A Multiscale Evaluation of the Detection Capabilities of High-Resolution Satellite Precipitation Products in West Africa

CLÉMENT GUILLOTEAU AND RÉMY ROCA

Laboratoire d’Etudes en Géophysique et Océanographie Spatiales, Université de Toulouse III, CNRS, CNES, IRD, Toulouse, France

MARIELLE GOSSET

Géoscience Environnement Toulouse, Université de Toulouse III, CNRS, IRD, Toulouse, France

(Manuscript received 14 August 2015, in final form 9 May 2016)

ABSTRACT

Validation studies have assessed the accuracy of satellite-based precipitation estimates at coarse scale (1° and 1 day or coarser) in the tropics, but little is known about their ability to capture the finescale variability of precipitation. Rain detection masks derived from four multisatellite passive sensor products [Tropical Amount of Precipitation with an Estimate of Errors (TAPEER), PERSIANN-CCS, CMORPH, and GSMaP] are evaluated against ground radar data in Burkina Faso. The multiscale evaluation is performed down to 2.8 km and 15 min through discrete wavelet transform. The comparison of wavelet coefficients allows identification of the scales for which the precipitation fraction (fraction of space and time that is rainy) derived from satellite observations is consistent with the reference. The wavelet-based spectral analysis indicates that the energy distribution associated with the rain/no rain variability throughout spatial and temporal scales in satellite products agrees well with radar-based precipitation fields. The wavelet coefficients characterizing very finescale variations (finer than 40 km and 2 h) of satellite and ground radar masks are poorly correlated. Coarse spatial and temporal scales are essentially responsible for the agreement between satellite and radar masks. Consequently, the spectral energy of the difference between the two masks is concentrated in fine scales. Satellite-derived multiyear mean diurnal cycles of rain occurrence are in good agreement with gauge data in Benin and Niger. Spectral analysis and diurnal cycle computation are also performed in the West Africa region using the TRMM Precipitation Radar. The results at the regional scale are consistent with the results obtained over the ground radar and gauge sites.

1. Introduction

The number of precipitation-relevant observation platforms and algorithmic developments has increased in recent decades, yielding a large corpus of satellite quantitative precipitation estimation (QPE) products over the tropics. The range of applications of the products includes climatology (Biasutti and Yuter 2013; Roca et al. 2014), hydrological modeling (Bitew and Gebremichael 2011; Cassé et al. 2015), vegetation monitoring (Pierre et al. 2011), and infectious disease risk management (Guilloteau et al. 2014). Many validation studies of these products have been undertaken based on comparison with ground data (Sapiano and Arkin 2009; Gebremichael et al. 2014) at resolutions down to 0.25° and 3 h, and occasionally finer (Hong et al. 2007; Behrangi et al. 2012; Habib et al. 2012). Few of these validation studies have focused on West Africa (Nicholson et al. 2003; Roca et al. 2010; Gosset et al. 2013), where rain gauge networks are very sparse (Lorenz and Kunstmann 2012). These validation studies over Africa have been performed at a coarse spatiotemporal resolution (i.e., 1° and 1 day or coarser). Some products provide much more finely gridded data (down to 2.8-km instantaneous estimates) but remain unevaluated at their full resolution in West Africa.

Because of the intermittent nature of rain, QPE can be thought of as a double exercise: 1) identification of rainy areas and 2) estimation of rain intensity. The rain/no rain...
discrimination from passive satelliteborne sensors is far from trivial. Microwave and infrared brightness temperatures measured by the sensors cannot be unambiguously associated with a unique hydrometeor profile (Stephens and Kummerow 2007). Spatiotemporal variability of accumulated rain depth is partially driven by the rain/no rain variability. For a given period and over a given area, the cumulated rain depth is the product of precipitation fraction (i.e., fraction of space and time that is precipitating) and the mean rain intensity. The relative importance of each term in explaining rainfall variability depends on the considered resolution and the type of rainfall regime. Over the tropical continents, where a few hours of rain per year can produce most of the annual rain depth, the variability of the precipitation fraction is a key determinant (Morrissey et al. 1994; D’Amato and Lebel 1998; Kebe et al. 2005). In West Africa, most rainfall is provided by organized mesoscale convective systems (Houze 2004) of spatial expansion ranging from 102 to 106 km2 and propagating from east to west over thousands of kilometers. These systems may live up to a few days, but they produce rainfall for only a few hours over a given point. These physical processes give rise to a specific finescale signature in the rain fields.

Hossain and Huffman (2008) recommended to systematically analyze the dependence of error metrics to scale when assessing satellite rainfall data. To this end, Turk et al. (2009) and Sohn et al. (2010) aggregated satellite rain fields at various spatiotemporal resolutions to compare them with ground data. In this paper, the rain/no rain discrimination ability of a suite of high-resolution products derived from spaceborne passive sensors is evaluated in West Africa. Products considered are the Tropical Amount of Precipitation with an Estimate of Errors (TAPEER) intermediate data rain mask, Climate Prediction Center morphing technique (CMORPH), Global Satellite Mapping of Precipitation (GSMaP), and Precipitation Estimation from Remotely Sensed Information Using Artificial Neural Networks–Cloud Classification System (PERSIANN-CCS). The last three products are also evaluated as rain masks. A multiscale approach based on discrete wavelet decomposition is used to investigate the scale dependence of the masks’ performance. Satellite-derived rain masks are compared with ground radar–derived rain masks over Burkina Faso and with TRMM Precipitation Radar (PR)-derived rain masks over the whole West Africa. The wavelet coefficients resulting from the decomposition of the masks and characterizing the variations of these masks at various scales are compared. At each scale, the variances and the covariance of the coefficients are computed. This is equivalent to analyzing the masks through various bandpass filters and comparable to what is done in Turk et al. (2009) and Sohn et al. (2010), where the aggregation can be seen as a low-pass filtering.

The objective is to determine the relative contribution of each scale to the masks’ variance and covariance and to the variance of their difference. The scales for which the variance of the difference of the coefficients is greater than the variance of the radar-derived coefficients can be considered as noninformative. At these scales the information provided by the satellites degrades the estimation of the precipitation fraction. The subdaily variations of the precipitation fraction are partially driven by the diurnal cycle that is prominent in the tropics (Nesbitt and Zipser 2003; Roca et al. 2010). The satellite-derived mean diurnal cycle of rain occurrence is evaluated here against gauges in Benin and Niger.

This paper is organized as follows. Section 2 presents the datasets used in this study. In section 3, the method for multiscale qualification through wavelet transform is explained. Results of the decomposition, multiscale skill scores, and mean diurnal cycle are presented in section 4, with emphasis on both local and regional scales.

2. Data

a. High-resolution multisatellite rainfall products

The products evaluated are time series of high-resolution (i.e., finer than 0.1°) mapped estimates. Each estimate is an instantaneous snapshot of the surface rain rate at a given time. All the products have a sampling period shorter than 1 h. For all the products, microwave (MW) radiances measured from satellites forming the GPM constellation (Hou et al. 2014) and infrared (IR) images from geostationary satellites are used as primary or auxiliary sources of information. Spatial resolution, sampling period, and other main characteristics of the products are summarized in Table 1.

