Rain Rate Estimation with SAR using NEXRAD measurements with Convolutional Neural Networks

Aurélien COLIN, Pierre TANDEO, Charles PEUREUX, Romain HUSSON, Nicolas LONGEPE, Ronan FABLET

a IMT Atlantique, Lab-STICC, UMR CNRS 6285, F-29238, France
b Collecte Localisation Satellites, Brest, France
c Φ-lab Explore Office, ESRIN, European Space Agency (ESA), Frascati, Italy

Corresponding author: Aurélien COLIN, acolin@groupcls.com
ABSTRACT: Remote sensing of rainfall events is critical for both operational and scientific needs, including for example weather forecasting, extreme flood mitigation, water cycle monitoring, etc. Ground-based weather radars, such as NOAA’s Next-Generation Radar (NEXRAD), provide reflectivity and precipitation measurements of rainfall events. However, the observation range of such radars is limited to a few hundred kilometers, prompting the exploration of other remote sensing methods, particularly over the open ocean, that represents large areas not covered by land-based radars. For a number of decades, C-band SAR imagery such as Sentinel-1 imagery has been known to exhibit rainfall signatures over the sea surface. However, the development of SAR-derived rainfall products remains a challenge. Here we propose a deep learning approach to extract rainfall information from SAR imagery. We demonstrate that a convolutional neural network, such as U-Net, trained on a colocated and preprocessed Sentinel-1/NEXRAD dataset clearly outperforms state-of-the-art filtering schemes. Our results indicate high performance in segmenting precipitation regimes, delineated by thresholds at 1, 3, and 10 mm/h. Compared to current methods that rely on Koch filters to draw binary rainfall maps, these multi-threshold learning-based models can provide rainfall estimation for higher wind speeds and thus may be of great interest for data assimilation weather forecasting or for improving the qualification of SAR-derived wind field data.
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 (Douville et al. 2021). 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 (Marshall et al. 1947). In previous work (Miyoshi et al. 2016), data assimilation schemes have been investiogated in order to reconstruct precipitation patterns from weather radar observation data.

Farther from the coast, through the use of satellite instruments, rain events can also be observed, though at lower spatial resolution. For example, brightness temperatures measured by microwave radiometers (Liu 2003), 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 (Jackson 2004). 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; (Melshelmer et al. 1996; Clemente-Colon et al. 1999; Lin et al. 2001; Melshelmer et al. 2001). The preparation and analysis of well colocated 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 (Wang et al. 2019b; Colin et al. 2022), 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. As illustrated by Zhao et al. (2021), 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.

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 preprocessing 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 Zhao et al. (2021). 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.

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.

### a. 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 (Wang et al. 2019a). 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-Range Weather Forecast (ECMWF) model estimates and wind inversions from SAR observations.

1. **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).

2. **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 (Mouche et al. 2017).

3. **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. As highlighted in Alpers et al. (2016), four physical processes contribute to the radar signature of rainfall events: 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
Figure 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.

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.

Figure 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.
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.

\textbf{b. SAR preprocessing}

Unlike Stripmap, in which the SAR sensor observes the entire swath simultaneously, TOPSAR (Zan and Guarnieri 2006) 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) (Hersbach 2010) 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 Wang et al. (2019b).

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-oceanic or land phenomena and would vary from one SAR product to another.

\textbf{c. Enhanced colocalized Sentinel-1/NEXRAD dataset}

Several improvement steps, described below, were applied to the initial colocations performed by Zhao et al. (2021) 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 farther 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-oceanic (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. As explained by Alpers et al. (2016), the bright rain signatures may also be due to strong reflectivity from the melting layer (4th physical process aforementioned), 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 (Williamson 2015), 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. Colocations 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 metoceanic 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 (Williamson 2015)). 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 Figure 2. This geographic repositioning corrects for remaining colocation 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 Figure 3.

Figure 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).

