A Spatiotemporal Correlation Technique to Improve Satellite Rainfall Accumulation

VELJKO PETKOVIC AND CHRISTIAN D. KUMMEROW

Colorado State University, Fort Collins, Colorado

(Manuscript received 30 August 2011, in final form 21 May 2012)

ABSTRACT

A spatiotemporal correlation technique has been developed to combine satellite rainfall measurements using the spatial and temporal correlation of the rainfall fields to overcome problems of limited and infrequent measurements while accounting for the measurement accuracies. The relationship between the temporal and spatial correlation of the rainfall field is exploited to provide information about rainfall beyond instantaneous measurements. The technique is developed using synthetic radar data. Nine months of Operational Program for the Exchange of Weather Radar (OPERA) data are used on grid sizes of 100, 248, and 500 km with pixel resolutions of 8, 12, and 24 km to simulate satellite fields of view and are then applied to the real satellite data over the Southwest to calculate 3-h rainfall accumulations. The results are compared with the simple averaging technique, which takes a simple mean of the measurements as a constant rainfall rate over the entire accumulation period. Using synthetic data, depending on the time separation of the measurements and their accuracy, a spatiotemporal correlation technique has shown the potential to yield improvements of up to 40% in absolute error and up to 25% in root-mean-square error when compared with the simple averaging technique. When applied to the real satellite data over the Southeast, the technique showed much less skill (general improvement of only 2%–6%).

1. Introduction

Demands on water resources have been increasing since the early days of humankind. Today, more than one billion people—almost one-fifth of the world’s population—still lack access to safe drinking water according to the United Nations World Water Development Report (UNESCO 2009). Precipitation variability dominates both water supply and the occurrence of droughts and floods, thus exerting great pressure on agriculture, as well as economic and social activities. This was and still is the basic motivation for precipitation measurements from early records in China dating as far back as 2000 BCE (Wang and Zhang 1988). Today’s motives have an added interest related to recent climate change science. An important part of the global hydrology cycle is the interaction between its small- and large-scale components, but neither climate nor weather forecast models are able to correctly capture the full spectrum of rainfall processes. Occurrence, intensity, and accumulation information is always in high demand. Of those parameters, rainfall accumulation is typically the most complex measurement. To provide an accurate accumulation, either dense frequent measurements or a very accurate model of the rainfall-producing system is required. However, even with the latest improvements in ground radar networks and satellite remote sensing technology, there is still not enough accuracy to satisfy all the needs of hydrology and meteorology (Hossain et al. 2004; Hossain and Anagnostou 2004).

Spatial scale requirements can be as small as 1 km × 1 km for purposes of nowcasting flash flood events (Huff 1993; Yates et al. 2000) and as large as 250 km × 250 km for global climate analysis and forecasting (Adler et al. 2003). At the same time, requirements on temporal sampling vary from 5 min (flash floods) to weeks or months (climatological means). There is significant demand for high accuracy at large scales. Rainfall trends, for example, require very accurate measurements over extended periods of time.

Currently, the best areal rainfall accumulations are provided by systems such as the Weather Surveillance Radar-1988 Doppler (WSR-88D) network in the United States (Crum et al. 1998) and the Automatic Meteorological...
Data Acquisition System (AMeDAS) in Japan (Makihara et al. 1996). Unfortunately, the majority of the world does not have such systems and is limited to satellite estimates. Using all currently available space-based measurements at even 3-h temporal and 0.25° spatial resolution remains a challenge (Joyce et al. 2004; Huffman et al. 2007). The primary objective of this study is to assess if spatial correlation patterns observed in rain systems can be exploited to improve 3-h accumulations from passive microwave sensors.

