Article

Decorrelation of Satellite Precipitation Estimates in Space and Time

Francisco J. Tapiador 1,*, Cecilia Marcos 2, Andres Navarro 1, Alfonso Jiménez-Alcázar 1, Raul Moreno Galdón 1 and Julia Sanz 3

1 University of Castilla-La Mancha, Earth and Space Sciences Group (ESS), Institute of Environmental Sciences (ICAM), 45071 Toledo, Spain; Andres.Navarro@uclm.es (A.N.); Alfonso.JAlcazar@uclm.es (A.J.-A.); raulmorenogaldon@gmail.com (R.M.G.)
2 National Meteorology Agency (AEMET), 28071 Madrid, Spain; cmarcosm@aemet.es
3 Laboratory of Remote Sensing (LATUV), University of Valladolid, 47071 Valladolid, Spain; julia@latuv.uva.es
* Correspondence: Francisco.Tapiador@uclm.es; Tel.: +34-925-268-800 (ext. 5762)

Received: 25 April 2018; Accepted: 10 May 2018; Published: 14 May 2018

Abstract: Precise estimates of precipitation are required for many environmental tasks, including water resources management, improvement of numerical model outputs, nowcasting and evaluation of anthropogenic impacts on global climate. Nonetheless, the availability of such estimates is hindered by technical limitations. Rain gauge and ground radar measurements are limited to land, and the retrieval of quantitative precipitation estimates from satellite has several problems including the indirectness of infrared-based geostationary estimates, and the low orbit of those microwave instruments capable of providing a more precise measurement but suffering from poor temporal sampling. To overcome such problems, data fusion methods have been devised to take advantage of synergisms between available data, but these methods also present issues and limitations. Future improvements in satellite technology are likely to follow two strategies. One is to develop geostationary millimeter-submillimeter wave soundings, and the other is to deploy a constellation of improved polar microwave sensors. Here, we compare both strategies using a simulated precipitation field. Our results show that spatial correlation and RMSE would be little affected at the monthly scale in the constellation, but that the precise location of the maximum of precipitation could be compromised; depending on the application, this may be an issue.

Keywords: precipitation; geostationary microwave sensors; polar systems

1. Introduction

The importance of precise estimation of precipitation is apparent for assessing water availability for ecosystems and agriculture, and for other human activities. The usefulness of quality precipitation estimates is also evident for nowcasting and for data assimilation into numerical models. Thus, simulations have demonstrated that the assimilation of precipitation data leads to improved forecasting of a tropical cyclone in terms of its intensity and kinematical and precipitation structures [1,2]. The products after assimilating results in, for instance, significantly improved cyclone prediction, reflecting mostly in the cyclone’s track, the associated frontal structure and the associated precipitation along the front [3]. Rainfall monitoring is also important to assess possible anthropogenic impacts on global climate [4–6], to monitor hydrometeorological natural disasters, such as flood and flash flood events [7–11], and to improve precipitation estimates in Earth System Models (ESMs) [12–14].

Satellites are the only means to provide homogeneous global estimates of precipitation. Gauges are limited to land areas, leaving oceans with little or no direct measurements; the same applies to ground radars. On the contrary, satellites cover the whole planet and have the potential to provide frequent
estimates, not only of surface hydrometeors, but also of 3D precipitation profiles [15,16]. Unfortunately, measuring precipitation from space is a difficult task [17–19].

Satellite remote sensing of precipitation has evolved from the use of visible and/or infrared (IR) algorithms [20], to more direct strategies using passive microwave (PMW) radiometry [21] and orbiting radars, such as the Global Precipitation Measurement (GPM) mission [22] core observatory (GPM-Core), which is unique. PMW sensors measure the natural electromagnetic Earth emissions at microwave wavelengths, which are affected by rain drops in several ways. This allows for a more direct estimate of precipitation from space using radiative transfer modelling [23] (see References [24,25] for an update). Nonetheless, PMW sensors have poor temporal and spatial resolution due to their low orbits and the antenna diffraction limit at microwave wavelengths [26]. Infrared geostationary satellites, on the other hand, provide an indirect measure of the rainfall by establishing a relationship between cloud top temperature and surface precipitation [27], and have good temporal sampling and spatial resolution comparable with ground radars. The use of data fusion methods in rainfall estimation permits merging both datasets, aiming to create a high spatial and temporal resolution product [28].