TAPEER is a 1°, 1-day quantitative rain estimation algorithm based on the Universally Adjusted Global Precipitation Index (UAGPI) technique (Xu et al. 1999; Chambon et al. 2013a). TAPEER was developed under the French–Indian Megha-Tropiques mission framework (Roca et al. 2015). TAPEER combines thermal infrared (10.8 µm) brightness temperatures (BTIR10.8) from geostationary imagers with passive microwave instantaneous rain estimates RMW (Chambon et al. 2012). For the West African region, infrared brightness temperatures are provided by the Spinning Enhanced Visible and Infrared Imager (SEVIRI; Schmetz et al. 2002) on board the Meteosat Second Generation geostationary platform every 15 min, with a 2.8-km resolution at nadir. Instantaneous
microwave rain rates (i.e., $R_{MW}$) are estimated using the Bayesian Rain Algorithm Including Neural Networks (BRAIN; Viltard et al. 2006; Kacimi et al. 2013). BRAIN is a Bayesian inversion algorithm associating hydrometeor profiles to microwave multispectral signatures measured by various radiometers on board the low-Earth-orbiting platforms forming the GPM constellation. The 2012–14 GPM constellation permits 6–10 overpasses per day over a fixed point at the surface (Chambon et al. 2013b).

As described in Chambon et al. (2012), the TAPEER degree-day estimation $R_{18,1d}$ relies on the rain detection from infrared brightness temperatures locally trained using $R_{MW}$ estimates:

1) For each $1^\circ \times 1^\circ \times 1$ day estimation volume, a $3^\circ \times 3^\circ \times 1$ day neighborhood or training volume is defined.
2) The $R_{MW}$ estimates and BTIR$_{10.8}$ pixels in the training volume are collocated.
3) Histograms of collocated data are computed and a rain/no rain threshold (IR TH) is defined such as $\Pr(R_{MW} > 0) > \Pr(BTIR_{10.8} < IR_{TH})$ over the training volume.
4) A binary indicator field, the rain mask $I_{TAPEER}$ is obtained by thresholding BTIR$_{10.8}$. The spatial resolution of $I_{TAPEER}$ is 2.8 km. Its sampling period is 15 min. The $1^\circ$, 1-day precipitation fraction estimate (PF$_{18,1d}$) is computed as the mean value of $I_{TAPEER}$ pixels over the 96 images in the $1^\circ \times 1^\circ \times 1$ day estimation volume.
5) Finally, $R_{18,1d} = PF_{18,1d} \times R_{cond,1d}$ is computed, where $R_{cond,1d}$ is the estimated average rain rate computed over rainy $R_{MW}$ pixels in the training volume:

$$R_{cond,1d} = \frac{1}{N_R} \sum_{X \in \{X^R\}} R_{MW}(X),$$  \hspace{1cm} (1)

where $\{X^R\}$ is the ensemble of rainy pixels in the MW training volume $\{X^R\} = \{X | R_{MW}(X) > 0\}$, $N_R$ is the number of elements of $\{X^R\}$, and $X = (x, y, t)^T$ is the time/position vector characterizing each pixel.

The final estimation $R_{18,1d}$ is not evaluated in this work. The intermediate product TAPEER rain mask is evaluated.

PERSIANN-CCS is produced by the Center for Hydrometeorology and Remote Sensing of the University of California, Irvine (Hong et al. 2004). The algorithm relies on the identification of rainy clouds using features such as cloud height, areal extent, and texture from IR images. PERSIANN-CCS algorithm also uses MW observations to statistically adjust IR-based estimates.
GSMaP is a JAXA product (Ushio et al. 2009). MW radiances are used to estimate rain rates through a radiative transfer model. To compensate for the sparse MW observations, rain fields are advected in space and time through a simple motion model called moving vector. Each new MW observation is assimilated using Kalman filtering. IR geostationary images are used for the computation of motion vectors. The near-real-time version of the product, which does not integrate gauges, is used.

CMORPH (Joyce et al. 2004) uses MW-based rainfall estimates as the primary source of information. From every MW observation, two “predictions” are made, forward and backward in time, using advection vectors computed from IR images. Estimated rain rates between two MW observations result from the merging of the two predictions (morphing). In this work, the gauge-free version 1.0 of CMORPH is used.

Integrated Multisatellite Retrievals for GPM (IMERG) recently developed by the U.S. Global Precipitation Mission team is a synthesis of PERSIANN-CCS and CMORPH algorithms (Huffman et al. 2015). The IMERG algorithm also inherited the correction procedure from gauge data of the TRMM Multisatellite Precipitation Analysis (TMPA) algorithm (Huffman et al. 2007). The PERSIANN-CCS algorithm is first run independently and PERSIANN-CCS rain fields are then optimally integrated into a CMORPH interpolation scheme using a Kalman filter. The data used here are the uncalibrated precipitation fields, which do not include correction from gauges. As this article is being written, IMERG has not been processed for 2012 and 2013 yet, so only the 2014 rainy season is considered. A limited assessment of IMERG detection capabilities focused on the energy spectrum is presented in section 4b.

b. Xport polarimetric radar data in Burkina Faso

Xport is an X-band (9.4 GHz) dual-polarization Doppler precipitation radar (Koffi et al. 2014). It operated in Ouagadougou, Burkina Faso (12.4°N, 1.5°W, Sahelian climate) during the 2012 rainy season (i.e., May–October) as a part of the Megha-Tropiques ground validation campaign. Its ground coverage is a 120-km-radius disk (Fig. 1). Its radial resolution is 200 m and its angular resolution is 1° (equivalent to 2-km width at the maximum range). The radar performs a complete scan of the surrounding area every 6 min. Rain rates used here are derived from differential phase shift between horizontal and vertical signals (Matrosov et al. 2002; Koffi et al. 2014). The intercomparison of several Xport rain fields derived from various independent variables shows a very good consistency in terms of rain detection at the resolutions considered in this study (i.e., 2.8 km and larger; Kacou 2014). Radar rain fields have also been validated against gauge data (Kacou 2014). Xport rain detection fields are considered as a reference dataset for the evaluation of TAPEER rain mask and the other satellite products presented above. Because of its high spatiotemporal resolution, the radar is an ideal tool to evaluate both temporal and spatial finescale rain variability. On the other hand, spatial scales larger than the radar coverage cannot be evaluated with a single radar, and conclusions obtained from local data cannot be extrapolated to a larger regional scale without additional information.

c. TRMM PR

The TRMM satellite carried a Ku-band radar from November 1997 to October 2014 to estimate the rain intensity from reflectivity profiles. The data used here are the TRMM 2A25, version 7, near surface rain (Iguchi et al. 2000) provided with a 5-km spatial resolution. The radar swath width at the ground is 247 km, not significantly greater than Xport’s range, but TRMM PR provides coverage of the whole West Africa region. More than 3000 orbit sections over West Africa during the 2012–14 rainy seasons are considered here (Fig. 1). The data are publicly available online (http://mirador.gsfc.nasa.gov/cgi-bin/mirador/presentNavigation.pl?tree=project&dataset=2A25%20%20Version%20007%20%20Radar%20Rainfall%20Rate%20and%20Profile%20%20PR%20&project=TRMM&dataGroup=Orbital&version=007).

d. AMMA-CATCH gauge networks in Benin and Niger

Two dense gauge networks setup in Benin (Sudanese climate) and in Niger (Sahelian climate) since the early 1990s have been operated as an element of the Couplage
de l’Atmosphère Tropicale et Cycle Hydrologique (CATCH) observatory of the African Monsoon Multidisciplinary Analysis (AMMA) program (Lebel et al. 2010). Both networks are made of 40–45 rain gauges covering a square area of 1° (Fig. 1). Gauges are automatic tipping buckets with a tip every 0.5 mm cumulated depth, leading to a delay of 30 min between two tips for a 1 mm h⁻¹ rain rate. Because of sparse spatial sampling and time-integrated measurements, rain gauges are weak in representing instantaneous rain fields at high resolution (Ciach 2003). Here, rain gauges are not used for a direct comparison with satellite instantaneous estimates, but only to infer statistical properties such as the multiyear mean diurnal cycle of rain occurrence.

e. Binary indicators generation

The comparison of two datasets with different resolutions is performed at the coarsest resolution: Xport data are aggregated to 2.8, 4.4, 8, and 11 km for the comparison with TAPEER rain mask, PERSIANN-CCS, CMORPH, and GSMaP. TAPEER rain mask is aggregated to 5-km resolution for the comparison with TRMM PR.