While some rainfall accumulation products are based on measurements provided by microwave, infrared (IR), and sometimes model results to compensate for poor temporal sampling of passive microwave sensors, this paper focuses on the improvements possible by these measurements alone. Temporal coverage of rainfall measurements from passive microwave sensors varies from about 30 min at high latitudes to about 6 h at low and midlatitudes. It is expected that 3-h time sampling (the goal set by the Global Precipitation Climatology Project) over all latitudes become available for the first time once the Global Precipitation Mission (GPM) is in place (Smith et al. 2007). To achieve this 3-h global coverage goal, this project plans to launch a core satellite to be used as a reference standard to a constellation of partner radiometers. Measurements from all available satellites, regardless of their capabilities, will be integrated to provide a unified global rainfall product. There is no doubt that a planned increase in the number of orbiting satellites itself will dramatically improve the quality of current rainfall products. However, in order to provide the best possible results, the measurements from all sensors need to be combined in such a manner that the most accurate ones are exploited to their maximum extent possible.

A number of techniques for combining the information from different instruments have been developed. Currently, techniques such as the Tropical Rainfall Measuring Mission (TRMM) Multisatellite Precipitation Analysis (TMPA; Huffman et al. 2007) and Climate Prediction Center morphing method (CMORPH; Joyce et al. 2004) are able to provide rainfall estimates based on the combination of microwave, radar, and IR data from space as well as gauge measurements from the ground (when available). These techniques use different methods to overcome poor spatiotemporal sampling, but ultimately use simple averages of individual observations, or advect the observed precipitation without considering the relative quality of the scene that is being advected. This study explores this topic, providing a relatively simple method for exploiting the temporal correlation of rain to improve rainfall accumulation of techniques mentioned above.

The goal of this study is to investigate the nature of the rainfall fields and, based on the findings, to develop a method that is capable of combining instantaneous rainfall estimates into a rainfall accumulation. Rather than taking a simple average of the measurements to represent the whole accumulation period, the method will recognize when one measurement should be trusted for longer periods. This solution can form a basis for improving accumulation techniques currently in use.

2. Data
   a. OPERA data

To develop the new technique, the data from the Operational Program for the Exchange of Weather Radar, stage 3 (OPERA 3; http://www.knmi.nl/opera), were employed. OPERA 3 radar data are available at 15-min time resolution over the majority of Europe (see Fig. 1) at a spatial resolution (here referred to as pixel resolution) of 4 km × 4 km. This makes OPERA data well suited for studying spatial and temporal properties of precipitating systems. This study uses OPERA 3 data from the beginning of September 2008 to the end of May 2009 over its entire available domain.

Figure 1 presents an arbitrarily chosen OPERA data scene. Scenes like this one are available every 15 min, except for intermittent problems affecting individual radars and for a period during January and February 2009 when no data are available due to archiving problems.

A total of approximately 100 radars are used, all of which operate at S, C, or X band in either Doppler or non-Doppler mode. About 5%–10% of radars are polarimetric radars. All data are assumed to be accurate in this study, with no qualitative differences between different types of radars. The data are used here only to develop and test the techniques so that uncertainties in the rain rates themselves would have little or no impact on the results. No time or space interpolation is made at times when missing data have occurred.

1) SUBDIVIDING DATA

All scenes are divided into a number of square, equally spaced subsounces, referred to as grids. In this study grid sizes of 100, 248, and 500 km are used (to roughly represent 1°, 2.5°, and 5° grids), which are sizes typically used in satellite studies. Throughout the study, only fully covered subsences containing at least one raining pixel are used as grids. Each grid contains a certain number of pixels. Beside the original 4-km radar resolution, three lower pixel resolutions are formed (8, 12, and 24 km) to correspond to available satellite field-of-view (FOV) sizes.
A grid size of 100 km provides approximately 1.5 million grids over the 6 months of available OPERA data. For 500-km grids, the number of grids is approximately 100,000, which still provides a relatively large database.

2) SIMULATED ERRORS

The methodology developed in this study is intended for satellite data application, which has a broad range of uncertainties. Therefore, the methodology must be tested on inaccurate data. Rather than using infrequent, limited, real satellite data where the accuracy is often difficult to estimate, inaccurate data are simulated by the OPERA data itself.