To determine why the colocations do not overlap perfectly, statistics were computed 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, colocation, 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 Figure 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 1. As noted, the dataset suffers a lack of data at higher reflectivity and wind speeds. Indeed, as 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.
Figure 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 over-estimating the wind speed for this particular case.

The reflectivity is divided into four intervals: [0, 24.7], [24.7, 31.5], [31.5, 38.8] and [38.8, +∞] dBZ. With the general NEXRAD radar formula (Williamson 2015), these intervals can be approximated in terms of rainfall by [0, 1], [1, 3], [3, 10], and [10, +∞] 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: ≥ 24.7 dBZ, ≥ 31.5 dBZ, and ≥ 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 Figure 4.b.

(i) GOES-16 Geostationary Lightning Mapper As illustrated in Table 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 colocated lightning data set was constructed. It has been demonstrated that precipitation can be approximated by a linear function of the lightning frequency (Soula 2009). The quality of this approximation can be very variable depending on the type of precipitation (convective/stratiform), the location (land/ocean) or the climate of the studied area (humid/arid). Still, the colocalization between weather radars and lightning
activity was also explored in a data assimilation framework by Lagouvardos et al. (2013), 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 (Goodman et al. 2013). 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

| 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 %          |
| ≥ 38.8 dBZ ≥ 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 %          |
| ≥ 16 m/s                         | 100 %         | 0.0 %             | 0.0 %       | 0.1 %          |
with a pair of events less than 330 ms and 16.5 km apart). In this study, only colocations 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.

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 (Koch 2004), the second one based on a U-Net architecture (Ronneberger et al. 2015).

a. Koch-filter-based architecture

Koch filters, introduced by Koch (2004), 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 (Zhao et al. 2021).

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. Zhao et al. (2021) 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. Figure 5 illustrates this Koch filter incorporated into the CNN. This formulation allows for segmentations for different rain thresholds, unlike the original rain detection (Zhao et al. 2021).
Figure 5. Architecture of the multi-threshold Koch filter as a convolutional neural network. $p_{i}^{VV,r}$ is the output of the high-pass filter $i$ at the resolution $r$, for the $VV$ polarization. $\sigma$ is the sigmoid function defined as $\sigma(x) = \left[ 1 + \exp(-4(x + 0.5)) \right]^{-1}$. $K_{j}^{VV,r}$ and $B_{j}^{VV,r}$ 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 (Koch 2004).

b. 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 (Ronneberger et al. 2015). 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 (de Gélis et al. 2021) and semantic segmentation (Colin et al. 2022).

The specific model used is shown in Figure 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.
(Luo et al. 2017), is small enough to ensure continuity. The number of weights, independent of input size and spatial resolution, was 3,117,731.

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

| Layer Type  | Configuration  |
|-------------|----------------|
| ConvBlock_32| (x, 3, 3, 32)  |
| DeConvBlock_32| (x, 3, 3, 32) |
| ConvBlock_64| (x, 3, 256)    |
| DeConvBlock_64| (x, 1, 1, 3)  |

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 at [https://github.com/CIA-Oceanix/SAR-Segmentation/tree/oceanix](https://github.com/CIA-Oceanix/SAR-Segmentation/tree/oceanix).
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-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 colocated 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 with Zhao et al. (2021). 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, in accordance with Bengio (2012).

5. Results and discussion

Table 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 multiclass segmentation. The binary Koch filter performs worse on the binary F1 score at each
| Model                        | Input resolution | Binary F1-score (> 1 mm/h) | Binary F1-score (> 3 mm/h) | Binary F1-score (> 10 mm/h) | Multiclass F1-score |
|-----------------------------|------------------|----------------------------|-----------------------------|-----------------------------|---------------------|
| 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 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.

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 Zhao et al. (2021), 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 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 km². 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 Figure 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 remains unrecognized by the model as being rain. A large standard deviation in the confusion matrix indicates instability in the training. The U-
Figure 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-registered dataset (bottom), each at 400 m/px.

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.