Thresholds for the rain intensity fields are defined to generate rain detection masks. When comparing several rain masks, the differences in rain detection sensitivity associated with various passive and active sensors and various detection methods must be accounted for. Figure 2 shows the probability of exceedance for rain rates between 0 and 5 mm h⁻¹ computed on Xport data and on collocated CMORPH, GSMaP, and PERSIANN-CCS data for the 2012 rainy season. The various datasets have different statistical distributions for low rain rates. All three products overestimate the occurrence of rain compared to Xport data. For the same period and area, the rate of rainy pixels in TAPEER rain mask is 12%. To generate the indicators $I_{\text{CMORPH}}$, $I_{\text{GSMaP}}$, and $I_{\text{PERSIANN}}$, a different threshold $R_{TH}$ is used for each product (2.6, 2.7, and 4.0 mm h⁻¹, respectively) so that the probability of exceeding $R_{TH}$ is also 12%. Thresholds are also defined for Xport fields at 2.8-, 4.4-, 8-, and 11-km resolution with $R_{TH}$ between 0.6 and 0.8 mm h⁻¹ to obtain the indicators $I_{\text{Xport2.8}}$, $I_{\text{Xport4.4}}$, $I_{\text{Xport8}}$, and $I_{\text{Xport11}}$. The resulting indicators’ probability distributions are therefore identical (i.e., 12% of 1 and 88% of 0). The aim of the present study is to analyze the products’ rain/no rain pattern variability and its scale dependence, rather than assessing detection biases between the datasets. The same method is applied to TRMM PR fields for the comparison with TAPEER rain mask on the West African scale. Rainy areas with intensity between 0 mm h⁻¹ and $R_{TH}$ are ignored because of the thresholds. These light rain areas represent between 50% and 75% of the total rainy area, but their contribution to the accumulated rain depth is only 2.2% for Xport data and 13% for each of CMORPH, GSMaP, and PERSIANN-CCS. Neglecting rainy areas below the threshold is suitable for West Africa, where the contribution of low rainfall intensities to the accumulated rain volume is marginal. In rain regimes dominated by low rain intensities, comparisons between radar and passive microwave such those as presented here would be partially driven by the relative sensitivities of passive sensor-based methods and radar measurements.

When aggregated to a resolution coarser than its original resolution, the indicator $I$ can be interpreted as a precipitation fraction $f \in [0, 1]$ (where the brackets indicate a closed interval). The precipitation fraction is a surface ratio and is therefore expressed in square meters per square meters. At 1° and 1-day resolution, the correlation between the precipitation fraction derived from $I_{\text{Xport8}}$ and Xport accumulated rain depth is 0.96. This means that in the area studied, the performance of the UAGPI method at TAPEER’s 1°, 1-day resolution mainly depends on the detection ability and the effect of rain intensity variability is secondary. This highlights the importance of the estimation of the precipitation fraction and supports the need to evaluate the scale at which this fraction can be estimated from passive spaceborne sensors.

3. Methodology

As stated in section 2e, the satellite detection fields $I_{\text{TAPEER}}$, $I_{\text{PERSIANN}}$, $I_{\text{CMORPH}}$, and $I_{\text{GSMaP}}$ are compared...
with radar detection fields \(I_{xport_4}, I_{xport_6}, I_{xport_8}, \) and \(I_{xport_{10}}\), respectively, for the 2012 rainy season. Detection fields are indicator fields such as \(I(x, y, t) \in \{0, 1\}\) (where the curly brackets indicate a finite ensemble or a list), where 1 is the value of the indicator when the measured rain is above the predetermined \(R_{TH}\). The various indicators have the same probability distribution of 0 and 1 by construction and signals have therefore the same total energy (see appendix A on Bernoulli distribution).

The pixel-to-pixel comparison of \(I_{CMORPH}, I_{GSMaP}, I_{PERSIANN}, \) and \(I_{TAPEER}\) with \(I_{xport}\) shows that false alarm rates (FARs) are 43%, 50%, 54%, and 49%, respectively. As a consequence, for all satellite masks, the mean-squared difference (MSD) with respect to the radar mask is of the same order of magnitude as the masks’ variance (see appendix A for the relation between FAR and MSD). Nash–Sutcliffe efficiency coefficients (see appendix A for the definition of these coefficients) of satellite masks with respect to the Xport mask are all negatives (between \(-0.23\) and 0). The negative values of the Nash–Sutcliffe efficiency coefficients (Krause et al. 2005) indicate that an unbiased “estimator” with a variance equal to zero (i.e., a constant climatic value) would perform better than the passive sensor satellite detection masks in terms of MSD with respect to the radar mask. Such pointwise evaluation, however, is of limited interest because it misses an important aspect of the precipitation process: its spatiotemporal organization. The estimated signal is auto-correlated in space and time and so is the error (Hossain and Anagnostou 2006; Teo and Grimes 2007). The “double penalty” phenomenon (Rossa et al. 2008), that is, spatial or temporal mismatch between satellite- and radar-observed patterns causing both false alarm and misdetection, affects pointwise verifications at high resolution, even if the two fields show good agreement at a coarser resolution.

Several methods are well suited to describe the spatiotemporal structure and the covariation of two variables. Among them are geostatistics (Grimes and Pardo-Igúzquiza 2010) and multiscale analysis through fractal theory (Lovejoy and Mandelbrot 1985) or wavelet transform (Kumar and Foufoula-Georgiou 1997; Venugopal et al. 2006). We chose the last one for its simplicity and its computational efficiency when dealing with massive data. Wavelet transform presents many similarities with the well-known Fourier transform (Flandrin 1998). In the Fourier domain, the wavelet decomposition is equivalent to a filter bank decomposition (Vetterli and Herley 1992). Wavelet spectra and Fourier spectra can be interpreted in a similar way. Wavelet transform is a lossless (reversible) operation. It does not rely on any approximation. It does not require any specific property such as data stationarity, which is questionable when dealing with precipitation data (Over and Gupta 1996). Wavelet coefficients are localized in space and time identically to the original data. The correlation between two series of wavelet coefficients can be interpreted in the space–time domain.

