IGARSS 2023 - 2023 IEEE International Geoscience and Remote Sensing Symposium | 979-8-3503-2010-7/23/$31.00 ©2023 IEEE | DOI: 10.1109/IGARSS52108.2023.10281499

PRESERVING EXTREMES IN SATELLITE QUANTITATIVE PRECIPITATION ESTIMATES: A MATTER OF SCALE

Clement Guilloteau[1] and Efi Foufoula-Georgiou[1,2]

[1]Department of Civil and Environmental Engineering, University of California, Irvine
[2]Department of Earth System Science, University of California, Irvine

ABSTRACT

Alternative metrics to the mean squared error (MSE) are assessed as optimization criteria for geophysical remote sensing applications. While the MSE promotes smooth solutions with compressed dynamical range, optimization criteria focusing on the preservation of the statistical properties of the target geospatial fields can be used to compensate this effect. These different optimization criteria are included as distance metrics within the objective function during the training of a deep learning algorithm for retrieval and mapping of precipitation rates from space-borne passive radiometric measurements. It is shown than imposing a constraint on the statistical distribution of the target variable at the "pixel" level only is insufficient, and that a multi-scale constraint (applied here through a discrete wavelet transform) is necessary to preserve the statistical properties of the target and the extremes across all scales.

Index Terms— Remote sensing, deep learning, convolutional neural network, optimization, precipitation, satellite, microwave

1. INTRODUCTION

In the global monitoring of precipitation, extreme high values, even if relatively infrequent, are of particular interest to the research community, because of their strong hydrological and climatological impact and their impact on human activities. Moreover, the uncertainty regarding their evolution under a changing climate reinforces the necessity to have accurate measurements and records of precipitation extremes all around the globe. Global satellite quantitative precipitation estimates (QPEs) have been produced and used by the research community for several decades and are now available at sub-hourly and sub-degree theoretical resolutions. However, regarding precipitation extremes, strong discrepancies still exist across the many available global datasets (which include observational products from remotely sensed information or gauges and reanalyzes) in particular at the finest spatial and temporal scales [1, 2]. The limited accuracy of satellite QPEs derives not only from the limited sampling and instrumental resolution provided by the multiple orbiting sensors, but also from the limited information content of the measured radiances, which makes the retrieval of precipitation from space a fundamentally underconstrained problem [3, 4, 5]. Numerous studies presenting evaluations of satellite QPEs against locally available high-accuracy ground measurements have revealed that, at sub-daily sub-degree resolutions, it is not uncommon, even the best-performing state-of-the art satellite QPEs, to have errors of order of magnitude similar to that of the precipitation signal itself [6, 7].

Satellite QPE products are generally designed and optimized to have the lowest possible mean squared error (MSE) at the pixel resolution. It is however well known that, when the residual errors cannot be reduced to a negligible quantity, MSE-optimal estimation algorithms tend to favor smooth solutions and to compress the dynamical range of the retrieved variable [8, 9]; the preservation of statistical extremes is generally "sacrificed" to the uncertainty. For satellite QPEs, the compressing effect of MSE minimization is scale-dependent: the variability at the finer spatial and temporal scales, which is generally associated larger uncertainty, is smoothed to a higher degree. This naturally impairs the preservation of extremes, particularly at fine scales. In the present study we assess the use of alternative metrics to the MSE as optimization criteria for satellite QPEs, focusing in particular on the statistical preservation of precipitation extremes, not only at the pixel scale, but also for spatially aggregated estimates, at all possible spatial scales.

In the recent years, the field of precipitation estimation from space has been massively migrating toward deep-learning algorithms, which in many cases have already demonstrated superior performance as compared to older techniques [10, 11, 12]. Beyond its already-demonstrated performance gain, an appealing aspect of the deep-learning framework is the fact that the criterion defining the optimal solution of a given deep neural network is not a-priori restricted by the data or by the network architecture, it is solely defined by the objective function chosen for the training.

In the present study, we use a mixed-scale dense convolutional neural network (MSDNet) [13] to perform the retrieval of instantaneous surface precipitation rates over oceans, from the brightness temperature measured by the passive GPM Microwave Imager (GMI) onboard the Global Precipitation Measurement (GPM) Core Observatory

satellite. The MSDNet is first trained using the MSE as the objective function to minimize. Then, keeping all other parameters identical, the training is run with different objective functions, which include constraints on the statistical distribution of precipitation rates, as well as multiscale constraints through the spatial wavelet power spectrum.

