Segmentation of Rainfall Regimes by Machine Learning on a Colocalized Nexrad/Sentinel-1 Dataset

Aurélien Colin, Charles Peureux, Romain Husson, Ronan Fablet, Pierre Tandeo

SEGMENTATION OF RAINFALL REGIMES BY MACHINE LEARNING ON A COLOCALIZED NEXRAD/SENTINEL-1 DATASET

Aurélien Colin[1,2], Charles Peureux[1], Romain Husson[1], Ronan Fablet[2], Pierre Tandeo[2]
[1]Collecte Localisation Satellites, Brest, France
[2]IMT Atlantique, Lab-STICC, UMR CNRS 6285, F-29238, France

ABSTRACT

Precipitation measurement is an important prior for several operational and scientific applications, including weather forecasting, hazard prevention, agriculture, etc. Weather radars, such as NEXRAD, observe the air volume reflectivity and infer precipitation intensity at high resolution. However, their capabilities are limited over the ocean. C-band SAR imagery, which is sensitive to ocean surface roughness, is known to be sensitive to the effect of rain. In this study, we improve existing NEXRAD/Sentinel-1 collocations and train a U-Net deep learning model to estimate NEXRAD radar reflectivity from Sentinel-1 observations. Precipitation forecasts are returned as segmentations with thresholds at 1, 3 and 10 mm/hr. The results indicate high performance over a wide range of wind speeds and thus can provide an accurate estimate of precipitation in the absence of weather radar.

Index Terms— SAR, Sentinel-1, NEXRAD, ocean

1. INTRODUCTION

Remote sensing of precipitation is of interest for a wide range of applications, from weather forecasting to agriculture. It can be measured by space satellites (GPM, TRMM, ...) or by ground-based weather radars. Weather radars are particularly interesting because their fixed position simplifies co-location with other instruments, but they have a short range and are affected by topography.

Synthetic aperture radars (SARs) are space-based imaging systems capable of measuring sea surface roughness at high resolution. Of these instruments, Sentinel-1 A and B, operated by the European Space Agency, have been acquiring data regularly since 2014 and 2016, respectively. [1] recently demonstrated that it is possible to apply Koch filters, usually used to detect heterogeneous areas in SAR observations, to detect the presence of rain, based on a collocated Sentinel-1/NEXRAD dataset.

In this study, we improve the dataset proposed in [1], adapt Koch filters to multiclass segmentation, and train a deep learning model. The resulting model is able to accurately detect different rainfall patterns, outperforming the Koch filters, although being sensitive to wind speed.

2. DATA

The dataset is composed of both SAR observations and weather radar measurements, acting as the input and output groundtruths respectively. The SAR observations were acquired as part of the Sentinel-1 mission, consisting of two satellites, Sentinel-1A and Sentinel-1B, whose instruments routinely acquire C-band (5.4 GHz) observations. Specifically, we use Interferometric Wide Swath (IW) Ground Range Detected High Resolution products (GRD-HR). These observations, with a spatial resolution of 20x22 m, span about 250 km in range and a few hundred in azimuth. The weather radar measurements are obtained from NEXRAD, a network of Doppler weather radars with a bandwidth between 2.7 and 3 GHz. The resolution is 1 km in range and 1° in azimuth. The initial collocations were performed by [1].

The collocated IWs are divided into 256x256 pixel patches and scaled to 100 m/px. Patches are removed if the collocated NEXRAD measurement does not indicate precipitation greater than 1 mm/h in any part of the area. This step prevents occlusions by topography or overestimation of the NEXRAD range from introducing undetected rain patches. Once extracted, the patches are inspected to manually ensure overlap between the rain signature on the SAR observation and the NEXRAD measurement. The manual correction of the collocations shows that the alignment error increases with the distance to the radar (R^2 = 0.404). On the other hand, neither the direction of the NEXRAD ground stations, nor the wind speed, nor its direction (obtained from the ECMWF) are correlated with the realignment vector.

After the manual realignment, the data are divided into training (79.5% of the patches), validation (9.6%) and test (10.9%) subsets. The dataset is divided at the IW level to eliminate all data leakage and balanced at the pixel level to have the same distribution of precipitation and wind speed in each subset. The number of IWs for the training, validation, and test subsets are 39, 7, and 7, respectively. The total number of patches in the data set is 1570.