The use of the wavelet transform for the comparison of observed or modeled fields has been proposed by Briggs and Levine (1997). Kumar and Foufoula-Georgiou (1993, 1997), Venugopal and Foufoula-Georgiou (1996), Turner et al. (2004), Casati et al. (2004), and Johnson et al. (2014) showed the applicability of this method specifically for the analysis of rain fields. In this study, the Haar wavelet is used because it is well suited for representing binary fields (Kumar and Foufoula-Georgiou 1997; Domingues et al. 2005). The Haar scaling function is a simple averaging operator, and the wavelet coefficients are computed as finite differences (see appendix B). The Haar wavelet has been used by Casati et al. (2004) and Saux Picart et al. (2012) to analyze binary images, as it is done in the present study.

### a. The discrete wavelet transform

The discrete wavelet transform is a spectral decomposition. It decomposes a signal into a sum of sub-signals. Each subsignal characterizes the variations of the original signal at a specific scale

\[
I(X) \Rightarrow W_i(m, X),
\]

where \(X = (x, y, t)^T\) is the time/position vector, WT refers to wavelet transform, the integer \(m \in [0, M]\) is the scale index (where the brackets indicate a closed interval), \(M\) is the depth (or number of levels) of the decomposition, and \(W_i(m, X)\) are wavelet coefficients characterizing signal variations at a specific scale.

In the following, scales will be designated by the length scale \(L_m\) rather than the scale index \(m:\)

\[
L_m = 2^m L_0,
\]

where \(L_0\) is the original sampling period (spatial or temporal) of the data. The notation \(L_m^s\) is used for the spatial scale and \(L_m^t\) is used for the temporal scale.

For the largest scale \(m = M\), wavelet coefficients should be interpreted differently from the case when \(m < M\). The coefficient series associated with the index \(M\) can be seen as the residual of an uncompleted decomposition (because this decomposition has a finite number of levels). The term \(W_i(L_M, X)\) encodes large-scale variations of the signal, including the dc component (continuous component). The mean value of \(W_i(L_M, X)\) is equal to the mean value of \(I(X)\). As the Haar wavelet (see appendix B) is used here, \(W_i(L_M, X)\) is actually equal to the precipitation fraction at the \(L_M\)
resolution. For all other scales $m < M$, the mean value of $W_I(L_m, \mathbf{X})$ is zero.

Here, the wavelet decomposition is consecutively applied along spatial and temporal dimensions. The temporal decomposition has one dimension. For the spatial decomposition, a two-dimensional wavelet is used. A two-dimensional wavelet decomposition decomposes the signal into three components (vertical, horizontal, and diagonal) at each scale (Fig. 3). Each coefficient resulting from the spatial wavelet transform is a vector of $\mathbb{R}^3$ (except for the coefficients of index $M$):

$$I(\mathbf{X}) \xrightarrow{2D \text{ WT}} W^S_I(L^S_m, \mathbf{X}) \quad \text{with}$$

$$W^S_{HL}$$

$$W^S_I(L^S_m, \mathbf{X}) = W^S_{LH} \quad \text{for} \quad m < M.$$  \hspace{1cm} (4)

The coefficients associated with the index $M$ are scalars of $\mathbb{R}$.

The comparison of satellite and radar detection fields is performed as follows:

1) Each instantaneous rain mask image is first decomposed by spatial scale. The depth $M$ of the decomposition is limited by the Xport radar coverage, which is around 200 km. Variable $M$ also depends on the products’ original resolution. For the TAPEER rain mask, as $L_0^S = 2.8$ km, $M = 6$ and the resulting spatial scales are $L^S = \{2.8, 5.6, 11, 22, 44, 90 \text{ km}\}$ (where the curly brackets indicate a finite ensemble or a list). By decomposing one time series of detection images, $M + 1$ time series of wavelet coefficients are obtained.

2) Each of the $M + 1$ time series is then decomposed through a temporal wavelet transform of depth $N$.

3) Finally, $(M + 1) \times (N + 1)$ time series of wavelet coefficients result from the two successive decompositions. Each of these series represents the signal’s variation at a given spatiotemporal scale and is designated by spatial and temporal length scales $L^S_m$ and $L^T_n$:

$$I(\mathbf{X}) \xrightarrow{2D \text{ WT}} W^S_I(L^S_m, \mathbf{X}) \xrightarrow{1D \text{ WT}} W^S_I(L^S_m, L^T_n, \mathbf{X}).$$  \hspace{1cm} (5)

In section 4a, the decomposition is applied to ground radar and satellite masks, which are then compared scale by scale. For the comparison with TRMM PR data in section 4b, only spatial decomposition is performed because the lack of temporal continuity of TRMM PR observations forbids temporal decomposition. Haar wavelet is used for both spatial and temporal decompositions.

b. Discrete wavelet energy spectrum and cospectrum

From a frequency point of view, discrete wavelet transform is equivalent to a filter bank analysis (Vetterli and Herley 1992). For each value of $m$, the $W_I(L_m, \mathbf{X})$ coefficients series can be seen as the result of a filtering of the signal analyzed. When $m < M$, the filtering function is bandpass [centered on frequency $1/(2 \times L_m)$]. When $m = M$, the filtering function is low pass (with a cutoff frequency close to $1/L_M$). The analysis is based on the computation of discrete wavelet energy spectra and cospectra of signals. The discrete wavelet energy spectrum of $I$, $S_I(L^S_m, L^T_n)$, is obtained by computing the variance of wavelet coefficients $W_I(L^S_m, L^T_n, \mathbf{X})$ for each spatiotemporal scale $(L^S_m, L^T_n)$. As will be shown in Figs. 6 and 7 (described in greater detail below), the spectrum can be represented as a $(M + 1) \times (N + 1)$ matrix. It shows how the energy of the signal is distributed through scales. The cospectrum (CoS) of two signals $I_1$ and $I_2$, $\text{CoS}_{I_1, I_2}(L^S_m, L^T_n)$, is the covariance of wavelet coefficients at each scale (Fig. 6, bottom; described in greater detail below). Detailed equations for the computation of spectrum and cospectrum are given in appendix B. At each scale, the spectral energy of the difference of two signals is a function of the spectral energy of both signals and the cospectral energy:

$$S_{I_1 - I_2}(L^S_m, L^T_n) = S_{I_1}(L^S_m, L^T_n) + S_{I_2}(L^S_m, L^T_n) - 2 \text{CoS}_{I_1, I_2}(L^S_m, L^T_n),$$  \hspace{1cm} (6)

and the uncentered correlation (CC$_{mc}$; see appendix A) of wavelet coefficients is related to the spectral values:

$$\text{CC}_{mc}[W_{I_1}(L^S_m, L^T_n), W_{I_2}(L^S_m, L^T_n)] = \frac{\text{CoS}_{I_1, I_2}(L^S_m, L^T_n)}{\sqrt{S_{I_1}(L^S_m, L^T_n)S_{I_2}(L^S_m, L^T_n)}}.$$  \hspace{1cm} (7)

The spectral and cospectral analysis is used to quantify the contribution of each scale to the total energy of each signal, to the total cospectral energy, and to total energy of the difference between the two signals.