Random errors are introduced in a controlled manner. Normally distributed random perturbations are added to the prescribed pixel rain rates. This is done following

\[ r = r_{\text{OPERA}}(1 + e_n), \]

where \( r \) denotes the rainfall rate having some uncertainty \( e_n \); \( r_{\text{OPERA}} \) is the original rainfall value provided by OPERA; \( n \) is a normally distributed random number; and \( e_n \) is the assigned error. The product \( e_n \) represents a perturbation that has been added to the original data. If the value of the perturbation results in a negative rainfall rate, the rainfall rate is set to zero. The value of \( e_n \) is allowed to vary from 0.1 to 0.9 (in increments of 0.1), corresponding to errors from 10% to 90%.

b. Validation data

The methodology is tested using real data in a case study over the southeastern United States. Satellite data are provided from currently available sensors consisting of the TRMM Microwave Imager (TMI; Kummerow et al. 1998), the Advanced Microwave Scanning Radiometer for Earth Observing System (AMSR-E) on board the National Aeronautics and Space Administration (NASA) Aqua satellite (Lobl 2001; NSIDC 2009; Kawanishi et al. 2003), and the Special Sensor Microwave Imagers (SSM/Is; Weng, and Grody 1994) carried aboard the Defense Meteorological Satellite Program’s (DMSP) satellites F-13, F-14, and F-15. The data cover a 10-day time period from 20 to 30 April 2006. The area used in the case study is 10° × 5°, placed in the southeastern United States (31°–36°N, 83°–93°W).

On average, each satellite passes over the area twice a day leading to an average sampling interval of 140 min. Spatial coverage of the area is also relatively poor. Satellite swaths are 700–1400 km, which corresponds to about half of the area. On the other hand, data resolution is relatively high at roughly 25 km. Because of its low inclination, TRMM has slightly more overpasses than the other satellites.

“Truth” for the case study over the southeastern United States is provided by the Next-Generation Weather Radar (NEXRAD) stage IV data (Lin and Mitchell 2005; Baldwin and Mitchell 1998). The NEXRAD stage IV data represent the best estimate of rainfall accumulation over this area, considering the spatial and temporal coverage. Stage IV data are radar data mosaicked into a national product at the National Centers for Environmental Prediction (NCEP) from the regional hourly–6-hourly multisensor (radar and gauges) precipitation analyses produced by the 12 River
Forecast Centers over the contiguous United States (CONUS). Data are available hourly on a 4-km grid, from 1 Jan 2002 to the present.

3. Methodology

a. Conceptual idea

The fundamental hypothesis of this study is that the spatial correlation coefficient, or rainfall homogeneity that can be observed by a satellite snapshot, is closely related to the temporal variability. Linking these two allows one to use instantaneous rainfall measurements to propagate information through time. The idea of using the link between the spatial and temporal properties of the rainfall field comes from studies such as Bell et al. (1990) and Habib et al. (2001).

b. Spatial correlation

There are a variety of techniques used to calculate the spatial correlation or the spatial autocorrelation of one entity. The correct technique depends on the purpose of the calculation and on the definition of the entity. Mathematically, spatial autocorrelation quantifies the similarity between the distributions of the values of two vectors’ elements. In this case, the elements of the two vectors are the grid’s pixels; while their values are rainfall rates (see Fig. 2). Therefore, the correlation coefficient represents the “uniformity” of the rainfall field enclosed by the grid.

Figure 2 depicts part of the OPERA grid with its pixels. Both the left and the right panel in Fig. 2 represent the grid at the same time point (t₀). The labels are applied to provide a physical explanation below.

The spatial correlation coefficient, \( r_{xy} \), of a rainfall field is calculated using

\[
 r_{xy} = \frac{N \sum x_i y_i - \sum x_i \sum y_i}{\sqrt{N \sum x_i^2 - (\sum x_i)^2} \sqrt{N \sum y_i^2 - (\sum y_i)^2}}
 = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{(N-1) S_x S_y},
\]

where \( \mathbf{x} \) and \( \mathbf{y} \) are two vectors: \( \mathbf{x} = (x_1, x_2, x_3, \ldots, x_N) \) is the vector that contains the grid’s pixels as marked in the left panel in Fig. 2 and \( \mathbf{y} = (y_1, y_2, y_3, \ldots, y_N) \) is the vector that contains the corresponding neighboring pixels shifted by one pixel to the right (east), with \( N \) representing the number of pixels. Here, \( S_x \) and \( S_y \) are standard deviations of the vectors \( \mathbf{x} \) and \( \mathbf{y} \), respectively.