Orbital radars are still scarce, and in spite of the effort in developing data fusion methods for IR and PMW sensors, the problem of precise remote sensing of precipitation is far to be solved in the near future. Routine comparisons between merged algorithms show large differences in performances between current methods depending on algorithm, season and location [29]. As space-borne sensors are the only means to homogeneously monitor land and ocean precipitation, the problem of reliably estimating global precipitation at appropriate spatial and temporal resolutions remains unsolved. Moreover, as precipitation estimates require a model error to be used when assimilated into NWP systems, algorithms and methods need to be physically based to be capable of estimating the covariance.

To further improve precipitation estimates from satellite, two major research directions are being followed. On the one hand, the GPM mission has increased the temporal resolution of the global estimates of precipitation by increasing the number of polar-orbiting microwave sensors, putting together a constellation of low-orbit satellites that reduce revisiting period. The contribution of all these satellites can generate improved (MW-based) and more frequent (more satellites involved) global precipitation estimates [30].

Another approach is the development of geostationary microwave sensors. While microwave antennas in the 6–90 GHz range would require antennas as large as 70 m for 10-km spatial resolution at 19 GHz [31], the exploitation of millimeter and submillimeter wavelengths would allow smaller (3-m) antennas, which are an affordable alternative for current engineering limits. Thus, a Geostationary Microwave Observatory (GEM) was proposed in 1998 [32] with a 2-m antenna yielding 15-km spatial resolution at nadir. In Europe, the Geostationary Observatory for Microwave Atmospheric Sounding (GOMAS) proposed a 3-m antenna aiming at 10-km spatial resolution [33]. Since then, other projects have been proposed and those are currently at different levels of maturity.

The rationale of millimeter and submillimeter estimation of precipitation differs from MW or PMW estimation. MW instruments such as the Precipitation Radar (PR) in TRMM measure the backscattered signal of a radar pulse, while PMW ones rely on the emission signature of cold hydrometeors over a warmer background (over the oceans), and on a variable relationship between the natural Earth PMW emission intercepted by hydrometeors and their emission signature over land. On the other hand, Geostationary Meteorological Satellite (GMS) estimate precipitation using absorption bands rather than windows to measure precipitation [34], which is an idea to measure how the atmospheric profiles are affected by the presence of hydrometeors. Preliminary studies show promising performances [31] in terms of not only precipitation identification but also raincell dynamics.

The aim of this paper is to analyze the theoretical differences in the spatial structure of precipitation between an ideal GMS and other alternatives such as a constellation of low-orbit PWM sensors or merged multi-satellite products. We follow a top-down approach by assuming an error-free GMS capable of providing precise measurements of precipitation. By degrading both the spatial resolution (to match the characteristics of hypothetical sensors spanning up to 250 km), and the revisiting periods
(up to 6.0 h), we can compare the loss of performances with a perfect GMS estimate. As there is not an actual GMS sensor to compare with, we have used a simulation to generate a realistic rainfall field at 15 min/0.05 degrees (about 5 km) resolution. The advantage over a pure synthetic stochastic precipitation field is that the simulation using an observed cloud cover can account for rare events affecting precipitation such as landing hurricanes or land-surfaces processes that would be too complex or cumbersome to be used in a pure stochastic model using, for instance, Poisson statistics.

2. Data

We have selected the precipitation in October 2005 in Spain as our empirical basis (Figure 1). Climatologically, October is the rainiest month in the country, so a statistically-significant number of precipitation events both from Atlantics fronts and from Mediterranean convective systems can be expected. October 2005 is also interesting because of two major meteorological events: the 11th of October hurricane (Vince) that landed in Southwestern Spain. This was the first time ever a hurricane headed towards the Iberian Peninsula and landed in mainland Spain [35]. Cordoba airport (37.85N, −4.85W, 170 km inland) received an unusual 84 mm of rain in 4 h, with a maximum intensity of 88.8 mm in 10 min. The short period in which the hurricane landfall generated high rainfall rates makes this event an anomaly in the normal climatology of the area.

Another event of interest is the major floods in Northeastern Spain (11–13 October) after the driest hydrological year on record. The driest year in Spain since 1947 was 2005, so monitoring high precipitation rates after such an event is relevant to erosion, urban drainage and agricultural analyses. Therefore, October 2005 presents a suitable benchmark for our study, as we have enough rain events in our comparison to be meaningful over the semi-arid environment of the Iberian Peninsula, and also we have the contribution of hurricane Vince and convective cells in the northeast. Satellite monitoring of
the high precipitation rates associated with Vince require high temporal sampling, whereas convective cells need a high spatial resolution.