4. Results

a. Local comparison of satellite rain masks with Xport rain masks

1) COARSE SPATIAL SCALE: $m = M$

As stated in section 3, the largest spatial-scale coefficient $W^S_{T \text{TAPEER}}(L^S_0, \mathbf{X})$, resulting from the $M = 6$ levels spatial decomposition of the mask $I_{\text{TAPEER}}$, is equal to
FIG. 3. Spatial wavelet transform applied to $I_{TAPEER}$ with $M = 6$. (a) Values of $I_{TAPEER}$ at 2.8-km resolution and aggregated to 5.6, 11, 22, 45, and 90 km. (b)–(d) Wavelets coefficients $W^S(L^S_m, X)$ (horizontal, vertical, and diagonal) for (from top to bottom) $m$ between 0 and 5; $W^S(L^S_0, X)$ is equal to the precipitation fraction at 180 km × 180 km resolution ($PF_{180km}$).
the instantaneous precipitation fraction at 180-km resolution. Figure 4 shows the temporal evolution (with 15-min time steps) of $W^S_{TAPEER}(L^S_6, \mathbf{X})$ and $W^S_{TAPEER}(L^S_6, \mathbf{X})$ over a 180 km $\times$ 180 km square area centered on the Xport radar position during the 2012 rainy season. The uncentered correlation coefficient between the two time series is 0.83. The Nash–Sutcliffe efficiency coefficient of the $W^S_{TAPEER}(L^S_6, \mathbf{X})$ time series with respect to the $W^S_{TAPEER}(L^S_6, \mathbf{X})$ time series is 0.40. The energy of $W^S_{TAPEER}(L^S_6, \mathbf{X})$ is 36% of the $I_{Xport2.8}$ total energy. The energy of $W^S_{TAPEER}(L^S_6, \mathbf{X})$ is 44% of the $I_{TAPEER}$ total energy. The large spatial scale $L^S_6$ accounts for 67% of the total cospectral energy between the two masks $I_{TAPEER}$ and $I_{Xport2.8}$, and for only 15% of the energy of the difference $I_{TAPEER} - I_{Xport2.8}$. Table 2 sums up the results of the comparison of all four satellite masks with $I_{Xport}$, considering only the largest spatial scale coefficients $W^S_{I}(L^S_M, \mathbf{X})$. Results are similar for all products. The largest spatial scale accounts for less than one-half of each rain mask energy but explains more than two-thirds of the covariance between satellite and radar masks. The satellite masks and the radar masks appear to be consistent, essentially for spatial scales larger than 120 km. As a consequence [because of Eqs. (6) and (B11)], the energy of the difference between the two masks is concentrated in the fine scales: the relative weight of the spatial scale $L^S_M$ in the MSD is less than 20% for all satellite masks. For all satellite masks $I$, the uncentered correlation between $W^S_{I}(L^S_M, \mathbf{X})$ and $W^S_{I_{Xport2.8}}(L^S_6, \mathbf{X})$ is higher than 0.72, while the uncentered correlation between $I$ and $I_{Xport}$ is lower than 0.54. This demonstrates again the better agreement between Xport and satellite rain masks for large spatial scales than for the remaining part of the spectrum. When the masks are aggregated at a spatial resolution larger than 120 km, all resulting time series of satellite precipitation fraction show a positive Nash–Sutcliffe efficiency coefficient with respect to the radar precipitation fraction.

Figure 5 shows the spectra resulting from the wavelet temporal decompositions of $W^S_{TAPEER}(L^S_6, \mathbf{X})$ and $W^S_{I_{Xport2.8}}(L^S_6, \mathbf{X})$. At 180-km resolution, the 4- and 8-h temporal scales contribute the most to the energy of the signals and to their covariance. Little energy is associated with temporal scales finer than 2 h, showing that, in the region studied, the 180-km precipitation fraction varies slowly with time. As expected, at 180-km resolution, the variance of the difference $W^S_{I_{Xport2.8}}(L^S_6, \mathbf{X}) - W^S_{TAPEER}(L^S_6, \mathbf{X})$ (red curve) is small when compared to the variance of each signal, except for temporal scales finer than 2 h.

2) Fine spatial scales: $m < M$

Figure 5 shows the spectra of $I_{TAPEER}$ and $I_{Xport2.8}$, limited to the largest spatial scale $L^S_6$. Figures 6 (top and middle) show the spectra of $I_{TAPEER}$ and $I_{Xport2.8}$ for all spatial and temporal scales. The energy distributions of the two masks $I_{TAPEER}$ and $I_{Xport2.8}$ are very similar, except that $I_{Xport2.8}$ exhibits more variance in very small scales ($L^S < 10$ km, $L^T < 30$ min). This difference indicates that while the expected value (0.12) and variance (0.11) of the two datasets are equal by construction (see section 2f), $I_{Xport2.8}$ exhibits a more scattered structure with more rain/no rain transitions than $I_{TAPEER}$. The spectra for $I_{CMORPH}$, $I_{GSMaP}$, and $I_{PERSIANN}$ shown in Fig. 7 exhibit the same characteristics as the $I_{TAPEER}$ spectrum. The feature of the IR-based rain detection producing fewer and bigger objects than the radar-based detection has already been presented by Nesbitt et al. (2006). Note that the MW-based algorithms CMORPH and GSMaP share this feature with the IR-based algorithms. The coupling between spatial and temporal scales can be observed on the energy spectra in Fig. 6. When considering the 45-km spatial scale [Figs. 6 (top, middle); third row from above], the energy maximum is associated with the 2-h temporal scale. For the 11-km spatial scale [Figs. 6 (top, middle); fifth row], the energy is maximal around the 30-min temporal scale. This coupling is due to the relation between the size and...
TABLE 2. Performance of satellite rain masks against Xport rain mask, considering only the coarse-spatial-scale coefficients $W_{\text{coarse}}(L^s, X)$. Evaluation criteria are 1) $W^2(L^s, X)$; 2) $W^2(L^s, X)$ contribution to total energy of the satellite rain mask and Xport rain mask; 3) $W^2(L^s, X)$ contribution to the total MSD between satellite rain mask and Xport rain mask; 4) $C_{229}$ between $W^2(L^s, X)$ and $W^2(L^s, X)$; 5) $S_{\text{coarse}}(L^s, X)$ coefficient time series.

<table>
<thead>
<tr>
<th>Satellite rain mask</th>
<th>Original spatial resolution (i.e., $L^s$)</th>
<th>Depth (i.e., $L^d$)</th>
<th>Decomposition $L^d = 2^m L^s$</th>
<th>$S_{\text{coarse}}(L^s, X)$ contribution to total energy with $I_{\text{Xport}}$</th>
<th>$S_{\text{coarse}}(L^s, X)$ contribution to total energy with $I_{\text{TAPEER}}$</th>
<th>$W^2(L^s, X)$ vs $S_{\text{coarse}}(L^s, X)$</th>
<th>$W^2(L^s, X)$ contribution to covariance with $I_{\text{Xport}}$</th>
<th>$W^2(L^s, X)$ contribution to covariance with $I_{\text{TAPEER}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>TAPEER</td>
<td>2.8 km</td>
<td>6</td>
<td>180 km</td>
<td>67%</td>
<td>44%</td>
<td>44%</td>
<td>44%</td>
<td>44%</td>
</tr>
<tr>
<td>GSMaP</td>
<td>11 km</td>
<td>4</td>
<td>180 km</td>
<td>70%</td>
<td>40%</td>
<td>40%</td>
<td>40%</td>
<td>40%</td>
</tr>
<tr>
<td>CMORPH</td>
<td>8 km</td>
<td>4</td>
<td>128 km</td>
<td>72%</td>
<td>51%</td>
<td>51%</td>
<td>51%</td>
<td>51%</td>
</tr>
<tr>
<td>PERSIANN</td>
<td>4.4 km</td>
<td>5</td>
<td>140 km</td>
<td>69%</td>
<td>40%</td>
<td>40%</td>
<td>40%</td>
<td>40%</td>
</tr>
</tbody>
</table>