Equation (2) can be applied to any or all directions. The example in Fig. 2 illustrates only the shift in the easterly direction. If more than one direction is used, then for each additional direction the \( \mathbf{x} \) vector is extended by repeating its elements while vector \( \mathbf{y} \) will contain sets of corresponding neighboring pixels in that direction. Results presented in this study are based on the spatial correlation coefficient being calculated using the four cardinal directions: west, east, north, and south. Therefore, vectors \( \mathbf{X} \) and \( \mathbf{Y} \) in Eq. (2) are made each by concatenating four vectors: \( \mathbf{X} = [\mathbf{x}, \mathbf{x}, \mathbf{x}, \mathbf{x}] \), where each \( \mathbf{x} \) is as in Fig. 2; \( \mathbf{y} = (y_1, y_2, y_3, \ldots, y_N) \) is vector \( \mathbf{x} \) shifted in the east direction (as shown in Fig. 2); and \( \mathbf{y}_N, \mathbf{y}_W, \) and \( \mathbf{y}_S \) are vectors made by shifting vector \( \mathbf{x} \) in the north, west, and south directions, respectively. While the results of this study are not dependent

![Figure 2](image-url)
on the direction used in the spatial correlation calculations, the four directions are used to reduce the noise.

The spatial correlation coefficient (hereinafter referred to as “uniformity”) has high values for highly uniform raining systems. To illustrate this, spatial correlation coefficients are calculated for two different types of raining events using Eq. (2). The rain event shown in Fig. 3 is an example of a uniform (stratiform) type of rain, while the event shown in Fig. 4 is an example of a nonuniform (convective) type of rain. Spatial correlation coefficients of the initial states (labeled 00 h 00 min) of these two rain events have values of 0.682 for the uniform rain and 0.529 for nonuniform rain. Calculations for both grids were done for resolutions of 4 km and for the purpose of these examples arbitrarily chosen grids of 150 km × 150 km.

The mean spatial correlation coefficients for OPERA data for three different pixel sizes (8, 12, and 24 km) and a grid size of 100 km are shown in Table 1. Table 1 suggests that the lower the pixel resolution, the smaller the correlation coefficient will be. Since spatial correlation describes the uniformity of the rainfall field within a grid, the higher the coefficient, the more

![Fig. 3: Evolution of the uniform rainfall field over a 3-h time interval in 15-min snapshots. The event occurred on 2 May 2009 over the area marked by the red rectangle in Fig. 1.](image)
uniform the field. However, it would be incorrect to conclude that the spatial correlation coefficients in Table 1 suggest that an increase in the resolution increases the uniformity of the field. On the contrary, a decrease in pixel resolution always results in an increase in uniformity. Table 1 merely reflects the fact that spatial correlation is defined based on “pixels” rather than absolute distance. The technique presented in Fig. 2 implies that the lower the resolution, the larger the shifts of the pixels in space will be. This causes the correlation coefficient values in Table 1 to decrease with a decrease in pixel resolution, since, in general, rainfall rates between further points differ more than between close ones. Being aware of this, the more appropriate conclusion from Table 1 is that as the pixel resolution decreases, less of the uniformity is captured. The same conclusion is drawn when correlation length is calculated (see Table 1) instead of correlation coefficient (Petkovic 2010). However, this was not found to be more efficient in determining the temporal variability.

**Table 1.** OPERA data mean spatial correlation coefficient and corresponding correlation length for grid size of 100 km $\times$ 100 km.

