawn and bathymetric slicks in Synthetic Aperture Radar (SAR) data from the Timor Sea, north-west Australia”, *in*: *International Journal of Remote Sensing* 27.*10* (May 2006), pp. 2063–2069, DOI: 10.1080/01431160500445308, URL: https://doi.org/10.1080/01431160500445308.

[84] Ivan Kapustin et al., “On Capabilities of Tracking Marine Surface Currents Using Artificial Film Slicks”, *in*: *Remote Sensing* 11.*7* (Apr. 2019), p. 840, DOI: 10.3390/rs11070840, URL: https://doi.org/10.3390/rs11070840.

[85] M.A. Karam et al., “A microwave scattering model for layered vegetation”, *in*: *IEEE Transactions on Geoscience and Remote Sensing* 30.*4* (July 1992), pp. 767–784, DOI: 10.1109/36.158872, URL: https://doi.org/10.1109/36.158872.

[86] Krishna Karra et al., “Global land use / land cover with Sentinel 2 and deep learning”, *in*: *2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS*, IEEE, July 2021, DOI: 10.1109/igarss47720.2021.9553499, URL: https://doi.org/10.1109/igarss47720.2021.9553499.

[87] J. Karvonen, “A sea ice concentration estimation algorithm utilizing radiometer and SAR data”, *in*: *The Cryosphere* 8.*5* (Sept. 2014), pp. 1639–1650, DOI: 10.5194/tc-8-1639-2014, URL: https://doi.org/10.5194/tc-8-1639-2014.

[88] Juha Karvonen, “Baltic Sea Ice Concentration Estimation Based on C-Band Dual-Polarized SAR Data”, *in*: *IEEE Transactions on Geoscience and Remote Sensing* 52.*9* (Sept. 2014), pp. 5558–5566, DOI: 10.1109/tgrs.2013.2290331, URL: https://doi.org/10.1109/tgrs.2013.2290331.

[89] Juha Karvonen et al., “A Comparison Between High-Resolution EO-Based and Ice Analyst-Assigned Sea Ice Concentrations”, *in*: *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 8.*4* (Apr. 2015), pp. 1799–1807, DOI: 10.1109/jstars.2015.2426414, URL: https://doi.org/10.1109/jstars.2015.2426414.

[90] Mohamed Kerkech, Adel Hafiane, and Raphael Canals, “Vine disease detection in UAV multispectral images using optimized image registration and deep learning segmentation approach”, *in*: *Computers and Electronics in Agriculture* 174 (July 2020), p. 105446, DOI: 10.1016/j.compag.2020.105446, URL: https://doi.org/10.1016/j.compag.2020.105446.

[91] Diederik P. Kingma and Jimmy Ba, “Adam: A Method for Stochastic Optimization”, *in*: *ICLR (Poster)*, 2015, URL: http://arxiv.org/abs/1412.6980.

[92] M. J. Kingsford and J. H. Choat, “Influence of surface slicks on the distribution and onshore movements of small fish”, *in*: *Marine Biology* 91.*2* (1986), pp. 161–171, DOI: 10.1007/bf00569432, URL: https://doi.org/10.1007/bf00569432.

[93] W. Koch, “Directional analysis of SAR images aiming at wind direction”, *in*: 42.*4* (Apr. 2004), pp. 702–710, DOI: 10.1109/tgrs.2003.818811, URL: https://doi.org/10.1109/tgrs.2003.818811.

[94] Igor Kozlov et al., "SAR observing large-scale nonlinear internal waves in the White Sea", *in*: *Remote Sensing of Environment* 147 (May 2014), pp. 99–107, DOI: 10.1016/j.rse.2014.02.017, URL: https://doi.org/10.1016/j.rse.2014.02.017.

[95] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton, "ImageNet Classification with Deep Convolutional Neural Networks", *in*: *Advances in Neural Information Processing Systems*, ed. by F. Pereira et al., vol. 25, Curran Associates, Inc., 2012, URL: https://proceedings.neurips.cc/paper/2012/file/c399862d3b9d6b76c8436e924a68c45b-Paper.pdf.

[96] Ryan Kruk et al., "Proof of Concept for Sea Ice Stage of Development Classification Using Deep Learning", *in*: *Remote Sensing* 12.*15* (Aug. 2020), p. 2486, DOI: 10.3390/rs12152486, URL: https://doi.org/10.3390/rs12152486.

[97] Irene L. Kruse, Jan O. Haerter, and Bettina Meyer, "Cold pools over the Netherlands: A statistical study from tower and radar observations", *in*: *Quarterly Journal of the Royal Meteorological Society* 148.*743* (Dec. 2021), pp. 711–726, DOI: 10.1002/qj.4223, URL: https://doi.org/10.1002/qj.4223.

[98] Jan-Hendrik Körber et al., "Natural oil seepage at Kobuleti Ridge, eastern Black Sea", *in*: *Marine and Petroleum Geology* 50 (Feb. 2014), pp. 68–82, DOI: 10.1016/j.marpetgeo.2013.11.007, URL: https://doi.org/10.1016/j.marpetgeo.2013.11.007.

[99] Tran Vu La et al., "Detection of convective systems through surface wind gust estimation based on Sentinel-1 images: A new approach", *in*: *Atmospheric Science Letters* 19.*12* (Oct. 2018), e863, DOI: 10.1002/asl.863, URL: https://doi.org/10.1002/asl.863.

[100] Tran Vu La et al., "Use of Sentinel-1 C-Band SAR Images for Convective System Surface Wind Pattern Detection", *in*: *Journal of Applied Meteorology and Climatology* 59.*8* (Aug. 2020), pp. 1321–1332, DOI: 10.1175/jamc-d-20-0008.1, URL: https://doi.org/10.1175/jamc-d-20-0008.1.

[101] K. Lagouvardos et al., "Study of a heavy precipitation event over southern France, in the frame of HYMEX project: Observational analysis and model results using assimilation of lightning", *in*: *Atmospheric Research* 134 (Dec. 2013), pp. 45–55, DOI: 10.1016/j.atmosres.2013.07.003, URL: https://doi.org/10.1016/j.atmosres.2013.07.003.

[102] Charis Lanaras et al., “Super-resolution of Sentinel-2 images: Learning a globally applicable deep neural network”, *in*: *ISPRS Journal of Photogrammetry and Remote Sensing* 146 (Dec. 2018), pp. 305–319, DOI: 10.1016/j.isprsjprs.2018.09.018, URL: https://doi.org/10.1016/j.isprsjprs.2018.09.018.

[103] Francesco Lattari et al., “Deep Learning for SAR Image Despeckling”, *in*: *Remote Sensing* 11.*13* (June 2019), p. 1532, DOI: 10.3390/rs11131532, URL: https://doi.org/10.3390/rs11131532.

[104] O. Yu. Lavrova and M. I. Mityagina, “Satellite monitoring of oil slicks on the Black Sea surface”, *in*: *Izvestiya, Atmospheric and Oceanic Physics* 49.*9* (Dec. 2013), pp. 897–912, DOI: 10.1134/s0001433813090107, URL: https://doi.org/10.1134/s0001433813090107.

[105] Yann LeCun, Yoshua Bengio, and Geoffrey Hinton, “Deep learning”, *in*: *Nature* 521.*7553* (May 2015), pp. 436–444, DOI: 10.1038/nature14539, URL: https://doi.org/10.1038/nature14539.

[106] Yoonjin Lee, Christian D. Kummerow, and Milija Zupanski, “A simplified method for the detection of convection using high-resolution imagery from GOES-16”, *in*: *Atmospheric Measurement Techniques* 14.*5* (May 2021), pp. 3755–3771, DOI: 10.5194/amt-14-3755-2021, URL: https://doi.org/10.51944/amt-14-3755-2021.