Fig. 5. Temporal spectra of coarse-spatial-scale components $W_{\text{TAPEER}}^2(L^s, X)$ and $W_{\text{Xport}}^2(L^s, X)$. Blue is $W_{\text{TAPEER}}^2(L^s, X)$ energy spectrum [$S_{\text{TAPEER}}(L^s, L^d)$]. Black is $W_{\text{Xport}}^2(L^s, X)$ energy spectrum [$S_{\text{Xport}}(L^s, L^d)$]. Green is $W_{\text{TAPEER}}^2(L^s, X)$ and $W_{\text{Xport}}^2(L^s, X)$ cospectrum [$C_{\text{TAPEER-Xport}}(L^s, L^d)$]. Red is energy spectrum of the difference $W_{\text{TAPEER}}^2(L^s, X) - W_{\text{Xport}}^2(L^s, X)$ [$S_{\text{TAPEER-Xport}}(L^s, L^d)$].

lifespan of rainy structures, varying from mesoscale systems down to individual convective cells (Mathon et al. 2002; Fiolleau and Roca 2013). This coupling is also associated with the displacement velocity of rainy systems, which is generally between 15 and 55 km h$^{-1}$ in the region studied (Depraetere et al. 2009). The area studied is not characterized by any significant climatic gradient in rain occurrence at the scale of the radar coverage. Long-term temporal averages tend to produce spatially homogeneous fields. As a consequence, the spectra show very little energy associated with small spatial scales and large temporal scales. The area studied is not suited for the evaluation of the ability of satellite estimates to capture finescale climatic patterns.

Figure 6 (bottom) shows the $I_{\text{TAPEER}}$ and $I_{\text{Xport}}$ power spectrum. The energy of the covariation is concentrated in few scales around (45 km, 2 h) and (45 km, 4 h). This concentration of energy reveals that, in the region studied, the TAPEER rain mask overall performance is essentially due to its skill at a few specific scales. The TAPEER IR-based detection is able to determine the edges of cloud systems but is very limited to map the rain/no rain variability inside of a cloud system. The scales of maximum covariation are in fact the scales corresponding to the dimensions of cloud systems. For each scale, Fig. 8 shows the correlation of $I_{\text{TAPEER}}$ and other satellite masks’ wavelet coefficients with $I_{\text{Xport}}$ wavelet coefficients. This correlation is lower than 0.5 for all products and for small scales such as $L^s < 45$ km and $L^d < 4$ h. At these specific scales, the variance of the difference $I_{\text{Xport}} - I_{\text{satellite}}$ is systematically higher than the variance of $I_{\text{Xport}}$. Wavelet coefficients of satellite masks at these scales behave as noise, meaning that
products do not represent the physical variability of actual rain/no rain patterns and that their variations are essentially random. If these scales were filtered out (i.e., wavelet coefficients set to zero), the resulting satellite fields would show lower MSD when compared with the radar rain mask. These small scales account for 24% of $I_{TAPEER}$ energy (and 38% of $I_{Xport2.8}$ energy). They account for only 7% of the covariance and up to 54% in the energy of the difference between the two masks. The performance of satellite products’ detection generally decreases as the temporal and spatial scales decrease, consistent with the well-known improvement of the performance by spatial or temporal averaging (Turk et al. 2009; Sohn et al. 2010; Hossain and Huffman 2008). Note that the scale/correlation dependency is not perfectly monotonic, as shown in Fig. 8. For instance, correlations between the satellite- and radar-based wavelet coefficients are higher for

**FIG. 6.** (top) Spatiotemporal energy spectrum $[\text{m}^2 \text{m}^{-2}]$ of $I_{TAPEER}$ $[S_{TAPEER}(L^m_m, L^x_x)]$. (middle) Spatiotemporal energy spectrum of $I_{Xport2.8}$ $[S_{Xport2.8}(L^m_m, L^x_x)]$. Each cell of the matrix shows the variance of wavelet coefficients $W(L^m_m, L^x_x, X)$. The sum of all cells is equal to the total energy of the mask. (bottom) Cospectrum of $I_{TAPEER}$ and $I_{Xport2.8}$ $[\text{CoS}(L^m_m, L^x_x)]$. Each cell shows covariance of wavelet coefficients $W_{TAPEER}(L^m_m, L^x_x, X)$ and $W_{Xport2.8}(L^m_m, L^x_x, X)$. The color scale is logarithmic. The curves on the top and on the right of each matrix are the temporal spectrum and spatial spectrum, respectively. They are respectively the sum of all lines of the matrix and the sum of all columns. Note that the first line of the matrices is redundant with Fig. 5.

**FIG. 7.** Spatiotemporal normalized energy spectra of (top) $I_{CMORPH}$, (middle) $I_{GSMAP}$, and (bottom) $I_{PERSIANN}$. The color scale is logarithmic. To be comparable, the spectra are normalized by dividing them by the total energy of each mask.
4- and 8-h scales than for scales between 16 and 32 h for TAPEER and GSMaP.

3) MULTIYEAR MEAN DIURNAL CYCLE OF RAIN OCCURRENCE

The spatiotemporal spectral analysis (Fig. 6) shows that 90% of $I_{TAPEER}$ and $I_{Xport2}$ energy is carried by temporal scales finer than 24 h. Part of this variability comes from the relation between the time of the day and the probability of precipitation, as revealed by the mean diurnal cycle. Because of the limited time span of the available Xport time series and because of temporal gaps in this radar data, rain gauges are used to compute mean diurnal cycles for the comparison with satellite data. Figure 9 shows 1-h mean diurnal cycles in terms of rain occurrence computed with TAPEER rain mask, CMORPH, GSMaP, and PERSIANN-CCS compared with AMMA-CATCH rain gauge networks (section 2e) in Niger and Benin. The mean diurnal cycles are computed for the July–September period for 2012–14. All products reproduce the main features of the mean diurnal cycle in both locations. TAPEER rain mask and GSMaP overestimate the cycle amplitude in Benin, particularly during the morning phase (0–12 h). The timing of maxima and minima are well reproduced.

This demonstrates that despite the low correlations shown in Fig. 8, finescale temporal variations are not purely random and the correlations are still significantly positive. Averaging several realizations of satellite detection series enables us to suppress random variability and reveals relevant finescale temporal variations. We may also expect temporal averaging to exhibit coherent small-scale spatial patterns. This cannot be verified with Xport radar data or AMMA-CATCH gauges, as covered regions do not exhibit a significant climatological gradient of rain occurrence, and long-term averages are spatially homogeneous.

b. Energy spectra at regional scale, comparison of $I_{TAPEER}$ with $I_{PR}$

The energy spectra are computed on a larger area. TAPEER rain mask is compared with TRMM PR to check if local results obtained with Xport data are still relevant on the West Africa regional scale. Figure 10
shows the energy spectra from spatial wavelet decomposition of the four satellite rain detection fields $I_{\text{TAPEER}}$, $I_{\text{CMORPH}}$, $I_{\text{PERSIANN}}$, and $I_{\text{GSMaP}}$, computed for May–October in 2012–14 over a region located between 5° and 25°N and between 10°W and 30°E in Sahelian and Sudanese climate zones (see Fig. 1). The spectra from Xport radar coverage in 2012 are superimposed as dashed lines in Fig. 10. For all datasets, the regional spectra are similar to those computed over Xport radar coverage in 2012. For the regional spectra, the relative weight of spatial scales finer than 64 km is systematically higher than for the local spectra. This small difference is related to the specific climate of the Guinean coast and Sudanese parts of West Africa, where mesoscale convective systems coexist with smaller rainy systems of a different nature (Fink et al. 2010; Depraetere et al. 2009). The regional spectrum was also computed on IMERG data for the 2014 rainy season. IMERG detection fields show significantly more variability than comparable products at the finest 11-km spatial scale. Compared to other products, IMERG spectrum thus appears to be more similar to the radar-derived spectrum. Whether this finescale variability is representative of actual rain intermittency or is random (as for other products tested in Burkina Faso) will be assessed once the IMERG product is made available for 2012.