Figure 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
Figure 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 colocated 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.

rain patches, not detected in NEXRAD. Three neighbouring 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 estimate the sensitivity to incidence angle, wind speed, and distance from shore. The result of this analysis is compiled in Figure 9. Figures 9.a and 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 (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 (figure 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 (Figure 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.

![Figure 9](image)

Figure 9. Probability of rain detection as a function of wind speed (a) and distance to the coast (d). F1-score as a function of wind speed (b) and wind speed (c), using GLM as a rain proxy.

Figure 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.

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 colocated 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.

Acknowledgments. We thank Alexis Mouche (Laboratoire d’Océanographie Physique et Spatiale, Ifremer) for the access to colocalizations between NEXRAD and Sentinel-1.
References

Alpers, W., B. Zhang, A. Mouche, K. Zeng, and P. W. Chan, 2016: Rain footprints on c-band synthetic aperture radar images of the ocean - revisited. Remote Sensing of Environment, 187, 169–185, https://doi.org/10.1016/j.rse.2016.10.015, URL https://doi.org/10.1016/j.rse.2016.10.015.

Bengio, Y., 2012: Practical recommendations for gradient-based training of deep architectures. CoRR, abs/1206.5533, URL http://arxiv.org/abs/1206.5533, 1206.5533.

Clemente-Colon, P., P. C. Manousos, W. G. Pichel, and K. S. Friedman, 1999: Observations of Hurricane Bonnie in spaceborne synthetic aperture radar (SAR) and next-generation Doppler weather radar (NEXRAD). Satellite Remote Sensing of Clouds and the Atmosphere IV, J. E. Russell, Ed., SPIE, International Society for Optics and Photonics, Vol. 3867, 63–70, URL https://doi.org/10.1117/12.373044.

Colin, A., R. Fablet, P. Tandeo, R. Husson, C. Peureux, N. Longépé, and A. Mouche, 2022: Semantic segmentation of metoceanic processes using SAR observations and deep learning. Remote Sensing, 14 (4), 851, https://doi.org/10.3390/rs14040851, URL https://doi.org/10.3390/rs14040851.

de Gélis, I., A. Colin, and N. Longépé, 2021: Prediction of categorized sea ice concentration from sentinel-1 sar images based on a fully convolutional network. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 14, 5831–5841, https://doi.org/10.1109/JSTARS.2021.3074068.

Douville, H., and Coauthors, 2021: 2021 : Water cycle changes. 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.

Goodman, S. J., and Coauthors, 2013: The goes-r geostationary lightning mapper (glm). Atmospheric Research, 125-126, 34–49, https://doi.org/https://doi.org/10.1016/j.atmosres.2013.01.006, URL https://www.sciencedirect.com/science/article/pii/S0169809513000434.

Hersbach, H., 2010: Comparison of c-band scatterometer CMOD5.n equivalent neutral winds with ECMWF. Journal of Atmospheric and Oceanic Technology, 27 (4), 721–736, https://doi.org/10.1175/2009jtecho698.1, URL https://doi.org/10.1175/2009jtecho698.1.
Jackson, C., 2004: *Synthetic aperture radar: marine user’s manual*. U.S. Dept. of Commerce, National Oceanic and Atmospheric Administration, National Environmental Satellite, Data, and Information Serve, Office of Research and Applications, Washington, D.C.

Koch, W., 2004: Directional analysis of SAR images aiming at wind direction. *Remote Sens. Environ.*, **42** (4), 702–710, https://doi.org/10.1109/tgrs.2003.818811, URL https://doi.org/10.1109/tgrs.2003.818811.

Lagouvardos, K., V. Kotroni, E. Defer, and O. Bousquet, 2013: Study of a heavy precipitation event over southern france, in the frame of HYMEX project: Observational analysis and model results using assimilation of lightning. *Atmospheric Research*, **134**, 45–55, https://doi.org/10.1016/j.atmosres.2013.07.003, URL https://doi.org/10.1016/j.atmosres.2013.07.003.