[107] Steven Leigh, Zhijie Wang, and David A. Clausi, “Automated Ice–Water Classification Using Dual Polarization SAR Satellite Imagery”, *in*: *IEEE Transactions on Geoscience and Remote Sensing* 52.*9* (Sept. 2014), pp. 5529–5539, DOI: 10.1109/tgrs.2013.2290231, URL: https://doi.org/10.1109/tgrs.2013.2290231.

[108] Huimin Li et al., “Calibration of the Normalized Radar Cross Section for Sentinel-1 Wave Mode”, *in*: *IEEE Transactions on Geoscience and Remote Sensing* 57.*3* (Mar. 2019), pp. 1514–1522, DOI: 10.1109/tgrs.2018.2867035, URL: https://doi.org/10.1109/tgrs.2018.2867035.

[109] Jianwei Li, Changwen Qu, and Jiaqi Shao, “Ship detection in SAR images based on an improved faster R-CNN”, *in*: *2017 SAR in Big Data Era: Models, Methods and Applications (BIGSARDATA)*, IEEE, Nov. 2017, DOI: 10.1109/bigsardata.2017.8124934, URL: https://doi.org/10.1109/bigsardata.2017.8124934.

[110] Lianfa Li, “Deep Residual Autoencoder with Multiscaling for Semantic Segmentation of Land-Use Images”, *in*: *Remote Sensing* 11.*18* (2019), p. 2142, DOI: `10.3390/rs11182142`, URL: `https://doi.org/10.3390/rs11182142`.

[111] Xiaofeng Li et al., “Deep-learning-based information mining from ocean remote-sensing imagery”, *in*: *National Science Review* 7.*10* (Mar. 2020), pp. 1584–1605, DOI: `10.1093/nsr/nwaa047`, URL: `https://doi.org/10.1093/nsr/nwaa047`.

[112] Haoning Lin, Zhenwei Shi, and Zhengxia Zou, “Maritime Semantic Labeling of Optical Remote Sensing Images with Multi-Scale Fully Convolutional Network”, *in*: *Remote Sensing* 9.*5* (2017), p. 480, DOI: `10.3390/rs9050480`, URL: `https://doi.org/10.3390/rs9050480`.

[113] I.-I. Lin et al., “An ERS-1 synthetic aperture radar image of a tropical squall line compared with weather radar data”, *in*: *IEEE Transactions on Geoscience and Remote Sensing* 39.*5* (May 2001), pp. 937–945, DOI: `10.1109/36.921411`, URL: `https://doi.org/10.1109/36.921411`.

[114] I.-I. Lin et al., “Evidence and quantification of the correlation between radar backscatter and ocean colour supported by simultaneously acquired in situ sea truth”, *in*: *Geophysical Research Letters* 29.*10* (2002), pp. 102–1–102–4, DOI: `10.1029/2001gl014039`, URL: `https://doi.org/10.1029/2001gl014039`.

[115] Tsung-Yi Lin et al., “Microsoft COCO: Common Objects in Context”, *in*: *Computer Vision – ECCV 2014*, Springer International Publishing, 2014, pp. 740–755, DOI: `10.1007/978-3-319-10602-1_48`, URL: `https://doi.org/10.1007/978-3-319-10602-1_48`.

[116] Guosheng Liu, “Satellite remote sensing | Precipitation”, *in*: *Encyclopedia of Atmospheric Sciences*, Elsevier, 2003, pp. 1972–1979, DOI: `10.1016/b0-12-227090-8/00352-3`, URL: `https://doi.org/10.1016/b0-12-227090-8/00352-3`.

[117] Liyuan Liu et al., *On the Variance of the Adaptive Learning Rate and Beyond*, 2019, DOI: `10.48550/ARXIV.1908.03265`, URL: `https://arxiv.org/abs/1908.03265`.

[118] Nian Liu and Junwei Han, “DHSNet: Deep Hierarchical Saliency Network for Salient Object Detection”, *in*: *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE, June 2016, DOI: `10.1109/cvpr.2016.80`, URL: `https://doi.org/10.1109/cvpr.2016.80`.

[119] A.E. Long, "C-band V-polarized radar sea-echo model from ERS-1 Haltenbanken campaign", *in*: *Journal of Electromagnetic Waves and Applications* 9.*3* (Jan. 1995), pp. 373–391, DOI: 10.1163/156939395x00532, URL: https://doi.org/10.1163/156939395x00532.

[120] Nicolas Longepe et al., "Comparative Evaluation of Sea Ice Lead Detection Based on SAR Imagery and Altimeter Data", *in*: *IEEE Transactions on Geoscience and Remote Sensing* 57.*6* (June 2019), pp. 4050–4061, DOI: 10.1109/tgrs.2018.2889519, URL: https://doi.org/10.1109/tgrs.2018.2889519.

[121] Nicolas Longépé et al., "Completing fishing monitoring with spaceborne Vessel Detection System (VDS) and Automatic Identification System (AIS) to assess illegal fishing in Indonesia", *in*: *Marine Pollution Bulletin* 131 (June 2018), pp. 33–39, DOI: 10.1016/j.marpolbul.2017.10.016, URL: https://doi.org/10.1016/j.marpolbul.2017.10.016.

[122] J. Louet and S. Bruzzi, "ENVISAT mission and system", *in*: *IEEE 1999 International Geoscience and Remote Sensing Symposium. IGARSS'99 (Cat. No.99CH36293)*, IEEE, DOI: 10.1109/igarss.1999.772059, URL: https://doi.org/10.1109/igarss.1999.772059.

[123] Yiru Lu et al., "A C-band Geophysical Model Function for Determining Coastal Wind Speed Using Synthetic Aperture Radar", *in*: *2018 Progress in Electromagnetics Research Symposium (PIERS-Toyama)*, IEEE, Aug. 2018, DOI: 10.23919/piers.2018.8598163, URL: https://doi.org/10.23919/piers.2018.8598163.

[124] Zhou Lu et al., "The Expressive Power of Neural Networks: A View from the Width", *in*: *Proceedings of the 31st International Conference on Neural Information Processing Systems*, NIPS'17, Long Beach, California, USA: Curran Associates Inc., 2017, 6232–6240, ISBN: 9781510860964.

[125] Wenjie Luo et al., *Understanding the Effective Receptive Field in Deep Convolutional Neural Networks*, 2017, arXiv: 1701.04128 [cs.CV].

[126] Luigi Tommaso Luppino et al., "Deep Image Translation With an Affinity-Based Change Prior for Unsupervised Multimodal Change Detection", *in*: *IEEE Transactions on Geoscience and Remote Sensing* 60 (2022), pp. 1–22, DOI: 10.1109/tgrs.2021.3056196, URL: https://doi.org/10.1109/tgrs.2021.3056196.

[127] Marko Makynen and Juha Karvonen, “Incidence Angle Dependence of First-Year Sea Ice Backscattering Coefficient in Sentinel-1 SAR Imagery Over the Kara Sea”, *in*: *IEEE Transactions on Geoscience and Remote Sensing* 55.*11* (Nov. 2017), pp. 6170–6181, DOI: 10.1109/tgrs.2017.2721981, URL: https://doi.org/10.1109/tgrs.2017.2721981.

[128] David Malmgren-Hansen et al., “A Convolutional Neural Network Architecture for Sentinel-1 and AMSR2 Data Fusion”, *in*: *IEEE Transactions on Geoscience and Remote Sensing* 59.*3* (Mar. 2021), pp. 1890–1902, DOI: 10.1109/tgrs.2020.3004539, URL: https://doi.org/10.1109/tgrs.2020.3004539.

[129] J. S. Marshall, R. C. Langille, and W. Mc K. Palmer, “Measurement of rainfall by radar”, *in*: *Journal of Meteorology* 4.*6* (Dec. 1947), pp. 186–192, DOI: 10.1175/1520-0469(1947)004<0186:morbr>2.0.co;2, URL: https://doi.org/10.1175/1520-0469(1947)004<0186:morbr>2.0.co;2.