3. Methods

Precipitation rates in the Auto-Estimator [27] were based on the cloud top temperature using the following empirical relationship:

\[ R = 1.1183 \cdot 10^{11} \cdot \exp(-3.6382 \cdot 10^{-2} \cdot T^{1.2}) \]

where \( R \) is the rainfall rate in mm h\(^{-1} \) and \( T \) is the cloud top brightness temperature in Kelvin (K). The algorithm was calibrated for radar rainfall estimates from the US operational network of 5 and 10 cm radar (WSR-57S, WSR-74C, WSR-88D), indicated to provide rainfall estimates for fast-moving deep convective systems during summertime.

To analyze the spatial variability of the estimates we calculated a semivariogram [36]. For each precipitation estimate \( r_i \), \( i = 1, \ldots, M \) located at a \( d \) distance from the others \( r_j \), \( j = 1, \ldots, M \) the empirical semivariogram is given by:

\[ \hat{\gamma}(d) \equiv \frac{1}{2N(d)} \sum_{(i,j) \in N(d)} |r_i - r_j|^2 \]

where \( N(d) \) denotes the set of estimates \((i,j)\) located at \( d \) distance in every direction (omnidirectional semivariogram). The semivariogram provides an estimate of the spatial variance, thus characterizing the spatial variability of the precipitation. Standard statistics, such as Pearson \( r^2 \), Root Mean Squared Error (RMSE) and bias, were used to compare many realizations. In addition, information entropy [37] was used to account for the informational content of the estimates. Entropy is defined, in this context, as:

\[ S \equiv - \sum_i p(R_i = r) \log[p(R_i = r)] \]

where \( p(R_i = r) \) indicates the probability of rainfall rate \( R_i \) being \( r \).

4. Results and Discussion

To build our simulated precipitation field, we used the Auto-Estimator. As an IR source, we used Meteosat-8 (formerly Meteosat Second Generation) data from the EUMETSAT archive. In spite of the limitations of using an IR-based method, the Auto-Estimator is well suited to generating our simulated precipitation, as it can provide pixel-based estimates at a high temporal sampling using only geostationary imagery, while other more powerful methods are less suited to this purpose and require additional data or ancillary information. It is worth mentioning that further enhancements of this techniques give way to a new product, called the Hydro-estimator. Such improvements include cloud-top geometry, available atmospheric moisture, stability parameters, radar, and local topography.

Several strategies have been devised to merge IR and PMW data, including neural networks [28,38,39], histogram matching [40] and multivariate probability matching techniques [41]. The aim of these methods is to reduce the temporal gap between rainfall estimates without sacrificing the quality of the more direct PMW estimate. Other methods to merge IR and PMW data include advection techniques such as morphing techniques that advect PMW estimates through the IR, as in Joyce et al. [42]. Their CMORPH method uses a correlation window algorithm to find the IR trajectories on an almost global scale, then advecting PMW estimates on those trajectories to fill the gaps between PMW successive overpasses. It has been shown that this morphing procedure can outperform other methods [43], though seasonal and spatial variations exist.
A comparison between the Auto-Estimator monthly estimates and morphing techniques such as the CMORPH [42] and the UCLM algorithm [44] shows an overall agreement (Figure 2) for the purposes of the research in this paper.

Figure 2. October 2005 precipitation estimates using a morphed microwave algorithm (top right, UCLM algorithm) and infrared-calibrated algorithms (bottom, CMORPH and AUTO-ESTIMATOR algorithms) compared to Climate Prediction Center (CPC) global land precipitation field (top left).

Daily comparisons (Figure 3) also show that the Auto-Estimator gives similar rainfall fields to those of other more complex algorithms in terms of detail and structure. Validation against the land-only, gauge-based Global Precipitation Climatology Project (GPCP) [45,46] (Figures 2 and 3, upper-left panel) also suggest that the Auto-Estimator is well-suited for use as a surrogate of a realistic global precipitation field.

Indeed, we are not claiming that the Auto-Estimator is a suitable instantaneous precipitation algorithm for the Iberian Peninsula. Our interest is not to describe the actual precipitation field, but to...
have a realistic precipitation field. The dependence of our results on the actual performance of the Auto-Estimator when compared with gauge data is of a second-order error. The reason for this is that we use the highest spatial and temporal resolution to build our simulated field.