2. DISTANCE METRICS AND OBJECTIVE FUNCTION

2.1. Definitions

The MSE is the most commonly used optimization criterion in remote sensing, and is often used "by default" as the objective function to minimize in deep learning. Let us consider a target tensor **T** and its prediction **P**. The prediction MSE, designated here as the distance metric D_1, is the mean squared difference between **P** and **T**:

$$D_1(\mathbf{P},\mathbf{T}) = \frac{1}{n}\sum_{i=1}^{n}(p_i - t_i)^2 \quad (1)$$

where n is the number of elements in the N-dimensional tensors **P** and **T**, p_i is the i^{th} element of the tensor **P** and t_i is the i^{th} element of the tensor **T**. The MSE can be decomposed as:

$$D_1(\mathbf{P},\mathbf{T}) = (\mu_P - \mu_T)^2 + \sigma_T^2 + \sigma_P^2 - 2\,Cov(\boldsymbol{P},\boldsymbol{T}) \quad (2)$$

here μ_T, μ_P, σ_T^2 and σ_P^2 are respectively the mean and variance of **P** and **T** and $Cov(\mathbf{P},\mathbf{T})$ denotes the covariance between **P** and **T**. A condition for D_1 to be minimized is that $\sigma_P^2 = Cov(\boldsymbol{P},\boldsymbol{T}) \leq \sigma_T^2$ (otherwise, there exists a linear transformation of **P** with smaller MSE). Hence the "compressing" effect of MSE-optimal estimators, with the variance of the prediction being lower than the variance of the target.

To avoid this compressing effect, we define an alternative distance metric between **P** and **T** as:

$$D_2(\mathbf{P},\mathbf{T}) = (\mu_P - \mu_T)^2 + |\sigma_T^2 - \sigma_P^2| + \sigma_T^2(1 - CC(\boldsymbol{P},\boldsymbol{T})) \quad (3)$$

where $CC(\mathbf{P},\mathbf{T})$ denotes the linear Pearson correlation coefficient between **P** and **T.** This distance is minimized under the condition that $\sigma_T^2 = \sigma_P^2$.

To preserve the statistical properties of the target variable it is useful to consider distances between the respective statistical distributions of **P** and **T.** Because our focus is on high extremes values, to represent the statistical distribution of a non-negative variable X, rather than the probability density functions (PDF) or the cumulative distribution functions (CDF), we use the cumulative contribution function (CCF), which is a modified CDF, defined as:

$$\Phi_X(v) = \frac{1}{E[X]}\int_0^v u \times f_x(u)du \quad (4)$$

where f_x is the PDF of X and E[X] is the expected value of X. We measure distance between the CCFs of the target **T** and prediction **P** tensors through the standard continuous L1 distance and define:

$$D_3(\mathbf{P},\mathbf{T}) = \int_0^\infty |\Phi_P(v) - \Phi_T(v)|\,dv \quad (5)$$

Finally, to characterize the multiscale structure of the target variable, we utilize the discrete Haar wavelet transform. Let $S_X(\lambda,\omega)$ be the Haar wavelet power of the N-dimensional variable X at the scale λ and in the direction ω. We measure the distance between the discrete wavelet power spectra of the tensors **P** and **T** through the standard discrete L1 distance:

$$D_4(\mathbf{P},\mathbf{T}) = \sum_{i=1}^{J}\sum_{k=1}^{M}|S_P(\lambda_i,\omega_k) - S_T(\lambda_i,\omega_k)| \quad (6)$$

where λ_i and ω_k are respectively the discrete scales and directions of a N-dimensional discrete wavelet transform of depth J. In N dimensions, the number of directions of a discrete wavelet transform is $M = 2^N - 1$. In the present study, we apply the wavelet transform on 2-dimensional precipitation fields ($N = 2, M = 3$), and with $J = 5$.