[130] Christian Melsheimer, Werner Alpers, and Martin Gade, “Simultaneous observations of rain cells over the ocean by the synthetic aperture radar aboard the ERS satellites and by surface-based weather radars”, *in*: *Journal of Geophysical Research: Oceans* 106.*C3* (Mar. 2001), pp. 4665–4677, DOI: 10.1029/2000jc000263, URL: https://doi.org/10.1029/2000jc000263.

[131] Christian Melsheimer and Gunnar Spreen, *AMSR2 ASI sea ice concentration data, Antarctic, version 5.4 (NetCDF) (July 2012 - December 2019)*, en, 2019, DOI: 10.1594/PANGAEA.898400, URL: https://doi.pangaea.de/10.1594/PANGAEA.898400.

[132] C. Melshelmer, W. Alpers, and M. Gade, “Investigation of multifrequency / multipolarization radar signatures of rain cells, derived from SIR-C/X-SAR data”, *in*: *IGARSS '96. 1996 International Geoscience and Remote Sensing Symposium*, IEEE, DOI: 10.1109/igarss.1996.516666, URL: https://doi.org/10.1109/igarss.1996.516666.

[133] Andrea Meraner et al., “Cloud removal in Sentinel-2 imagery using a deep residual neural network and SAR-optical data fusion”, *in*: *ISPRS Journal of Photogrammetry and Remote Sensing* 166 (Aug. 2020), pp. 333–346, DOI: 10.1016/j.isprsjprs.2020.05.013, URL: https://doi.org/10.1016/j.isprsjprs.2020.05.013.

[134] Takemasa Miyoshi et al., ““Big Data Assimilation” Revolutionizing Severe Weather Prediction”, *in*: *Bulletin of the American Meteorological Society* 97.*8* (Aug. 2016), pp. 1347–1354, DOI: 10.1175/bams-d-15-00144.1, URL: https://doi.org/10.1175/bams-d-15-00144.1.

[135] Domen Mongus and Borut Žalik, “Segmentation schema for enhancing land cover identification: A case study using Sentinel 2 data”, *in*: *International Journal of Applied Earth Observation and Geoinformation* 66 (Apr. 2018), pp. 56–68, DOI: 10.1016/j.jag.2017.11.004, URL: https://doi.org/10.1016/j.jag.2017.11.004.

[136] Louis de Montera et al., “High-resolution offshore wind resource assessment at turbine hub height with Sentinel-1 synthetic aperture radar (SAR) data and machine learning”, *in*: *Wind Energy Science* 7.*4* (July 2022), pp. 1441–1453, DOI: 10.5194/wes-7-1441-2022, URL: https://doi.org/10.5194/wes-7-1441-2022.

[137] Taylor Mordan et al., “Revisiting Multi-Task Learning with ROCK: a Deep Residual Auxiliary Block for Visual Detection”, *in*: *Advances in Neural Information Processing Systems*, ed. by S. Bengio et al., vol. 31, Curran Associates, Inc., 2018, URL: https://proceedings.neurips.cc/paper/2018/file/7f5d04d189dfb634e6a85bb9d9adf2
Paper.pdf.

[138] L C Morena, K V James, and J. Beck, “An introduction to the RADARSAT-2 mission”, *in*: *Canadian Journal of Remote Sensing* 30.*3* (2004), pp. 221–234, DOI: 10.5589/m04-004, URL: https://doi.org/10.5589/m04-004.

[139] Grayson R. Morgan et al., “Deep Learning of High-Resolution Aerial Imagery for Coastal Marsh Change Detection: A Comparative Study”, *in*: *ISPRS International Journal of Geo-Information* 11.*2* (Feb. 2022), p. 100, DOI: 10.3390/ijgi11020100, URL: https://doi.org/10.3390/ijgi11020100.

[140] A. Mouche, P. Vincent, and G. Hajduch, *Sentinel-1 Ocean Wind Fields (OWI) Algorithm Definition*, tech. rep., Nov. 2017, URL: https://sentinels.copernicus.eu/documents/247904/2142675/Thermal-Denoising-of-Products-Generated-by-Sentinel-1-IPF.pdf.

[141] Adugna G. Mullissa, Claudio Persello, and Valentyn Tolpekin, “Fully Convolutional Networks for Multi-Temporal SAR Image Classification”, *in*: *IGARSS 2018 - 2018 IEEE International Geoscience and Remote Sensing Symposium*, IEEE, July 2018, DOI: `10.1109/igarss.2018.8518780`, URL: `https://doi.org/10.1109/igarss.2018.8518780`.

[142] Edoardo Nemni et al., “Fully Convolutional Neural Network for Rapid Flood Segmentation in Synthetic Aperture Radar Imagery”, *in*: *Remote Sensing* 12.*16* (Aug. 2020), p. 2532, DOI: `10.3390/rs12162532`, URL: `https://doi.org/10.3390/rs12162532`.

[143] Shuhratchon Ochilov and David A. Clausi, “Operational SAR Sea-Ice Image Classification”, *in*: *IEEE Transactions on Geoscience and Remote Sensing* 50.*11* (Nov. 2012), pp. 4397–4408, DOI: `10.1109/tgrs.2012.2192278`, URL: `https://doi.org/10.1109/tgrs.2012.2192278`.

[144] Florin Onea and Eugen Rusu, “Wind energy assessments along the Black Sea basin”, *in*: *Meteorological Applications* 21.*2* (2012), pp. 316–329, DOI: `10.1002/met.1337`, URL: `https://doi.org/10.1002/met.1337`.

[145] Lucas Prado Osco et al., “Semantic segmentation of citrus-orchard using deep neural networks and multispectral UAV-based imagery”, *in*: *Precision Agriculture* (Jan. 2021), DOI: `10.1007/s11119-020-09777-5`, URL: `https://doi.org/10.1007/s11119-020-09777-5`.

[146] Manohara M.M. Pai et al., “Automatic Segmentation of River and Land in SAR Images: A Deep Learning Approach”, *in*: *2019 IEEE Second International Conference on Artificial Intelligence and Knowledge Engineering (AIKE)*, IEEE, June 2019, DOI: `10.1109/aike.2019.00011`, URL: `https://doi.org/10.1109/aike.2019.00011`.

[147] Jeong-Won Park et al., “Classification of sea ice types in Sentinel-1 synthetic aperture radar images”, *in*: *The Cryosphere* 14.*8* (Aug. 2020), pp. 2629–2645, DOI: `10.5194/tc-14-2629-2020`, URL: `https://doi.org/10.5194/tc-14-2629-2020`.

[148] Jeong-Won Park et al., “Efficient Thermal Noise Removal for Sentinel-1 TOPSAR Cross-Polarization Channel”, *in*: *IEEE Transactions on Geoscience and Remote Sensing* 56.*3* (Mar. 2018), pp. 1555–1565, DOI: `10.1109/tgrs.2017.2765248`, URL: `https://doi.org/10.1109/tgrs.2017.2765248`.

[149] Kim C Partington et al., "Dual-Polarization C-Band Radar Observations of Sea Ice in the Amundsen Gulf", *in*: *IEEE Transactions on Geoscience and Remote Sensing* 48.*6* (June 2010), pp. 2685–2691, DOI: 10.1109/tgrs.2009.2039577, URL: https://doi.org/10.1109/tgrs.2009.2039577.

[150] Razvan Pascanu, Tomas Mikolov, and Yoshua Bengio, *On the difficulty of training Recurrent Neural Networks*, 2012, DOI: 10.48550/arxiv.1211.5063, URL: https://arxiv.org/abs/1211.5063.