<table>
<thead>
<tr>
<th>Pixel resolution (km)</th>
<th>24</th>
<th>12</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spatial correlation coeff</td>
<td>0.599</td>
<td>0.624</td>
<td>0.635</td>
</tr>
<tr>
<td>Spatial correlation length (km)</td>
<td>26.4</td>
<td>23.3</td>
<td>18.6</td>
</tr>
</tbody>
</table>
c. Temporal variability

The temporal variability $e_x$ is given by

$$e_x = \frac{\sum_{i=1}^{N} x_i - \sum_{i=1}^{N} y_i}{\sum_{i=1}^{N} x_i},$$

where $x_i$ and $y_i$ have the same meanings as in Eq. (2), except that now vector $y$ corresponds to the same pixels as vector $x$, just at some later time $t_1$. Since vectors $x$ and $y$ are made of instantaneous rainfall rate values, their totals are instantaneous totals. Therefore, the temporal variability of a raining grid corresponds to the change of rainfall during a time interval $\Delta t$. This change is presented as a ratio between the average instantaneous rain of the initial state and the difference in the average instantaneous rain between the initial and final states. The temporal variability also can be defined as a temporal sampling error that would be introduced into the estimate based on the assumption that the initial rainfall is constant in time.

The temporal variability must be defined relative to a time separation ($t_1 - t_0$). Each measurement time of the OPERA data defines one state of the raining event, so for $m$ consecutive measurements (states of the raining event) this would result in $m$ temporal variability states ($e_{x1}$, ..., $e_{xm}$), where $e_{x1}$ has a value of zero corresponding to temporal variability calculated for the measurement itself. An example is shown in Fig. 5. Temporal variabilities in Fig. 5 are calculated directly from the 15-min OPERA data for a single stratiform raining event (shown in Fig. 3), which occurred on 2 May 2009, and plotted (dashed line) as a function of time. Negative values of temporal variability, as per Eq. (2), indicate an increase in the rainfall relative to its initial value at 0 min.

Another example, for a single nonuniform raining event (the one presented in Fig. 4), is shown in Fig. 5 as a solid line. It is easy to notice that the temporal variability of the nonuniform event has larger amplitude and more rapid changes than the one that corresponds to the uniform event.

A number of factors influence the change in temporal variability. Aside from the rainfall type and storm velocity, temporal variability behavior depends largely on the size of the chosen grid.

d. Grid size and pixel resolution dependence

The following grid sizes and pixel resolutions are used to simulate typical grid sizes and satellite FOVs:

- 500 km × 500 km grid for pixel resolutions of 8, 12, and 24 km;
- 248 km × 248 km grid for pixel resolutions of 8, 12, and 24 km; and
- 100 km × 100 km grid for pixel resolutions of 8, 12, and 24 km.

On average, approximately 100,000 raining events are used in each calculation. Table 2 presents values of the spatial correlation coefficients for different grid sizes and pixel resolutions.

By looking at any of the rows in Table 2, one can see that the correlation coefficient increases with finer pixel resolution. As previously explained, this is brought about simply because grids are shifted by one pixel to compute the spatial correlation. The spatial shift of the 8-km pixel is therefore only $1/3$ of the shift used to compute the correlation when the pixel size is 24 km.

This analysis is not capable of recognizing finescale gradients (e.g., orography, coastal boundaries) in the field. Calculations (not shown here) are made for some smaller grid sizes (75 and 50 km). Results show the same trend of increasing correlation coefficients with decreasing grid size. This can be explained by the fact that
smaller grids have fewer raining systems. Rain systems over smaller areas increase the probability of uniform raining systems. Generally, rainfall rates change gradually within a raining system, implying that smaller portions of a system are generally more uniform. However, this is not the case when the 500-km grid is compared to the 248-km grid. This is caused by the definition of raining grids requiring at least one raining pixel. Very large grids have a higher chance of containing isolated rain pixels; this results in higher correlation values because of the large number of zeros in those grids. The minimum around 248-km grids appears to be related to synoptic scales. That is, there is usually some rain at these scales but not always at the smaller scales considered in this study.

Similar behavior can be seen in the temporal variability results. Table 3 shows the average variability for the same grid sizes and pixel resolutions as are shown in Table 2. The same data are used, choosing 180-min-long raining events to start at 0000, 0600, 1200, and 1800 UTC. A raining event is defined as the time period during which all of the grids are raining. Grids without continuous rain were not considered.