2.2. Sensitivity of the distance metrics to different types of error

In this section we "corrupt" precipitation fields with different types of error, namely, multiplicative white noise, random spatial displacement (translation of the field) and Gaussian blurring. We then compare the corrupted fields to the original ones through the D_1, D_2, D_3 and D_4 metrics, to evaluate the response of each metric to the different types of error, and to the varying magnitude of the error. The original fields are radar-observed precipitation fields at 5 km resolution and are equated to the "truth".

The results of this evaluation of the error metrics with synthetic errors are presented in Figure 1. It can be seen that D_1 and D_2 have relatively similar responses to multiplicative white noise. The D_2 metric is however less sensitive to spatial displacement than D_1, but more sensitive to Gaussian blurring. The D_3 and D_4 metrics are by design insensitive to spatial displacement, they show a relatively low sensitivity to multiplicative noise but high sensitivity to Gaussian blurring. For many applications, small displacement errors are arguably tolerable, while the statistical biases introduced by

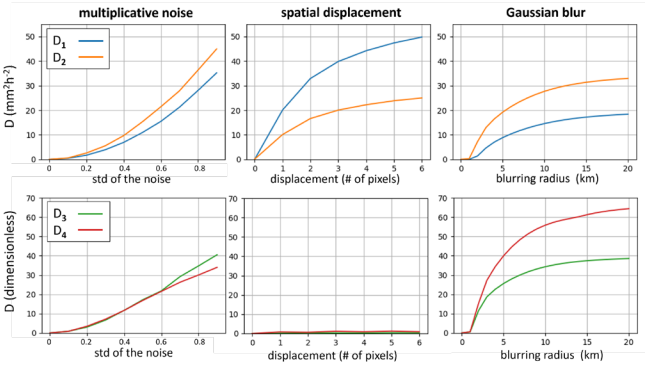

Fig. 1. Evolution of the computed distances the D_1, D_2, D_3 and D_4 between the original and "corrupted" precipitation fields as a function of the magnitude of the errors for different types of synthetic errors. The original precipitation fields comprise 1000 radar-observed precipitation scenes of dimension 245 km × 245 km at 5 km resolution. Distance metrics are averaged over all 1000 scenes.

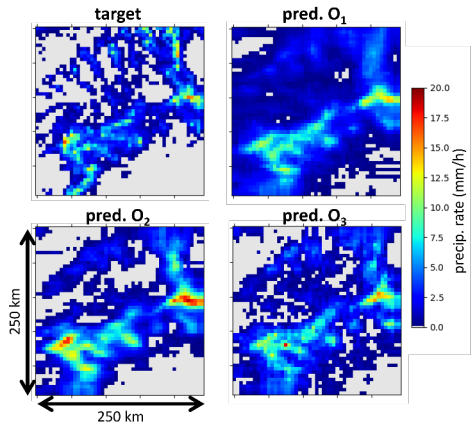

Fig. 2. One precipitation scene of the testing dataset as retrieved by the CORRA algorithm (target), and as predicted by the MSDNet from GMI brightness temperature only, trained with the different objective functions O_1, O_2 and O_3.

Gaussian blurring and other similar smoothing effects may be much more problematic.

From this, we can foresee that the D_1 metric (MSE) tends to promote smooth solutions when used as an optimization criterion. It is particularly intolerant to displacement errors, even of small magnitude. When the MSE is used as objective function in deep learning, the neural network is therefore likely to apply spatial smoothing (blurring) to mitigate the effect of high-frequency noise and/or spatial displacement errors. Including one or several other metrics (such as D_2, D_3 and D_4) in the objective function can however impede this effect.

3. DEEP LEARNING APPLICATION FOR SATELLITE PRECIPITATION ESTIMATION

3.1. Data and neural network architecture

The application chosen here to evaluate the different objective functions is the retrieval of instantaneous surface precipitation rates from the GMI instrument onboard the GPM satellite. GMI is a passive conical-scanning microwave imager, measuring brightness temperatures at the top of the atmosphere with vertical and horizontal polarizations at 7 frequencies between 10 and 183 GHz. These brightness temperatures are the product of the surface emission, emission and absorption by water vapor and liquid rain drops, and scattering by ice particles in the atmosphere. The middle of the 885-km wide GMI swath overlaps with the 245-km wide swath of the active Dual-frequency Precipitation Radar (DPR). Inside the DPR swath, estimates of the surface precipitation rate at 5 km horizontal resolution are provided by the NASA GPM Combined Radar-Radiometer Algorithm (CORRA [14]). We use these high-accuracy estimates to train the MSDNet to perform the retrieval of surface precipitation at 5 km resolution, from GMI only, over its full swath. 7000 unique precipitation scenes of dimension 245 km × 245 km randomly sampled over global oceans serve as training data and 1000 other scenes are used for the testing. The MSDNet comprises 70 densely connected convolution layers (see [15] for more details about the data and the MSDNet architecture).