Thus, the analysis presented here is independent of the precipitation algorithm used, providing that the algorithm can generate estimates at suitable spatial and temporal resolution, that is, at ~5 km/15 min. The assumption made here is that those estimates are the real precipitation and that the GMS system is capable of measuring precipitation with no error. Aggregated/subsampled estimates are compared in relative not absolute terms so the effects of varying resolutions can be investigated.

Simulated estimates were generated at 0.05°, with a 15-min resolution for October 2005, resulting in 2976 samples for analyses. These are taken both as real precipitation and as the GMS simulated estimates; that is, the estimates that a perfect GMS instrument would retrieve if on orbit and using a perfect precipitation retrieval algorithm.

Simplifying assumptions for polar instruments are in the form of perfect geometry, a wide swath to cover the Iberian Peninsula, no parallax error or instrumental biases, perfect retrieval and negligible beam-filling effects. All these simplifications can only benefit alternatives to the GMS estimate, as any polar system would suffer from these problems. Therefore, our results are to be considered as baseline estimates, meaning that the loss of performances we observe are the lowest limit over a perfect retrieval. The higher limit is the sum of all possible sources of error as mentioned above (imperfect geometry, narrow swath, etc.) up to the null hypothesis of the satellite retrieval perfectly matching actual precipitation.

To build our simulated estimates at several spatial and temporal resolutions we sampled the reference 0.05° rainrate field at 15-min intervals. The idea was to generate time-degraded estimates. These corresponded with a sensor having the same spatial resolution of the hypothetical GMS, but operating at different temporal resolution, therefore missing a variable number of continuous GMS estimates. Similarly, for the spatial resolution, we upsampled the 0.05° GMS field in 0.15° intervals up to 1.55° to simulate different satellite spatial resolutions. The resulting coarser resolution fields were then processed as before, assuming different temporal resolutions from 15 min to 6 h. This generated a host of different estimates corresponding to many possible satellite sampling/resolution characteristics. A sample of the results is plotted in Figure 4.
By hypothesis, a perfect GMS instrument measuring with no retrieval error would capture all the 2976 samples of the month at the original resolution yielding the totals depicted in the top/left plot of Figure 4. It is worth noting that the rainrate field was used instead of the field of radiances because it is assumed that the algorithm to derive the rainrates from radiances is perfect at each scale and resolution. That allows to isolate the effects of the changes in spatial and temporal resolution in the retrievals and therefore provides the best-case scenario in the event of degrading both variables.

As temporal resolution degrades, more and more samples are missed so the monthly precipitation estimate would deviate from the nominal truth as many relevant but short-living events will not contribute to the final sum. The spatial coarsening also deteriorates the perfect estimate, as small high precipitation events are blurred into the grid mean. It is clear from the figure that degrading the temporal sampling produces a patchy field, while the effect of spatial coarsening is smoothing the field.

Figure 5 explores all the combination between the estimates in terms of correlation, RMSE, bias and entropy. The correlation plot shows the importance of the spatial sampling to maintain the GMS performances. The correlation degrades faster in the spatial resolution direction, going below 0.90 $r^2$ at half a degree (for a 15-min temporal sampling), which is a working polar MW spatial resolution. In contrast, a revisiting period of 3 h decreased the correlation to 0.87 $r^2$. In environmental applications where precipitation estimates are integrated into models, multiplicative errors would noticeably worsen the final output in the latest case. In term of correlation, worsening the temporal resolution does not affect that much the estimates: Correlation isolines appear almost parallel to the x-axis.
The RMSE evolution parallels that of the correlation. The bias, however, shows the effects of the spatial aggregation using the pixel average. Negative biases are negligible, and correspond with temporal sampling at original spatial resolution, and the same applies to bias and RMSE. This smoothing in spatial resolution cannot compensate for the missing samples: the differences between the 0.05°/6.0 h estimate and the reference estimate are minute, whereas the 1.55°/6.0 h estimate cannot capture the precipitation features in the reference 0.05°/15 min estimate.