[151] Laryssa Patten, Greta Burger, and Patricia Voumard, "RADARSAT-1 mission planning: Meeting customer needs over 5 years of evolving operations", *in*: *Canadian Journal of Remote Sensing* 28.*2* (2002), pp. 109–113, DOI: 10.5589/m02-016, URL: https://doi.org/10.5589/m02-016.

[152] Daifeng Peng, Yongjun Zhang, and Haiyan Guan, "End-to-End Change Detection for High Resolution Satellite Images Using Improved UNet", *in*: *Remote Sensing* 11.*11* (June 2019), p. 1382, DOI: 10.3390/rs11111382, URL: https://doi.org/10.3390/rs11111382.

[153] Ge Peng and Walter N. Meier, "Temporal and regional variability of Arctic sea-ice coverage from satellite data", *in*: *Annals of Glaciology* 59.*76pt2* (Nov. 2017), pp. 191–200, DOI: 10.1017/aog.2017.32, URL: https://doi.org/10.1017/aog.2017.32.

[154] F. Perazzi et al., "Saliency filters: Contrast based filtering for salient region detection", *in*: *2012 IEEE Conference on Computer Vision and Pattern Recognition*, IEEE, June 2012, DOI: 10.1109/cvpr.2012.6247743, URL: https://doi.org/10.1109/cvpr.2012.6247743.

[155] Ernest W. Peterson and Joseph P. Hennessey, "On the Use of Power Laws for Estimates of Wind Power Potential", *in*: *Journal of Applied Meteorology* 17.*3* (Mar. 1978), pp. 390–394, DOI: 10.1175/1520-0450(1978)017<0390:otuopl>2.0.co;2, URL: https://doi.org/10.1175/1520-0450(1978)017<0390:otuopl>2.0.co;2.

[156] Naef A. A. Qasem et al., "A Comprehensive Review of Saline Water Correlations and Data: Part II—Thermophysical Properties", *in*: *Arabian Journal for Science and Engineering* 46.*3* (Jan. 2021), pp. 1941–1979, DOI: 10.1007/s13369-020-05020-5, URL: https://doi.org/10.1007/s13369-020-05020-5.

[157] Anja Reitz et al., “Sources of fluids and gases expelled at cold seeps offshore Georgia, eastern Black Sea”, *in*: *Geochimica et Cosmochimica Acta* 75.*11* (June 2011), pp. 3250–3268, DOI: 10.1016/j.gca.2011.03.018, URL: https://doi.org/10.1016/j.gca.2011.03.018.

[158] Japan Association on Remote Sensing, *Remote Sensing Note*, Japan Association on Remote Sensing, 1993, URL: http://sar.kangwon.ac.kr/etc/rs_note/rsnote/contents.htm.

[159] Mario Fuentes Reyes et al., “SAR-to-Optical Image Translation Based on Conditional Generative Adversarial Networks—Optimization, Opportunities and Limits”, *in*: *Remote Sensing* 11.*17* (Sept. 2019), p. 2067, DOI: 10.3390/rs11172067, URL: https://doi.org/10.3390/rs11172067.

[160] Maria Belmonte Rivas et al., “A scatterometer record of sea ice extents and backscatter: 1992–2016”, *in*: *The Cryosphere* 12.*9* (Sept. 2018), pp. 2941–2953, DOI: 10.5194/tc-12-2941-2018, URL: https://doi.org/10.5194/tc-12-2941-2018.

[161] David Rolnick et al., *Deep Learning is Robust to Massive Label Noise*, 2017, DOI: 10.48550/ARXIV.1705.10694, URL: https://arxiv.org/abs/1705.10694.

[162] Luis Salgueiro Romero, Javier Marcello, and Verónica Vilaplana, “Super-Resolution of Sentinel-2 Imagery Using Generative Adversarial Networks”, *in*: *Remote Sensing* 12.*15* (July 2020), p. 2424, DOI: 10.3390/rs12152424, URL: https://doi.org/10.3390/rs12152424.

[163] Olaf Ronneberger, Philipp Fischer, and Thomas Brox, “U-Net: Convolutional Networks for Biomedical Image Segmentation”, *in*: *Lecture Notes in Computer Science*, Springer International Publishing, 2015, pp. 234–241, DOI: 10.1007/978-3-319-24574-4_28, URL: https://doi.org/10.1007/978-3-319-24574-4_28.

[164] Olga Russakovsky et al., “ImageNet Large Scale Visual Recognition Challenge”, *in*: *International Journal of Computer Vision (IJCV)* 115.*3* (2015), pp. 211–252, DOI: 10.1007/s11263-015-0816-y.

[165] Giuseppe Scarpa et al., “A CNN-Based Fusion Method for Feature Extraction from Sentinel Data”, *in*: *Remote Sensing* 10.*2* (Feb. 2018), p. 236, DOI: 10.3390/rs10020236, URL: https://doi.org/10.3390/rs10020236.

[166] Timothy J. Schmit et al., “A Closer Look at the ABI on the GOES-R Series”, *in*: *Bulletin of the American Meteorological Society* 98.*4* (July 2017), pp. 681–698, DOI: 10.1175/bams-d-15-00230.1, URL: https://doi.org/10.1175/bams-d-15-00230.1.

[167] J. Schulz-Stellenfleth, T. König, and S. Lehner, “An empirical approach for the retrieval of integral ocean wave parameters from synthetic aperture radar data”, *in*: *Journal of Geophysical Research* 112.*C3* (Mar. 2007), DOI: 10.1029/2006jc003970, URL: https://doi.org/10.1029/2006jc003970.

[168] Shinta SETO et al., “The Precipitation Rate Retrieval Algorithms for the GPM Dual-frequency Precipitation Radar”, *in*: *Journal of the Meteorological Society of Japan. Ser. II* 99.*2* (2021), pp. 205–237, DOI: 10.2151/jmsj.2021-011, URL: https://doi.org/10.2151/jmsj.2021-011.

[169] Weizeng Shao et al., “An Empirical Algorithm to Retrieve Significant Wave Height from Sentinel-1 Synthetic Aperture Radar Imagery Collected under Cyclonic Conditions”, *in*: *Remote Sensing* 10.*9* (Aug. 2018), p. 1367, DOI: 10.3390/rs10091367, URL: https://doi.org/10.3390/rs10091367.

[170] Hui Shen et al., “Detection of macroalgae blooms by complex SAR imagery”, *in*: *Marine Pollution Bulletin* 78.*1-2* (2014), pp. 190–195, DOI: 10.1016/j.marpolbul.2013.10.044, URL: https://doi.org/10.1016/j.marpolbul.2013.10.044.

[171] Olga V. Shomina et al., “Slick bands kinematics due to marine current and wind: study and simulation”, *in*: *Remote Sensing of the Ocean, Sea Ice, Coastal Waters, and Large Water Regions 2019*, ed. by Charles R. Bostater, Xavier Neyt, and Françoise Viallefont-Robinet, SPIE, Oct. 2019, DOI: 10.1117/12.2533150, URL: https://doi.org/10.1117/12.2533150.

[172] J. C. B. da Silva et al., “Role of surface films in ERS SAR signatures of internal waves on the shelf: 1. Short-period internal waves”, *in*: *Journal of Geophysical Research: Oceans* 103.*C4* (Apr. 1998), pp. 8009–8031, DOI: 10.1029/97jc02725, URL: https://doi.org/10.1029/97jc02725.

[173] Karen Simonyan and Andrew Zisserman, *Very Deep Convolutional Networks for Large-Scale Image Recognition*, 2014, DOI: 10.48550/ARXIV.1409.1556, URL: https://arxiv.org/abs/1409.1556.

[174] Abhineet Singh et al., “River Ice Segmentation With Deep Learning”, *in*: *IEEE Transactions on Geoscience and Remote Sensing* 58.*11* (Nov. 2020), pp. 7570–7579, DOI: 10.1109/tgrs.2020.2981082, URL: https://doi.org/10.1109/tgrs.2020.2981082.