3.2. Results

The MSDNet is first trained using the MSE as objective function to minimize:

$$O_1 = D_1(\boldsymbol{P}, \boldsymbol{T})$$

As expected, this leads to spatially smooth estimates, with compression of the extremes (figures 2 and 3). The training is then run with an objective function O_2, which combines the D_2 and D_3 metrics:

$$O_2 = D_2(\boldsymbol{P}, \boldsymbol{T}) + \gamma_3 D_3(\boldsymbol{P}, \boldsymbol{T})$$

Finally, a third training run is performed with an objective function O_3, which adds constraint on the wavelet power spectrum through the D_4 metric:

$$O_3 = D_2(\boldsymbol{P}, \boldsymbol{T}) + \gamma_3 D_3(\boldsymbol{P}, \boldsymbol{T}) + \gamma_4 D_4(\boldsymbol{P}, \boldsymbol{T})$$

γ_3 and γ_4 are weights used to balance the different terms of the objective functions, they are et to 3 and 1 respectively.

Only with this last training function the general texture of the CORRA precipitation fields, which includes sharp spatial gradients and numerous small-scale features, is well reproduced by the MSDNet (Figure 2). Figure 3 shows the probability of exceedance of precipitation rates between 0 and 100 mm/h computed over the testing dataset for the CORRA reference and for the 3 MSDNet predictions with the 3 different objective functions, at the 5 km pixel resolution, and at the 20 km aggregated resolution. The MSE-optimal prediction (O_1) dramatically underestimates the probability of exceeding rain rates higher than 10 mm/h, at both the 5 km and 20 km resolutions. The prediction O_2 better preserves the probability of exceedance at the 5 km resolution, but significantly overestimates the occurrence of precipitation rates above 5 mm/h at the 20 km resolution. The prediction O_3 reproduces quasi-perfectly the probability of exceedance of the CORRA reference fields at both the 5 km and 20 km resolutions.

While, as expected, the O_1 prediction has the lowest MSE when compared with the CORRA reference with a value of 12.8 mm^2h^2, the MSE values for the O_2 and O_3 predictions

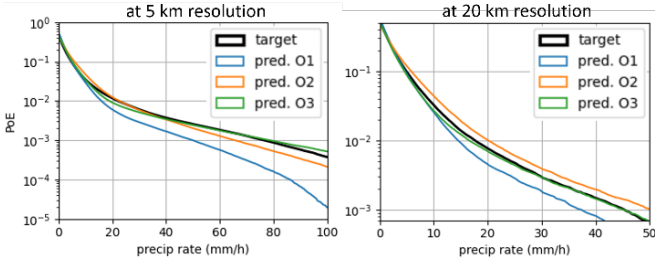

Fig. 3. Probability of exceedance (PoE) of the precipitation rates predicted by the three MSDNets trained with the O_1, O_2 and O_3 objective functions, along with the PoE for the CORRA algorithm (target), at 5 km resolution, and aggregated at 20 km resolution. All PoEs are computed over the 1000 scenes of the testing dataset.

are only slightly higher at 14.3 and 14.8 $\mathrm{mm^2h^2}$ respectively. In terms of coefficient of linear correlation with the CORRA reference, the three predictions are equivalent with values of 0.77, 0.77 and 0.76 for O_1, O_2 and O_3 respectively.

4. CONCLUSION

The present study demonstrates the limits of the MSE as an "all-purpose" optimization criterion (objective function) for remote sensing applications of deep learning, with the particular case of estimation of precipitation rates from passive satellite radiometric measurements. When the retrieval problem is underconstrained and the error cannot be reduced to a negligible value, the MSE-optimal solution tends to be smooth, with reduced variance, and compression of the extreme values as compared to the "truth". We show that, including a multiscale constraint in the objective function through the discrete spatial Haar wavelet power spectrum of the fields, combined with a constraint on the statistical distribution of precipitation rates, allows to statistically preserve the extremes across all scales, at the cost of a marginally degraded MSE.