Finally, the output ground truths are thresholded to provide precipitation segmentations for the intervals [1, $+\infty$], [3, $+\infty$] and [10, $+\infty$] mm/h.

A secondary dataset containing Sentinel-1 observations collocated with the Geostationnary Lightning Mapper (GLM) boarded on GOES-16 is also built. Although it does not contain precipitation information, lightning is known to be closely related to precipitation: [2, 3]. Furthermore, because GLM covers the entire Western Hemisphere with continuous observations, a large number of collocations can be obtained (189 IWs while only 7 were present in the test subset). These collocations are used to evaluate the impact of wind speed and incidence angle.

3. METHODS

Koch filters [4] are a standard filtering method used in SAR imaging. They consist of four different high-pass sub-filters that detect heterogeneous areas and indicate non-wind related phenomena. Koch filter has been shown to specifically detect rain by optimizing its thresholds on a NEXRAD/Sentinel-1 dataset [1]. To be able to compare the Koch filter in the context of multiclass segmentation, we tine-tune the Koch filter parameters on the improved NEXRAD/Sentinel-1 collocations, training new filters for each rain regime.

The deep learning model uses the U-Net architecture [5]. It has been successfully applied for semantic segmentation [6] and sea ice concentration estimation [7]. Our model contains three convolution blocks (each containing three convolutional layers activated by ReLUs) of 32, 64, and 128 3x3 kernels, respectively. Each block is followed by 2x2 max pooling layers. The central part is similarly composed of convolution layers with 256 3x3 kernels. The upsample part of the network is the symmetric of the downsample, as is the case with U-Net architectures. The output layer is a convolution layer with three 1x1 kernels, activated by a sigmoid function.

To compare the binary Koch filters, originally designed for binary segmentation, each rain rate threshold is considered as a threshold for binary segmentation to compute an F1 score. The F1 score is also given in the multiclass framework to compare the fine-tuned Koch filter with the deep learning model. In both cases, the F1 score is defined as the harmonic mean of precision and recall. Precision (resp. recall) is the average diagonal value of the column (resp. row) normalized confusion matrix.

4. RESULTS

The results are compiled in Table 1. They are given with the standard deviation over five trainings. The table shows that the U-Net architecture outperforms both the binary Koch filter and the fine-tuned filter. The best results were obtained at 400 m/px.

The figure 1 indicates the result of the segmentation on whole IW. To obtain this segmentation, the IW is divided in overlapping tiles of 20x20 km, segmented by the model, and fused to retrieve the whole IW. This process takes approximately 15 seconds per IW on a GTX 1050 Ti. The figure shows agreement between the NEXRAD measurement and the prediction.

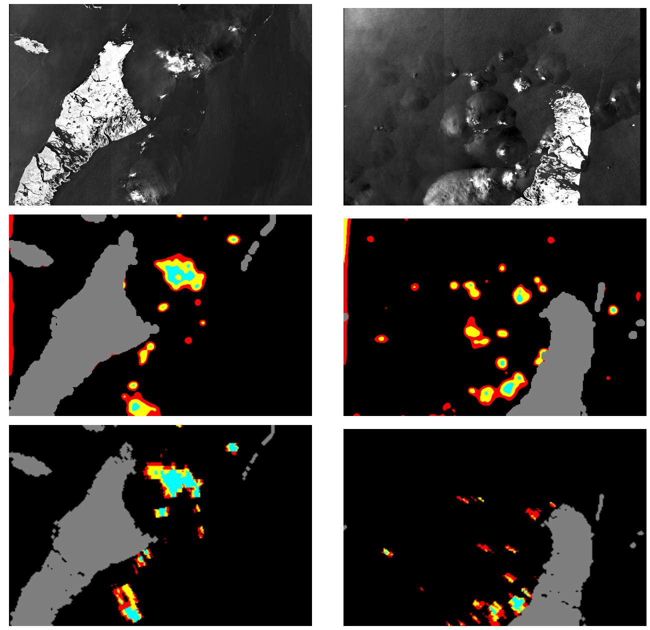

Fig. 1. From top to bottom: Sentinel-1 SAR observations, NEXRAD emulation result, and NEXRAD thresholded colocated reflectivity (c). The first and second columns were acquired on 2018-04-24 11:10:12 and 2018-08-19 23:19:09, respectively.