[175] D. M. Smoth, “Extraction of winter total sea-ice concentration in the Greenland and Barents Seas from SSM/I data”, *in*: *International Journal of Remote Sensing* 17.*13* (Sept. 1996), pp. 2625–2646, DOI: 10.1080/01431169608949096, URL: https://doi.org/10.1080/01431169608949096.

[176] Serge Soula, “Lightning and Precipitation”, *in*: *Lightning: Principles, Instruments and Applications*, Springer Netherlands, pp. 447–463, DOI: 10.1007/978-1-4020-9079-0_20, URL: https://doi.org/10.1007/978-1-4020-9079-0_20.

[177] G. Spreen, L. Kaleschke, and G. Heygster, “Sea ice remote sensing using AMSR-E 89-GHz channels”, *in*: *Journal of Geophysical Research* 113.*C2* (Jan. 2008), DOI: 10.1029/2005jc003384, URL: https://doi.org/10.1029/2005jc003384.

[178] B. Stevens et al., “EUREC^{4}A”, *in*: *Earth System Science Data* 13.*8* (2021), pp. 4067–4119, DOI: 10.5194/essd-13-4067-2021, URL: https://essd.copernicus.org/articles/13/4067/2021/.

[179] Ad Stoffelen and David Anderson, “Scatterometer data interpretation: Estimation and validation of the transfer function CMOD4”, *in*: *Journal of Geophysical Research: Oceans* 102.*C3* (Mar. 1997), pp. 5767–5780, DOI: 10.1029/96jc02860, URL: https://doi.org/10.1029/96jc02860.

[180] Ad Stoffelen et al., “The CMOD7 Geophysical Model Function for ASCAT and ERS Wind Retrievals”, *in*: *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 10.*5* (May 2017), pp. 2123–2134, DOI: 10.1109/jstars.2017.2681806, URL: https://doi.org/10.1109/jstars.2017.2681806.

[181] J. E. Stopa and A. Mouche, “Significant wave heights from Sentinel-1 SAR: Validation and applications”, *in*: *Journal of Geophysical Research: Oceans* 122.*3* (Mar. 2017), pp. 1827–1848, DOI: 10.1002/2016jc012364, URL: https://doi.org/10.1002/2016jc012364.

[182] Carole H. Sudre et al., “Generalised Dice Overlap as a Deep Learning Loss Function for Highly Unbalanced Segmentations”, *in*: *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*, Springer Inter-

national Publishing, 2017, pp. 240–248, DOI: 10.1007/978-3-319-67558-9_28, URL: https://doi.org/10.1007/978-3-319-67558-9_28.

[183] Lin Sun et al., “Satellite data cloud detection using deep learning supported by hyperspectral data”, *in*: *International Journal of Remote Sensing* 41.*4* (Sept. 2019), pp. 1349–1371, DOI: 10.1080/01431161.2019.1667548, URL: https://doi.org/10.1080/01431161.2019.1667548.

[184] Xian Sun et al., “AIR-SARShip-1.0: High-resolution SAR Ship Detection Dataset (in English)”, *in*: *Journal of Radars* 8.*R19097* (2019), p. 852, ISSN: 2095-283X, DOI: 10.12000/JR19097, URL: https://radars.ac.cn/en/article/doi/10.12000/JR19097.

[185] J. Svejkovsky and J. Shandley, “Detection of offshore plankton blooms with AVHRR and SAR imagery”, *in*: *International Journal of Remote Sensing* 22.*2-3* (2001), pp. 471–485, DOI: 10.1080/014311601450040, URL: https://doi.org/10.1080/014311601450040.

[186] Christian Szegedy et al., “Going deeper with convolutions”, *in*: *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE, June 2015, DOI: 10.1109/cvpr.2015.7298594, URL: https://doi.org/10.1109/cvpr.2015.7298594.

[187] Christian Szegedy et al., “Rethinking the Inception Architecture for Computer Vision”, *in*: *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE, June 2016, DOI: 10.1109/cvpr.2016.308, URL: https://doi.org/10.1109/cvpr.2016.308.

[188] Simon P. de Szoeke et al., “Cold Pools and Their Influence on the Tropical Marine Boundary Layer”, *in*: *Journal of the Atmospheric Sciences* 74.*4* (Apr. 2017), pp. 1149–1168, DOI: 10.1175/jas-d-16-0264.1, URL: https://doi.org/10.1175/jas-d-16-0264.1.

[189] Alan A. Thompson∗, “Overview of the RADARSAT Constellation Mission”, *in*: *Canadian Journal of Remote Sensing* 41.*5* (2015), pp. 401–407, DOI: 10.1080/07038992.2015.1104633, URL: https://doi.org/10.1080/07038992.2015.1104633.

[190] Gavin H. Tilstone et al., "Performance of Ocean Colour Chlorophyll a algorithms for Sentinel-3 OLCI, MODIS-Aqua and Suomi-VIIRS in open-ocean waters of the Atlantic", *in*: *Remote Sensing of Environment* 260 (2021), p. 112444, DOI: `10.1016/j.rse.2021.112444`, URL: `https://doi.org/10.1016/j.rse.2021.112444`.

[191] Adrian M. Tompkins, "Organization of Tropical Convection in Low Vertical Wind Shears: The Role of Cold Pools", *in*: *Journal of the Atmospheric Sciences* 58.*13* (July 2001), pp. 1650–1672, DOI: `10.1175/1520-0469(2001)058<1650:ootcil>2.0.co;2`, URL: `https://doi.org/10.1175/1520-0469(2001)058<1650:ootcil>2.0.co;2`.

[192] Rasmus Tonboe and John Lavelle, *The EUMETSAT OSI SAF Sea Ice Concentration Algorithm, The EUMETSAT OSI SAF Sea Ice Concentration Algorithm Product OSI-401-b*, tech. rep., Ocean Sea Ice SAF, Danish Meteorological Institute, 2016.

[193] Rasmus Tonboe et al., *Validation Report for OSI SAF Global Sea Ice Concentration Product OSI-401-b*, tech. rep., Ocean Sea Ice SAF, Danish Meteorological Institute, 2016.

[194] Rasmus T. Tonboe et al., "The EUMETSAT sea ice concentration climate data record", *in*: *The Cryosphere* 10.*5* (Sept. 2016), pp. 2275–2290, DOI: `10.5194/tc-10-2275-2016`, URL: `https://doi.org/10.5194/tc-10-2275-2016`.

[195] Konstantinos Topouzelis, "Oil Spill Detection by SAR Images: Dark Formation Detection, Feature Extraction and Classification Algorithms", *in*: *Sensors* 8.*10* (Oct. 2008), pp. 6642–6659, DOI: `10.3390/s8106642`, URL: `10.3390/s8106642`.

[196] Giuseppe Torri, Zhiming Kuang, and Yang Tian, "Mechanisms for convection triggering by cold pools", *in*: *Geophysical Research Letters* 42.*6* (Mar. 2015), pp. 1943–1950, DOI: `10.1002/2015gl063227`, URL: `https://doi.org/10.1002/2015gl063227`.

[197] Paolo Trivero et al., "Automatic oil slick detection from SAR images: Results and improvements in the framework of the PRIMI pilot project", *in*: *Deep Sea Research Part II: Topical Studies in Oceanography* 133 (Nov. 2016), pp. 146–158, DOI: `10.1016/j.dsr2.2016.03.003`, URL: `https://doi.org/10.1016/j.dsr2.2016.03.003`.

[198] Julia Uitz et al., "Estimates of phytoplankton class-specific and total primary production in the Mediterranean Sea from satellite ocean color observations", *in*: *Global Biogeochemical Cycles* 26.*2* (2012), n/a–n/a, DOI: 10.1029/2011gb004055, URL: https://doi.org/10.1029/2011gb004055.