Perhaps more importantly for environmental applications, the entropy plot depicts the expected loose in informational content. The Boltzmann-Gibbs-Shannon entropy is a good estimate of how peaked the histogram of the precipitation estimates is [47]. Flat histograms correspond with a maximum entropy state, giving us little information on the structure of the precipitation, whereas a peaked histogram indicates null uncertainty (zero entropy) on that variable. Therefore, increased entropy would mean a smoother distribution, and less entropy that the reference state would indicate a change to a peaked histogram. Here, Figure 5 shows that temporal resolution is less important for conserving the informational content. Thus, a 0.05°/6 h estimate conserves most of the entropy of the original estimate. If the estimates are to be used to characterize the statistical properties of the precipitation over a region, isentropes mark how to preserve the informational content of the estimates when both spatial and temporal resolution are changed.

To better appreciate the variation in the spatial structure of the precipitation, it is necessary to use a spatial statistic measure such as the variogram. Figure 6 depicts the variograms for the 0.05° and 1.55° spatial resolution. This measurement provides complementary information on the statistical properties of the precipitation, showing the extent at which temporal sampling increases the variance of the field, as some short-living events are missed by the sensor. In both panels, temporal resolution decreases, as does the semivariance.

A practical difference between a GMS and other alternatives can be seen in Figure 7, which gathers the errors committed in the estimation of the maximum precipitation rates location for several spatial/temporal samplings combinations. Contrary to Figures 5 and 6, no clear pattern emerges. While it is true that high spatial resolution avoids large errors, it is also true that the 5 h 30 min sampling seems to work up to a 1.45° spatial resolution. This has implications for natural hazards monitoring, early warning services and ecological models using remotely-sensed data.
Figure 6. Semivariograms for 0.05° and 1.55° spatial resolution simulations, covering all the temporal samplings.

Figure 7. Errors in the estimated location of the maximum precipitation at the spatial and temporal resolutions explored in the paper.

Figure 8 complements Figure 7 in a more visual way. The figure depicts the errors in the location of the maximum monthly precipitation as temporal resolution degrades. The nominal truth locates the maxima over mainland Spain, in Catalonia. A 3-h sampling places the maximum over the Gulf of Leon, and a 6-h sampling over mainland France; the errors are noticeable.
Figure 8 complements Figure 7 in a more visual way. The figure depicts the errors in the location of the maximum monthly precipitation as temporal resolution degrades. The nominal truth (30-min, 0.1° resolution, upper/left) locates the maxima over mainland Spain, in Catalonia. A 3-h sampling places the maximum over the Gulf of Leon (upper/right), and a 6-h sampling over mainland France (bottom/left).

5. Conclusions

We have presented a simulation of an ideal GMS precipitation estimate and have compared it with degraded estimates made up by coarsening the spatial resolution at 0.15° intervals from 0.05–1.55° and by diminishing the temporal resolution from 15 min up to 6 h in 15-min intervals.

The aim of this experiment was to compare the relative merits of GMS estimates with alternatives such as a constellation of MW sensors on polar orbit, since no GMS sensor still exists. The results show that in the best-case scenario (absence of geometry errors and perfect instrumentation and retrieval algorithms) $r^2$ is expected to worsen up to 0.86, with a 4 mm/month bias and a 17 mm/month RMSE for a 3 h/0.45 deg sensor. The entropy of the alternatives would be slightly smaller, where major differences can be found in the location of monthly maxima: In our case study, a difference of 350 km was observed, albeit this value was deemed as highly dependent on the atmospheric situation.
The results show that spatial resolution is more important than temporal sampling in order to capture the climatology, whereas temporal resolution is critical to identify extreme events. While both the geostationary sounder and the polar constellation strategies can help a better understanding of precipitation and the water cycle, it depends on applications as to whether or not such values and uncertainties are acceptable.

Author Contributions: F.J.T. led the research, outlined the draft of the manuscript and made the amendments suggested by the referees. A.N., C.M., A.J.-A., R.M.G. and J.S. contributed to analysis, plotting and manuscript writing.

Acknowledgments: Funding from projects CGL2013-48367-P, CGL2016-80609-R (Ministerio de Economía y Competitividad, Ciencia e Innovación) is gratefully acknowledged. ANM acknowledges support from grant FPU 13/02798 for carrying out his PhD. We want to thank the anonymous referees for their valuable comments. Their comments have greatly improved the paper under review.

Conflicts of Interest: The authors declare no conflicts of interest.

References


10. Katsanos, D.; Retalis, A.; Tymvios, F.; Michaelides, S. Analysis of precipitation extremes based on satellite (CHIRPS) and in situ dataset over Cyprus. *Nat. Hazards* 2016, 83, 53–63. [CrossRef]


© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).