Table 3 shows only two time separations (30 and 45 min). Similar behavior is seen for both longer and shorter separations. For all grid sizes and pixel resolutions, increased time separation results in increased temporal variability. Simply, the longer it has been, the more the precipitation will change. By looking at any of the columns in Table 3, one can conclude that smaller grids experience larger changes in instantaneous accumulations. This is plausible since raining systems need less time to move across smaller grids. Table 3 also confirms that the temporal variability is not dependent on pixel resolution. This is expected since changes in the average rain within the grid are constant no matter what the pixel resolution is, as long as the grid captures the same area.

Similar results, as seen here for spatial correlations and temporal variability, can be found in the works of Bell et al. (1990) and Habib et al. (2001). In these two studies the correlation length is used instead of the temporal variability.

e. Temporal variability dependence on spatial correlation coefficient

Figure 6 shows the average absolute values of temporal variability and their standard deviations, as a function of spatial correlation coefficient for six separation times. The sample size used in this calculation is depicted as the green curve in Fig. 6. This result is based on the full OPERA data domain for a grid size of 248 km and pixel resolution of 12 km. Other combinations of grid sizes and pixel resolutions (not shown here) show the same trends. It is clear that as the separation time increases, the temporal variability increases as well. Also, for any given separation time, higher values of the spatial correlation coefficient correspond to lower temporal variability. Errors for separation time of 0 min are equal to zero since they correspond to the difference of the initial state of the grid to itself.

The relationship shown in Fig. 6 is the basis of the spatiotemporal correlation technique used to accumulate rainfall from multiple satellites.

4. Spatiotemporal correlation technique

The spatiotemporal correlation technique computes the total rainfall accumulation by weighting the available measurements based on both the measurements’ accuracies and the temporal variability of the measurements. The calculation of total rainfall accumulation is presented in two parts. The first part introduces the technique assuming that the sensors are 100% accurate, focusing only on temporal variability, and the second part deals with imperfect measurements and its addition to temporal variability. Finally, both parts are combined to provide an estimate of instantaneous grid average rainfall and its uncertainty.

Figure 7 shows an example that is used here to introduce the spatiotemporal correlation technique. In this example, the accumulation time period is chosen to be 3 h with a 15-min sampling rate. The accumulation interval is chosen to match the 3-h Global Precipitation Climatology Project (GPCP) interval, while the time sampling matches the OPERA sampling. Only two measurements are available within the accumulation interval although the technique can be modified to use a larger number of measurements. Each measurement provides an instantaneous grid average rainfall \( r \) and the spatial correlation coefficient of the grid \( \lambda \). To estimate the total rainfall accumulation during this period, the instantaneous grid average rainfall is calculated first.

### Table 3. Dependence of temporal variability (%) on grid size (increases with a decrease in the grid size) and pixel resolution (independent) for time separations of 30 and 45 min.

<table>
<thead>
<tr>
<th>Grid size (km)</th>
<th>Time separation (min)</th>
<th>Pixel resolution (km)</th>
</tr>
</thead>
<tbody>
<tr>
<td>500 × 500</td>
<td>30</td>
<td>24 24 24</td>
</tr>
<tr>
<td></td>
<td>45</td>
<td>34 34 34</td>
</tr>
<tr>
<td>248 × 248</td>
<td>30</td>
<td>74 74 74</td>
</tr>
<tr>
<td></td>
<td>45</td>
<td>109 109 109</td>
</tr>
<tr>
<td>100 × 100</td>
<td>30</td>
<td>252 252 252</td>
</tr>
<tr>
<td></td>
<td>45</td>
<td>287 287 287</td>
</tr>
</tbody>
</table>

Figure 6 shows the average absolute values of temporal variability and their standard deviations, as a function of spatial correlation coefficient for six separation times. The sample size used in this calculation is depicted as the green curve in Fig. 6. This result is based on the full OPERA data domain for a grid size of 248 km and pixel resolution of 12 km. Other combinations of grid sizes and pixel resolutions (not shown here) show the same trends. It is clear that as the separation time increases, the temporal variability increases as well. Also, for any given separation time, higher values of the spatial correlation coefficient correspond to lower temporal variability. Errors for separation time of 0 min are equal to zero since they correspond to the difference of the initial state of the grid to itself.