[199] Fawwaz T Ulaby, Richard K Moore, and Adrian K Fung, *Microwave remote sensing: From theory to applications v. 3*, Remote sensing library, Norwood, MA: Artech House, Feb. 1986.

[200] P.W. Vachon, J.A. Johannessen, and D.P. Browne, "ERS-1 SAR images of atmospheric gravity waves", *in*: *IEEE Transactions on Geoscience and Remote Sensing* 33.*4* (July 1995), pp. 1014–1025, DOI: 10.1109/36.406687, URL: https://doi.org/10.1109/36.406687.

[201] C. Vaduva, I. Gavat, and M. Datcu, "Deep learning in very high resolution remote sensing image information mining communication concept", *in*: *2012 Proceedings of the 20th European Signal Processing Conference (EUSIPCO)*, 2012, pp. 2506–2510.

[202] M Vargas-Yáñez et al., "About the seasonal variability of the Alboran Sea circulation", *in*: *Journal of Marine Systems* 35.*3-4* (2002), pp. 229–248, DOI: 10.1016/s0924-7963(02)00128-8, URL: https://doi.org/10.1016/s0924-7963(02)00128-8.

[203] Ashish Vaswani et al., *Attention Is All You Need*, 2017, DOI: 10.48550/ARXIV.1706.03762, URL: https://arxiv.org/abs/1706.03762.

[204] Paul Viola and Michael J. Jones, "Robust Real-Time Face Detection", *in*: *International Journal of Computer Vision* 57.*2* (May 2004), pp. 137–154, DOI: 10.1023/b:visi.0000013087.49260.fb, URL: https://doi.org/10.1023/b:visi.0000013087.49260.fb.

[205] Chen Wang et al., "A labelled ocean SAR imagery dataset of ten geophysical phenomena from Sentinel-1 wave mode", *in*: *Geoscience Data Journal* 6.*2* (July 2019), pp. 105–115, DOI: 10.1002/gdj3.73, URL: https://doi.org/10.1002/gdj3.73.

[206] Chen Wang et al., "An assessment of marine atmospheric boundary layer roll detection using Sentinel-1 SAR data", *in*: *Remote Sensing of Environment* 250

(Dec. 2020), p. 112031, DOI: 10.1016/j.rse.2020.112031, URL: https://doi.org/10.1016/j.rse.2020.112031.

[207] Chen Wang et al., “Classification of the global Sentinel-1 SAR vignettes for ocean surface process studies”, *in*: *Remote Sensing of Environment* 234 (Dec. 2019), p. 111457, DOI: 10.1016/j.rse.2019.111457, URL: https://doi.org/10.1016/j.rse.2019.111457.

[208] Chen Wang et al., *Labeled SAR imagery dataset of ten geophysical phenomena from Sentinel-1 wave mode (TenGeoP-SARwv)*, 2018, DOI: 10.17882/56796, URL: https://www.seanoe.org/data/00456/56796/.

[209] Lei Wang, K. Scott, and David Clausi, “Sea Ice Concentration Estimation during Freeze-Up from SAR Imagery Using a Convolutional Neural Network”, *in*: *Remote Sensing* 9.*5* (Apr. 2017), p. 408, DOI: 10.3390/rs9050408, URL: https://doi.org/10.3390/rs9050408.

[210] Lei Wang et al., “Sea Ice Concentration Estimation During Melt From Dual-Pol SAR Scenes Using Deep Convolutional Neural Networks: A Case Study”, *in*: *IEEE Transactions on Geoscience and Remote Sensing* 54.*8* (Aug. 2016), pp. 4524–4533, DOI: 10.1109/tgrs.2016.2543660, URL: https://doi.org/10.1109/tgrs.2016.2543660.

[211] Sherrie Wang et al., “Weakly Supervised Deep Learning for Segmentation of Remote Sensing Imagery”, *in*: *Remote Sensing* 12.*2* (Jan. 2020), p. 207, DOI: 10.3390/rs12020207, URL: https://doi.org/10.3390/rs12020207.

[212] Yuanyuan Wang et al., “A SAR Dataset of Ship Detection for Deep Learning under Complex Backgrounds”, *in*: *Remote Sensing* 11.*7* (Mar. 2019), p. 765, DOI: 10.3390/rs11070765, URL: https://doi.org/10.3390/rs11070765.

[213] Shunjun Wei et al., “HRSID: A High-Resolution SAR Images Dataset for Ship Detection and Instance Segmentation”, *in*: *IEEE Access* 8 (2020), pp. 120234–120254, DOI: 10.1109/access.2020.3005861, URL: https://doi.org/10.1109/access.2020.3005861.

[214] Sisi Wei et al., “Multi-Temporal SAR Data Large-Scale Crop Mapping Based on U-Net Model”, *in*: *Remote Sensing* 11.*1* (Jan. 2019), p. 68, DOI: 10.3390/rs11010068, URL: https://doi.org/10.3390/rs11010068.

[215] Samuel Williamson, *Doppler radar meteorological observations. Part B, Doppler Radar Theory and Meteorology*, tech. rep., FCM-H11B-2005, Office of the Federal Coordinator for Meteorological Services and Supporting Research, 2015.

[216] Fei Wu et al., "Inshore Ship Detection Based on Convolutional Neural Network in Optical Satellite Images", *in*: *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 11.*11* (Nov. 2018), pp. 4005–4015, DOI: `10.1109/jstars.2018.2873190`, URL: `https://doi.org/10.1109/jstars.2018.2873190`.

[217] Jia Xu, Alexander G. Schwing, and Raquel Urtasun, "Tell Me What You See and I Will Show You Where It Is", *in*: *2014 IEEE Conference on Computer Vision and Pattern Recognition*, IEEE, June 2014, DOI: `10.1109/cvpr.2014.408`, URL: `https://doi.org/10.1109/cvpr.2014.408`.

[218] Yuan Xu et al., "Change detection based on auto-encoder model for VHR images", *in*: *MIPPR 2013: Pattern Recognition and Computer Vision*, ed. by Zhiguo Cao, SPIE, Oct. 2013, DOI: `10.1117/12.2031104`, URL: `https://doi.org/10.1117/12.2031104`.

[219] Yufang Ye, Georg Heygster, and Mohammed Shokr, "Improving Multiyear Ice Concentration Estimates With Reanalysis Air Temperatures", *in*: *IEEE Transactions on Geoscience and Remote Sensing* 54.*5* (May 2016), pp. 2602–2614, DOI: `10.1109/tgrs.2015.2503884`, URL: `https://doi.org/10.1109/tgrs.2015.2503884`.

[220] David RL Yzenga, "The physical basis for estimating wave energy spectra from SAR imagery", *in*: *gan* 1001 (1987), p. 48107.

[221] Natalia Zakhvatkina et al., "Operational algorithm for ice–water classification on dual-polarized RADARSAT-2 images", *in*: *The Cryosphere* 11.*1* (Jan. 2017), pp. 33–46, DOI: `10.5194/tc-11-33-2017`, URL: `https://doi.org/10.5194/tc-11-33-2017`.

[222] F. De Zan and A. Monti Guarnieri, "TOPSAR: Terrain Observation by Progressive Scans", *in*: *IEEE Transactions on Geoscience and Remote Sensing* 44.*9* (Sept. 2006), pp. 2352–2360, DOI: `10.1109/tgrs.2006.873853`, URL: `https://doi.org/10.1109/tgrs.2006.873853`.

[223] Andrea Zanchetta and Stefano Zecchetto, “Wind direction retrieval from Sentinel-1 SAR images using ResNet”, *in*: *Remote Sensing of Environment* 253 (Feb. 2021), p. 112178, DOI: 10.1016/j.rse.2020.112178, URL: https://doi.org/10.1016/j.rse.2020.112178.