TRMM PR is used to evaluate the spatial patterns in TAPEER rain mask, with a regional perspective. A threshold with $R_{\text{TH}} = 1.15 \text{ mm h}^{-1}$ (see section 2e) is applied to the PR rain intensity fields to obtain an indicator field $I_{\text{PR}}$. The two detection fields $I_{\text{TAPEER}}$ and $I_{\text{PR}}$ are collocated and compared. Because of the temporal discontinuity of TRMM PR observations, the multiscale analysis is performed along spatial dimensions only. More than 3000 overpasses during the 2012–14 rainy seasons (May–October) were processed. TAPEER rain mask is aggregated to TRMM PR 5-km spatial resolution before comparison. Spatial wavelet decomposition is applied with depth $M = 5$ (largest scale $L_5$ is 160 km). The large-spatial-scale coefficients $W_S^S(I_{\text{TAPEER}}(L_5, X))$, that is, precipitation fraction at 160-km resolution, account for 25% of $I_{\text{PR}}$ energy and 37% of $I_{\text{TAPEER}}$ energy. The uncentered correlation coefficient between the two series $W_S^S(I_{\text{TAPEER}}(L_5, X))$ and $W_S^S(I_{\text{PR}}(L_5, X))$ is 0.79 and the Nash–Sutcliffe efficiency coefficient of $W_S^S(I_{\text{TAPEER}}(L_5, X))$ against $W_S^S(I_{\text{PR}}(L_5, X))$ is 0.44, which is consistent with what was found locally in section 4a. Figure 11 shows the spatial wavelet energy spectra and cospectrum of $I_{\text{TAPEER}}$ and $I_{\text{PR}}$. The deficit of variance in TAPEER rain mask at 10- and 5-km scales (as seen in
5. Conclusions

The wavelet transform highlights the contribution of each spatial and temporal scale to signals’ variances and covariances. All four satellite-based, high-resolution rain masks considered exhibit energy spectra that are consistent with each other and with the ground radar spectrum. In particular, despite a small deficit of variance at scales $L^S < 10$ km and $L^T < 30$ min compared to the radar-based rain mask, they all show substantial variability in fine scales $L^S < 40$ km and $L^T < 2$ h. This is quite remarkable considering the expected low-pass filtering effect of merging procedures such as Kalman filter for GSMaP or morphing for CMORPH. Nevertheless, the comparison with ground radar data reveals that this variability is not representative of actual rain/no rain variability, but is essentially random. These fine scales roughly account for 40%–60% of the mean-squared error while their contribution in covariance with radar fields is negligible (<10%).

The multiscale method highlights the fact that standard pointwise scores such as correlation and mean quadratic difference are a combination of different scale-specific performances. A few specific temporal and spatial scales dominate the satellite product’s overall performance as revealed by the concentration of spectral energy.

All satellite-based products show a similar evolution of performance regarding spatial and temporal scales. The correlation between satellite detection fields and radar detection fields is essentially explained by their good agreement at coarse spatial and temporal scales. The relation between surface rain and BT$_{IR10.8}$ is known to be valid only from a statistical point of view (Kidd 2001). However, the direct use of BT$_{IR10.8}$ through a simple adaptive threshold in TAPEER shows skill in terms of rain detection at fine scales comparable with MW-based algorithms GSMaP and CMORPH or with PERSIANN-CCS.

All products considered have skills in reproducing the mean diurnal cycle of rain occurrence at 1-h temporal scale in Niger and Benin. This demonstrates that averaging along one dimension enables us to separate meaningful small-scale information from random variations. Similar analysis with radar data from a region with a strong local climatic gradient, such as a coastal or mountainous area, would be interesting to assess the effect of temporal averaging on finescale spatial patterns.

The results on the local scale were confirmed using TRMM PR data over the whole West African region and over a 3-yr period. The results presented focused on rainfall detection above a fixed rain intensity threshold. They are relevant for the question of quantitative estimation of rainfall amounts because precipitation fraction and rainfall cumulated depth are related. How much of the rainfall variability can be captured by the rain mask alone depends on the scales considered, on the precipitation regime, and on the value of the selected intensity threshold. The method presented could be applied to evaluate rain amounts (rather than a single rain mask) by repeating the multiscale binary indicators analysis with varying values of the intensity threshold (Casati et al. 2004).

In the region studied, the spectral contents of satellite rain masks are not optimal regarding the mean-squared difference between satellite and radar data. The
satellite-based rainfall estimate cannot be optimal in terms of mean-squared error and at the same time preserve statistical properties of actual rain. An adapted filtering of the finescale variations would reduce the mean-squared difference but would decrease the consistency between the radar and satellite power spectra. It would also degrade the mean diurnal cycle dynamics, which is currently satisfactory.
These results suggest that deterministic approaches to rain detection from passive sensors at high resolution are intrinsically limited by the nature of the precipitation process and by the inherent ambiguity of the relation between surface rain and cloud-related variables measured by spaceborne passive sensors (Stephens and Kummerow 2007). The development of a probabilistic approach for high-resolution rain detection deserves further investigation. Ensemble generation by wavelet transform associated with stochastic methods (Perica and Foufoula-Georgiou 1996), with rain mask as constraining data, will be investigated for this purpose.

**Acknowledgments.** The authors would like to acknowledge the support from CNRS, IRD, and Université de Toulouse III Paul Sabatier. Part of this work was supported by the CNES. The authors thank the AMMA-CATCH team for providing gauges data, and Dr. K. Hsu and Mr. D. Braithwaite from the University of California, Irvine, CHRIS for providing PERSIANN-Dr. K. Hsu and Mr. D. Braithwaite from the University AMMA-CATCH team for providing gauges data, and was supported by the CNES. The authors thank the correction work. Audine Laurian is thanked for her proofreading and available CMORPH, GSMaP, and IMERG data. Precipitation Mission team for respectively making measurement Mission science team, and the U.S. Global Prediction Center, the JAXA Precipitation Measurement Mission team for making mask data. The authors also thank the NOAA/Climate Lorant are thanked for providing the TAPEER rain mask data. Audine Laurian is thanked for providing the TAPEER rain mask data. The authors also thank the NOAA/Climate Prediction Center, the JAXA Precipitation Measurement Mission science team, and the U.S. Global Precipitation Mission team for respectively making available CMORPH, GSMaP, and IMERG data. Audine Laurian is thanked for her proofreading and correction work.