Lin, I.-I., W. Alpers, V. Khoo, H. Lim, T. Lim, and D. Kasilingam, 2001: An ers-1 synthetic aperture radar image of a tropical squall line compared with weather radar data. *IEEE Transactions on Geoscience and Remote Sensing*, **39** (5), 937–945, https://doi.org/10.1109/36.921411.

Liu, G., 2003: SATELLITE REMOTE SENSING | precipitation. *Encyclopedia of Atmospheric Sciences*, Elsevier, 1972–1979, https://doi.org/10.1016/b0-12-227090-8/00352-3, URL https://doi.org/10.1016/b0-12-227090-8/00352-3.

Luo, W., Y. Li, R. Urtasun, and R. Zemel, 2017: Understanding the effective receptive field in deep convolutional neural networks. 1701.04128.

Marshall, J. S., R. C. Langille, and W. M. K. Palmer, 1947: MEASUREMENT OF RAINFALL BY RADAR. *Journal of Meteorology*, **4** (6), 186–192, https://doi.org/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.

Melsheimer, C., W. Alpers, and M. Gade, 2001: Simultaneous observations of rain cells over the ocean by the synthetic aperture radar aboard the ers satellites and by surface-based weather radars. *Journal of Geophysical Research*, **106**, 4665–4678, https://doi.org/10.1029/2000JC000263.

Melsheimer, C., W. Alpers, and M. Gade, 1996: Investigation of multifrequency/multipolarization radar signatures of rain cells, derived from sir-c/x-sar data.
Miyoshi, T., and Coauthors, 2016: “big data assimilation” revolutionizing severe weather prediction. *Bulletin of the American Meteorological Society*, 97 (8), 1347–1354, https://doi.org/10.1175/bams-d-15-00144.1, URL https://doi.org/10.1175/bams-d-15-00144.1.

Mouche, A., P. Vincent, and G. Hajduch, 2017: Sentinel-1 ocean wind fields (owi) algorithm definition. Tech. rep. URL https://sentinels.copernicus.eu/documents/247904/2142675/Thermal-Denoising-of-Products-Generated-by-Sentinel-1-IPF.pdf.

Ronneberger, O., P. Fischer, and T. Brox, 2015: U-net: Convolutional networks for biomedical image segmentation. *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, Springer, LNCS, Vol. 9351, 234–241, URL http://lmb.informatik.uni-freiburg.de/Publications/2015/RFB15a, (available on arXiv:1505.04597 [cs.CV]).

Soula, S., 2009: Lightning and precipitation. *Lightning: Principles, Instruments and Applications*, Springer Netherlands, 447–463, https://doi.org/10.1007/978-1-4020-9079-0_20, URL https://doi.org/10.1007/978-1-4020-9079-0_20.

Wang, C., and Coauthors, 2019a: Classification of the global sentinel-1 SAR vignettes for ocean surface process studies. *Remote Sensing of Environment*, 234, 111457, https://doi.org/10.1016/j.rse.2019.111457, URL https://doi.org/10.1016/j.rse.2019.111457.

Wang, C., and Coauthors, 2019b: A labelled ocean SAR imagery dataset of ten geophysical phenomena from sentinel-1 wave mode. *Geoscience Data Journal*, 6 (2), 105–115, https://doi.org/10.1002/gdj3.73, URL https://doi.org/10.1002/gdj3.73.

Williamson, S., 2015: 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.

Zan, F. D., and A. M. Guarnieri, 2006: Topsar: Terrain observation by progressive scans. *IEEE Transactions on Geoscience and Remote Sensing*, 44 (9), 2352–2360, https://doi.org/10.1109/TGRS.2006.873853.
Zhao, Y., N. Longépé, A. Mouche, and R. Husson, 2021: Automated rain detection by dual-polarization sentinel-1 data. Remote Sensing, 13 (16), https://doi.org/10.3390/rs13163155, URL https://www.mdpi.com/2072-4292/13/16/3155.