[224] Matthew D. Zeiler and Rob Fergus, “Visualizing and Understanding Convolutional Networks”, *in*: *Computer Vision – ECCV 2014*, Springer International Publishing, 2014, pp. 818–833, DOI: 10.1007/978-3-319-10590-1_53, URL: https://doi.org/10.1007/978-3-319-10590-1_53.

[225] Qiang Zhang et al., “Thick cloud and cloud shadow removal in multitemporal imagery using progressively spatio-temporal patch group deep learning”, *in*: *ISPRS Journal of Photogrammetry and Remote Sensing* 162 (Apr. 2020), pp. 148–160, DOI: 10.1016/j.isprsjprs.2020.02.008, URL: https://doi.org/10.1016/j.isprsjprs.2020.02.008.

[226] Wei Zhang et al., “Image processing of human corneal endothelium based on a learning network”, *in*: *Applied Optics* 30.*29* (1991), p. 4211, DOI: 10.1364/ao.30.004211, URL: https://doi.org/10.1364/ao.30.004211.

[227] Xueting Zhang et al., “Multi-Scale Cropping Mechanism for Remote Sensing Image Captioning”, *in*: *IGARSS 2019 - 2019 IEEE International Geoscience and Remote Sensing Symposium*, IEEE, 2019, DOI: 10.1109/igarss.2019.8900503, URL: https://doi.org/10.1109/igarss.2019.8900503.

[228] Zhanpeng Zhang et al., “Facial Landmark Detection by Deep Multi-task Learning”, *in*: *Computer Vision – ECCV 2014*, Springer International Publishing, 2014, pp. 94–108, DOI: 10.1007/978-3-319-10599-4_7, URL: https://doi.org/10.1007/978-3-319-10599-4_7.

[229] Zhao et al., “Use of Unmanned Aerial Vehicle Imagery and Deep Learning UNet to Extract Rice Lodging”, *in*: *Sensors* 19.*18* (Sept. 2019), p. 3859, DOI: 10.3390/s19183859, URL: https://doi.org/10.3390/s19183859.

[230] Yan Zhao et al., “Attention Receptive Pyramid Network for Ship Detection in SAR Images”, *in*: *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 13 (2020), pp. 2738–2756, DOI: 10.1109/jstars.2020.2997081, URL: https://doi.org/10.1109/jstars.2020.2997081.

[231] Yuan Zhao, Xiao-Ming Li, and Jin Sha, "Sea surface wind streaks in spaceborne synthetic aperture radar imagery", *in*: *Journal of Geophysical Research: Oceans* 121.*9* (Sept. 2016), pp. 6731–6741, DOI: `10.1002/2016jc012040`, URL: `https://doi.org/10.1002/2016jc012040`.

[232] Yuan Zhao et al., "Automated Rain Detection by Dual-Polarization Sentinel-1 Data", *in*: *Remote Sensing* 13.*16* (Aug. 2021), p. 3155, DOI: `10.3390/rs13163155`, URL: `https://doi.org/10.3390/rs13163155`.

[233] Bolei Zhou et al., "Learning Deep Features for Discriminative Localization", *in*: *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE, June 2016, DOI: `10.1109/cvpr.2016.319`, URL: `https://doi.org/10.1109/cvpr.2016.319`.

[234] Zongwei Zhou et al., "A Nested U-Net Architecture for Medical Image Segmentation", *in*: *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*, Springer International Publishing, 2018, pp. 3–11, DOI: `10.1007/978-3-030-00889-5_1`, URL: `https://doi.org/10.1007/978-3-030-00889-5_1`.

[235] Xiao Xiang Zhu et al., "Deep Learning in Remote Sensing: A Comprehensive Review and List of Resources", *in*: *IEEE Geoscience and Remote Sensing Magazine* 5.*4* (Dec. 2017), pp. 8–36, DOI: `10.1109/mgrs.2017.2762307`, URL: `https://doi.org/10.1109/mgrs.2017.2762307`.

[236] Xiao Xiang Zhu et al., "Deep Learning Meets SAR: Concepts, models, pitfalls, and perspectives", *in*: *IEEE Geoscience and Remote Sensing Magazine* 9.*4* (Dec. 2021), pp. 143–172, DOI: `10.1109/mgrs.2020.3046356`, URL: `https://doi.org/10.1109/mgrs.2020.3046356`.

[237] Qin Zou et al., "DeepCrack: Learning Hierarchical Convolutional Features for Crack Detection", *in*: *IEEE Transactions on Image Processing* 28.*3* (2019), pp. 1498–1512, DOI: `10.1109/tip.2018.2878966`, URL: `https://doi.org/10.1109/tip.2018.2878966`.

[238] H. Jay Zwally and Per Gloersen, "Passive microwave images of the polar regions and research applications", *in*: *Polar Record* 18.*116* (May 1977), pp. 431–450, DOI: `10.1017/s0032247400000930`, URL: `https://doi.org/10.1017/s0032247400000930`.

GLOSSARY

ABI Advanced Baseline Imager. A radiometer observing the Earth with 16 spectral bands between 450 and 1300 nm. . 19, 84, 184

ENL Equivalent Number of Looks. Number of looks that would produce similar speckle than the product. The higher this number is, the lower both the speckle and the resolution are. . 27, 184

EW Extra-Wideswath. The default observation mode of the Sentinel-1 mission in polar areas. EW are products of 400 km in range. . 27, 81, 184

GAN Generative Adversarial Network. A model architecture in which the loss (or part of the loss) is not defined explicitly but estimated from a secondary model (called the discriminator) training simultaneously. . 42, 184

GLM Geostationary Lightning Mapper. A near-infrared optical sensor onboard the GOES-16 and GOES-17 satellites. It aims at detecting lightning events. . 18, 19, 184

GMF Geophysical Model Function. In the context of SAR imagery, it refers to a function linking the wind speed, its direction relative to the satellite heading, the incidence angle, and the reflectivity. It is notably used to retrieve the wind speed from the SAR observation, assuming a wind direction obtained by a atmospheric model.. 132, 143, 184

GOES-16/17 Geostationary Operational Environmental Satellite 16 and 17. Part of the GOES constellation. Both are geostationnary satellites and embark a radiometer (ABI) and a near-infrared detector (GLM). . 19, 184

GPM-DPR Global Precipitation Measurement/Dual-frequency Precipitation Radar. A weather radar aboard the GPM-core satellite. . 18, 19, 133, 184

GRD Ground Range Detected. A SAR product where the observation have been projected to the ground using a Earth model, with a square spatial resolution. . 19, 25, 49, 133, 153, 184

incidence angle The angle between the direction of the measurement and the nadir. . 18, 53, 132, 133, 184

IW Interferometric Wideswath. The default observation mode of the Sentinel-1 mission in coastal areas. IW are products of 250 km in range. . 19, 27, 46, 80, 132, 133, 184

L2 Level-2 Product. Geolocated geophysical products derived from an observation. . 50, 133, 184

MAE Mean Absolute Error. An error function defined as $MAE(x, \widetilde{x}) = \frac{1}{n} \sum_{i=0}^{n} abs(x - \widetilde{x})$. 138, 139, 184

metocean Contraction of meteorologic and oceanic processes. Processes than originate either from atmospheric or oceanic conditions. . 8, 184

MSE Mean Square Error. An error function defined as $MSE(x, \widetilde{x}) = \frac{1}{n} \sum_{i=0}^{n} (x_i - \widetilde{x_i})^2$. 82, 83, 84, 184

NDBC National Data Buoy Center. An american network of buoys and coastal stations. . 17, 184

NEXRAD NEXt generation RADar. An american network of ground-based weather radars. . 17, 133, 184