5. REFERENCES

[1] Masunaga, H., Schröder, M., Furuzawa, F. A., Kummerow, C., Rustemeier, E., & Schneider, U. (2019). Inter-product biases in global precipitation extremes. *Environmental Research Letters*, 14(12), 125016.

[2] Derin, Y., Nikolopoulos, E., & Anagnostou, E. N. (2019). Estimating extreme precipitation using multiple satellite-based precipitation products. In *Extreme Hydroclimatic Events and Multivariate Hazards in a Changing Environment* (pp. 163-190). Elsevier.

[3] Kirstetter, P. E., Karbalaee, N., Hsu, K., & Hong, Y. (2018). Probabilistic precipitation rate estimates with space-based infrared sensors. *Quarterly Journal of the Royal Meteorological Society*, 144, 191-205.

[4] Guilloteau, C., & Foufoula-Georgiou, E. (2020). Beyond the pixel: Using patterns and multiscale spatial information to improve the retrieval of precipitation from spaceborne passive microwave imagers. *Journal of atmospheric and oceanic technology*, 37(9), 1571-1591.

[5] Turk, F. J., Ringerud, S. E., You, Y., Camplani, A., Casella, D., Panegrossi, G., ... & Peters-Lidard, C. (2021). Adapting passive microwave-based precipitation algorithms to variable microwave land surface emissivity to improve precipitation estimation from the GPM constellation. *Journal of Hydrometeorology*, 22(7), 1755-1781.

[6] Tang, G., Clark, M. P., Papalexiou, S. M., Ma, Z., & Hong, Y. (2020). Have satellite precipitation products improved over last two decades? A comprehensive comparison of GPM IMERG with nine satellite and reanalysis datasets. *Remote sensing of environment*, 240, 111697.

[7] Derin, Y., Anagnostou, E., Berne, A., Borga, M., Boudevillain, B., Buytaert, W., ... & Yilmaz, K. K. (2019). Evaluation of GPM-era global satellite precipitation products over multiple complex terrain regions. *Remote Sensing*, *11*(24), 2936.

[8] DeGroot, M. H. (2005). *Optimal statistical decisions*. John Wiley & Sons.

[9] Carroll, R. J., Ruppert, D., Stefanski, L. A., & Crainiceanu, C. M. (2006). *Measurement error in nonlinear models: a modern perspective*. Chapman and Hall/CRC.

[10] Sanò, P., Panegrossi, G., Casella, D., Di Paola, F., Milani, L., Mugnai, A., ... & Dietrich, S. (2015). The Passive microwave Neural network Precipitation Retrieval (PNPR) algorithm for AMSU/MHS observations: description and application to European case studies. *Atmospheric Measurement Techniques*, 8(2), 837-857.

[11] Gorooh, V. A., Asanjan, A. A., Nguyen, P., Hsu, K., & Sorooshian, S. (2022). Deep Neural Network High Spatiotemporal Resolution Precipitation Estimation (Deep-STEP) Using Passive Microwave and Infrared Data. *Journal of Hydrometeorology*, 23(4), 597-617.

[12] Rahimi, R., Vahedizadeh, S., & Ebtehaj, A. (2022). A Deep Learning Architecture for Passive Microwave Precipitation Retrievals using CloudSat and GPM Data. arXiv preprint arXiv:2212.02236.

[13] Pelt, D. M., & Sethian, J. A. (2018). A mixed-scale dense convolutional neural network for image analysis. *Proceedings of the National Academy of Sciences*, 115(2), 254-259.

[14] Grecu, M., Olson, W. S., Munchak, S. J., Ringerud, S., Liao, L., Haddad, Z., ... & McLaughlin, S. F. (2016). The GPM combined algorithm. *Journal of Atmospheric and Oceanic Technology*, 33(10), 2225-2245.

[15] Guilloteau, C., Le, P.V.V, & Foufoula-Georgiou, E. (2023), Constraining the multiscale structure of geophysical fields in machine-learning: the case of precipitation. *IEEE Geoscience and Remote Sensing Letters,* in press.