The colocalizations with GLM allow to evaluate the F1-score of the deep learning method with the binary lightning map, acting as a proxy for precipitation. Figure 2 indicates that the model performs better at higher incidence angle. The influence of incidence appears to be more important for heavier precipitation, especially for rain rates above 10 mm/h. The threshold at 1 mm/h, on the contrary, is not affected by the incidence angle.

Higher wind speeds also decrease the performance of the model. This is due to the lack of data at high wind intervals (81.4% of the pixels in the training subset are below 8 m/s, 98.4% are below 12 m/s). The decrease in performance can also be explained by the direct effect of wind. Since wind and rain increase the roughness of the sea surface, they have a negative impact on the SAR signature of the co-occurring phenomenon. It can be noted that the deep learning model performs particularly well at low wind speeds. On the contrary, the Koch binary filter performs less well when the wind speed is lower than 5 m/s.

5. CONCLUSIONS AND PERSPECTIVES

The deep learning model trained on the enhanced version of the Sentinel-1/NEXRAD dataset is able to segment entire Interferometric Wide Swath and retrieve precipitation in four

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 (co-pol.)	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 (co-pol.)	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 (co-pol.)	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%)

Table 1. Evaluation of the binary Koch's filter, the fine-tuned filters and the U-Net model on the patches of the test subset.

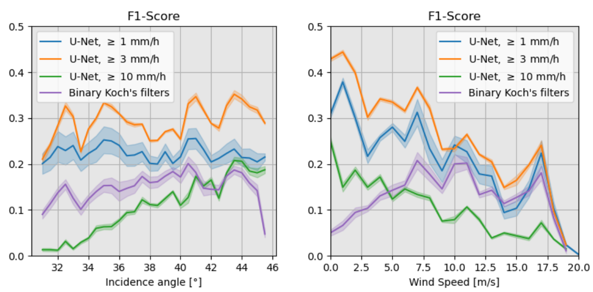

Fig. 2. F1-score of the rainfall segmentation for different incidence angles and wind speeds, using GLM as a proxy for the rainfall groundtruth.

different regimes: [0, 1[, [1, 3[, [3, 10[and [10, $+\infty$] mm/hr. It outperforms existing methods even when the state of the art is adapted to this multiclass segmentation problem and refined on this dataset. It outperforms existing methods even when the state of the art is adapted to the multiclass segmentation problem and fine-tuned on this dataset. The qualitative evaluation of IW segmentations highlights its importance in areas where weather radar is absent or too remote to provide accurate measurements.

In particular, the deep learning model performs best at low wind speeds, unlike the Koch filters. The influence of wind speed, and the weaker detection of rain events greater than 10 mm/h for low incidence angles, indicates that future work should focus on incorporating these parameters as priors in the network. Additional collocations, especially at higher wind speeds, could further improve segmentation.

6. THANKS

We thanks Nicolas Longépé (Φ-lab Explore Office, ESRIN, European Space Agency ESA) and Alexis Mouche (Laboratoire d'Oceanographie Physique et Spatiale, Ifremer) for the access to colocalizations between NEXRAD and Sentinel-1.

7. REFERENCES

[1] Yuan Zhao, Nicolas Longépé, Alexis Mouche, and Romain Husson, "Automated Rain Detection by Dual-Polarization Sentinel-1 Data," *Remote Sensing*, vol. 13, no. 16, 2021.

[2] Serge Soula, "Lightning and Precipitation," in *Lightning: Principles, Instruments and Applications*, pp. 447–463. Springer Netherlands.

[3] K. Lagouvardos, V. Kotroni, E. Defer, and O. Bousquet, "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*, vol. 134, pp. 45–55, Dec. 2013.

[4] W. Koch, "Directional analysis of SAR images aiming at wind direction," vol. 42, no. 4, pp. 702–710, Apr. 2004.

[5] O. Ronneberger, P.Fischer, and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," .

[6] Aurélien Colin, Ronan Fablet, Pierre Tandeo, Romain Husson, Charles Peureux, Nicolas Longépé, and Alexis Mouche, "Semantic Segmentation of Metoceanic Processes Using SAR Observations and Deep Learning," *Remote Sensing*, vol. 14, no. 4, pp. 851, Feb. 2022.

[7] Iris de Gélis, Aurélien Colin, and Nicolas Longépé, "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*, vol. 14, pp. 5831–5841, 2021.