NRCS The Normalized Radar Cross-Section, also called σ_0, is the radiometric calibrated radar cross-section in respect to the thermal noise, the incidence angle, and various other geophysical parameters [108]. . 22, 184

off-nadir Observation performed with a non-null incidence angle. . 18, 184

OLCI Ocean and Land Colour Instrument. A spectrometer using 21 bands from 400 nm to 1070 nm used notably to measure the ocean colour. . 19, 50, 184

RMSE Root Mean Square Error. An error function defined as $RMSE(x, \widetilde{x}) = \sqrt{\frac{1}{n} \sum_{i=0}^{n} (x_i - \widetilde{x_i})^2}$. 142, 184

SAR Synthetic Aperture Radar. An imaging technique relying on the movement of the sensor to provide high-resolution observations. . 19, 45, 79, 132, 184

Sentinel-3 A satellite serie part of the Copernicus Program. It embarks the OLCI used to estimate, *inter alia*, the chlorophyll concentration. . 19, 50, 184

SLC Single Look Complex. A SAR product where the observation is in the image plane of satellite data acquisition, and without multi-looking. Both phase and amplitude information are conserved, but the speckle is higher than in GRD products. . 19, 25, 184

Weakly-supervised learning A machine learning strategy under which some characterics of the desired output are better than those in the actual groundtruth (for example, higher spatial resolution or lower noise). . 44, 184

WV Wave Mode. The default observation mode in open ocean. WV products are observations of 20x20 km. . 19, 27, 79, 184

RELATED PUBLICATIONS

The following is a list of publications to which this manuscript's author contributed as either author or co-author.

Accepted papers

- Colin, A., Fablet, R., Tandeo, P., Husson, R., Peureux, C., Longépé, N., & Mouche, A. (2022). Semantic Segmentation of Metoceanic Processes Using SAR Observations and Deep Learning. In Remote Sensing (Vol. 14, Issue 4, p. 851). MDPI AG. https://doi.org/10.3390/rs14040851
- de Gelis, I., Colin, A., & Longepe, N. (2021). Prediction of Categorized Sea Ice Concentration From Sentinel-1 SAR Images Based on a Fully Convolutional Network. In IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing (Vol. 14, pp. 5831–5841). Institute of Electrical and Electronics Engineers (IEEE). https://doi.org/10.1109/jstars.2021.3074068

Conference papers

- Colin, A., Peureux, C., Husson, R., Fablet, R., & Tandeo, P. (2022). Segmentation of Rainfall Regimes by Machine Learning on a Colocalized Nexrad/Sentinel-1 Dataset. In IGARSS 2022 - 2022 IEEE International Geoscience and Remote Sensing Symposium. IGARSS 2022 - 2022 IEEE International Geoscience and Remote Sensing Symposium. IEEE. https://doi.org/10.1109/igarss46834.2022.9884881
- Colin, A., Peureux, C., Husson, R., Longepe, N., Rauzy, R., Fablet, R., Tandeo, P., Saoudi, S., Mouche, A., & Dibarboure, G. (2021). Segmentation of Sentinel-1 SAR Images Over the Ocean, Preliminary Methods and Assessments. In 2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS. IGARSS 2021 - 2021 IEEE International Geoscience and Remote Sensing Symposium. IEEE. https://doi.org/10.1109/igarss47720.2021.9553429

- Husson, R., Longepe, N., Mouche, A., Berger, H., Lin, C., Archer, O., & Colin, A. (2021). Wind Direction Estimation and Accuracy Retrieval from Sentinel-1 SAR Images Under Thermal and Dynamical Unstable Conditions. In 2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS. IGARSS 2021 - 2021 IEEE International Geoscience and Remote Sensing Symposium. IEEE. https://doi.org/10.1109/igarss47720.2021.9555045

Papers in preparation

- Colin, A., Tandeo, P., Peureux, C., Husson, R., Longepe, N., & Fablet, R. (2022). Rain Rate Estimation with SAR using NEXRAD measurements with Convolutional Neural Networks. arXiv. https://doi.org/10.48550/ARXIV.2207.07333
- Colin, A., Tandeo, P., Peureux, C., Husson, R., & Fablet, R. (2023). Reduction of rain-induced errors for wind speed estimation on SAR observations using convolutional neural networks.
- Colin, A., Tandeo, P., , Husson, R., & Fablet, R. Peureux, C. (2023). MediSAR: an exhaustive dataset of Sentinel-1 SAR observations of the Mediterranean Sea and the Black Sea regions.

DOCTORAT
BRETAGNE
LOIRE / MATHSTIC

IMT Atlantique
Bretagne-Pays de la Loire
École Mines-Télécom

Titre : De l'Utilisation de l'Apprentissage profond pour la Segmentation d'Images SAR océaniques

Mot clés : Télédetection, Radar à Synthèse d'Ouverture, Apprentissage Profond, Océanographie

Résumé : Nombreux sont les phénomènes météorologiques prenant naissance dans les océans, dont 71% de la surface de la Terre est couverte. L'observation des étendues marines est primordiale pour accroître notre compréhension du système Terre, améliorer les modèles météorologiques et atténuer l'effet des catastrophes naturelles. Depuis le lancement d'ERS-1 en 1991, les radars à synthèse d'ouverture (SAR, d'après l'acronyme anglais) en bande C sont utilisés pour observer les surfaces océaniques. La bande C est, en effet, particulièrement utile pour obtenir des informations sur les vagues, le vent, les précipitations, la banquise, que ce soit à méso- ou à sous-méso échelle.

La thèse ci-présente traite de la segmentation, c'est-à-dire de la classification pixel par pixel, des observations SAR en bande C dans un contexte océanographique. Les modèles d'apprentissage profond permettant d'inférer des informations à partir de larges ensembles de données, nous mettons à profit les acquisitions produites par les satellites Sentinel-1A et Sentinel-1B pour générer des cartes de segmentation. Ces deux satellites du programme Copernicus ont été lancés au cours des années 2014 et 2016. Les observations du capteur SAR sont co-localisées avec des sources secondaires (par exemple des radars côtiers, des bouées ou encore d'autres satellites), parfois manuellement annotées, ou encore reliées à des modèles météorologiques. Elles sont ensuite utilisées pour entraîner des modèles d'apprentissage profond et pour s'assurer de leurs performances à travers de multiples comparaisons.

Ces études mettent en exergue la possibilité pour l'imagerie SAR de déduire de nouvelles informations prometteuses. Elles proposent également un cadre pour la construction de jeux de données de segmentation et de l'entraînement des modèles afférents.

Title: On the Use of Deep Learning for Ocean SAR Image Semantic Segmentation

Keywords: Remote Sensing, Synthetic Aperture Radar, Deep Learning, Oceanography

Abstract: With water covering 71% of the surface of the Earth, and most meteorological processes stemming from the oceans, their observation is primordial to enhance our comprehension of the Earth system, improve meteorological models, and prevent hazards. Since ERS-1 (launched in 1991), C-Band Synthetic Aperture Radar (SAR) has been used to observe the ocean surfaces. This particular electromagnetic band is especially useful for deriving information on waves, winds, precipitation, sea ice, and more at meso- and sub-mesoscale.

The subject of this thesis is the segmentation, or pixel-per-pixel classification, of the ocean surface C-Band SAR observations. The generation of segmentation maps is possible through the use of machine learning frameworks that are able to extract information from the large data volume produced by the satellites Sentinel-1A and Sentinel-1B, which were launched in 2014 and 2016 as part of ESA's Copernicus program. These observations are collocated with third-party sensors (ground stations, buoys, satellite-boarded instruments), manually annotated segmentations, or meteorological models to be able to train deep learning models and ensure their capacity through extensive tests.

These studies show promising uses of new SAR-derived information and propose guidelines for building dedicated segmentation datasets and models.