**APPENDIX A**

**Standard Metrics**

When comparing two variable fields $Y_1$ and $Y_2$, with $Y_1$ being the evaluated variable and $Y_2$ the reference variable, standard metrics mean-squared difference (MSD), correlation (CC), and uncentered correlation (CCnc) are defined as follows:

$$\text{MSD}(Y_1, Y_2) = \frac{1}{N_X} \sum_{X} |Y_1(X) - Y_2(X)|^2,$$  \hspace{1cm} (A1)

where $N_X$ is the number of elements of the datasets $Y_1$ and $Y_2$;

$$\text{CC}(Y_1, Y_2) = \frac{\sum (Y_1 - E[Y_1])(Y_2 - E[Y_2])}{\sqrt{(\sum (Y_1 - E[Y_1])^2)(\sum (Y_2 - E[Y_2])^2)}} = \frac{\text{cov}(Y_1, Y_2)}{\sqrt{\text{var}(Y_1) \text{var}(Y_2)}},$$  \hspace{1cm} (A2)

and

$$\text{CCnc}(Y_1, Y_2) = \frac{\sum Y_1Y_2}{\sqrt{\sum Y_1^2} \sqrt{\sum Y_2^2}},$$  \hspace{1cm} (A3)

where $\text{CCnc} \in [-1, 1]$ (where the brackets indicate a closed interval) and $E[X]$ is the standard mathematical notation for the expected value of the variable $X$. If $Y_1$ and $Y_2$ are positive, $\text{CCnc} > 0$.

Uncentered correlation is preferably used when both series contain many zeroes. The classical correlation would be artificially enhanced by the large number of correctly detected zeros $[Y_1(X) = Y_2(X) = 0]$. Well-detected zeros do not affect the value of $\text{CCnc}$; they can be removed from the series before the computation. Note that if both compared signals have a zero mean, CC and $\text{CCnc}$ are equivalent.

The Nash–Sutcliffe efficiency coefficient

$$\text{NS} = 1 - \frac{\text{MSD}}{\text{var}(Y_2)},$$  \hspace{1cm} (A4)

ranges from $-\infty$ to 1. A Nash–Sutcliffe efficiency coefficient of 1 indicates that $\forall X, Y_1(X) = Y_2(X)$. An NS of less than 0 indicates that an unbiased null variance estimator, that is, a constant value $Y_0(X) = E[Y_2]$, $\forall X$, would perform better than $Y_1$ in term of mean-squared error.

When considering two indicators $I_1$ and $I_2$, whose values are in $\{0, 1\}$ (where the curly brackets indicate a finite ensemble or a list), detection rate (DR) and false alarm rate (FAR) are specifically used to evaluate two-class detection:

$$\text{DR} = \frac{\sum I_1I_2}{\sum I_1} \quad \text{and}$$

$$\text{FAR} = 1 - \frac{\sum I_1I_2}{\sum I_1}. \hspace{1cm} (A6)$$

For Bernoulli distribution:

$$\text{var}(I) = E[I](1 - E[I]),$$  \hspace{1cm} (A7)

if $E[I_1] = E[I_2]$:

$$\text{CCnc}(I_1, I_2) = \text{DR},$$  \hspace{1cm} (A8)

$$\text{FAR} = 1 - \text{DR}, \quad \text{and}$$

$$\text{MSD}(I_1, I_2) = 2 \times E[I_1] \times \text{FAR}. \hspace{1cm} (A10)$$

**APPENDIX B**

**Discrete Wavelet Transform, Haar Wavelet**

For $m < M$, wavelet coefficients correspond to the convolution of analyzed signal with a wavelet function $\psi^m$ (Mallat 1999):
The associated scaling function is
\[ W(x, m < M) = \int_{-\infty}^{+\infty} f(\tau)\phi^M(\tau - x) \, d\tau. \]  
(B1)

The coefficient \( W(x, M) \) corresponding to the larger scale is obtained by convolution with a scaling function \( \phi^M \):
\[ W(x, M) = \int_{-\infty}^{+\infty} f(\tau)\phi^M(\tau - x) \, d\tau. \]  
(B2)

Each wavelet of the wavelet basis is obtained by dilation and translation of a mother wavelet \( \psi^0 \):
\[ \psi^m(x - \tau) = \frac{1}{\sqrt{2^m}} \psi^0 \left( \frac{x - \tau}{2^m} \right). \]  
(B3)

The discrete wavelet transform is decomposed on an orthogonal basis. If \( m_1 \neq m_2 \),
\[ \int_{-\infty}^{+\infty} \psi^{m_1}(\tau)\psi^{m_2}(\tau - x) \, d\tau = 0. \]  
(B4)

The simple Haar mother wavelet is
\[ \psi^0(x) = \begin{cases} +1 & \text{if } x \in [0, 0.5] \\ -1 & \text{if } x \in [0.5, 1] \\ 0 & \text{otherwise} \end{cases}. \]  
(B5)

The associated scaling function is
\[ \phi^M(x) = \begin{cases} \frac{1}{\sqrt{2^M}} & \text{if } x \in [0, 2^M] \\ 0 & \text{otherwise} \end{cases}. \]  
(B6)

The discrete wavelet transform can be generalized in two dimensions using three separate mother wavelets characterizing signal variations along three directions (vertical, horizontal, and diagonal).

The two-dimensional Haar mother wavelets and associated scaling function are, in matrix form:
\[ \psi^0_{LH} = \frac{1}{2} \begin{pmatrix} +1 & -1 \\ 1 & -1 \end{pmatrix}, \quad \psi^0_{HL} = \frac{1}{2} \begin{pmatrix} +1 & +1 \\ 1 & -1 \end{pmatrix}, \quad \psi^0_{HH} = \frac{1}{2} \begin{pmatrix} +1 & -1 \\ -1 & +1 \end{pmatrix}. \]
\[ \psi^0 = \frac{1}{2} \begin{pmatrix} +1 & +1 \\ +1 & +1 \end{pmatrix}, \quad \psi^0 = \frac{1}{2} \begin{pmatrix} -1 & -1 \\ 1 & 1 \end{pmatrix}, \quad \psi^0 = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}. \]

Wavelet transform conserves energy:
\[ \sum_{X} |I(X)|^2 = \sum_{m} \sum_{n} \sum_{X} |W_{I}(L^S_{m}, L^T_{n}, X)|^2. \]  
(B7)

From that, we can define the wavelet energy spectrum of \( I(X) \) as the variance of wavelet coefficients:
\[ S_I(L^S_{m}, L^T_{n}) = \sum_{X} |W_{I}(L^S_{m}, L^T_{n}, X)|^2. \]  
(B8)

The energy cospectrum of \( I_1 \) and \( I_2 \) is the covariance of wavelet coefficients:
\[ \text{CoS}_{I_1I_2}(L^S_{m}, L^T_{n}) = \sum_{X} W_{I_1}(L^S_{m}, L^T_{n}, X)W_{I_2}(L^S_{m}, L^T_{n}, X)^*. \]  
(B9)

The asterisk operator denotes conjugate transpose.
The mean-squared difference is also analyzed through the spectrum of the difference between \( I_1 \) and \( I_2 \) (Turner et al. 2004):
\[ \text{MSD}(I_1, I_2) = \frac{1}{N_X} \sum_{m} \sum_{n} S_{I_1-I_2}(L^S_{m}, L^T_{n}), \]  
(B11)

where \( N_X \) is the number of elements in the datasets \( I_1 \) and \( I_2 \).

REFERENCES