The relationship shown in Fig. 6 is the basis of the spatiotemporal correlation technique used to accumulate rainfall from multiple satellites.
a. Combining two perfect measurements

An instantaneous grid average rainfall for any specified time is estimated by a weighted mean (Rabinovich 2005) of the two measured instantaneous grid average rainfalls. The weighted mean \( r \) is given by

\[
r = \frac{(r_1 w_1 + r_2 w_2)}{w_1 + w_2},
\]

(4)

where \( r_1 \) and \( r_2 \) are measured instantaneous grid average rainfalls at times \( t_1 \) and \( t_2 \), while \( w_1 \) and \( w_2 \) are the corresponding weights. The weighting accounts for a temporal variability using weights given by

\[
w_i = \frac{1}{e_{ij}^2},
\]

(5)

where \( e_{ij} \) is the temporal variability, \( i \) is the measurement index corresponding to one of the two measurements used in this study, and \( j \) is the separation time between the measured and the estimated instantaneous grid average rainfall.

In Fig. 7, the green points mark the times \( t_1 = 45 \) min and \( t_2 = 150 \) min when two measurements of instantaneous grid average rainfall are made, while the red point marks the time of the instantaneous grid average rainfall estimate. In this scenario, subscript \( j \) in \( e_{1j} \) has the value of 45, since there is 45-min separation time between the measurement at \( t_1 \) and the estimate. Similarly, subscript \( j \) in \( e_{2j} \) has a value of 60. Temporal variability \( e \) which is used to form the weights, is precalculated based on large sets of data for all given resolutions and

![Fig. 6. Temporal variability (ratio) vs spatial correlation coefficient (grid size, 248 km; pixel resolution, 12 km). The ratio between rainfall amount at initial time \( t_0 \) (first panel) and consecutive 15-min time steps is plotted with its standard deviations (bars). The sample size used in calculation is shown in green.](image)

![Fig. 7. Accumulation period time line with two rainfall measurements marked in green (at 45 and 150 min) and the desired time of the estimate (at 90 min) is marked in red.](image)
grid sizes, as described in section 3, and stored in lookup tables available for use. Once the grid size and pixel resolution are chosen and the spatial correlation coefficient computed, a corresponding lookup table provides a temporal variability value for any given time separation.

b. Accounting for instrument errors

Using weights given by Eq. (5), the rain can be estimated at \( t = 90 \) min from rain and its spatial correlation at \( t_1 = 45 \) min and \( t_2 = 150 \) min. However, perfect measurements are rarely available. If two imperfect measurements are combined according to the previous method, the error of any of them can easily bias the estimate, making it worse than if a simple mean had been used. Therefore, the weights that are defined in Eq. (5) must be modified according to well-established techniques (e.g., Xie and Arkin 1996) to account for measurement errors. However, when the weights are adjusted for measurement errors, temporal variability must be preserved. Before the two adjustments are combined, the weight for the measurement error adjustment alone is presented:

\[
w_j = \frac{1}{a_i^2}, \tag{6}
\]

where \( a_i \) is the uncertainty of the instrument \( i \), given in a form of a positive decimal number. For example, 10% uncertainty is given by \( a = 0.1 \).

Rabinovich (2005) explains how to combine two independent errors of a measurement. Since measurement error and temporal variability are independent, and temporal variability, while not an error, can be expressed as one (see section 3c), Rabinovich (2005) suggests their combination to form the weights, as shown in Eq. (7):

\[
w_i = \frac{1}{e_{ij}^2 + a_i^2}, \tag{7}
\]

where the symbols are the same as previously used. Using the weighted mean with weights as defined in Eq. (7), one can estimate \( r \) at any time in Fig. 7. Summing the instantaneous grid average rainfalls within the accumulation period forms the estimate of the total rainfall accumulation.

Before proceeding to the results, it should be noted that if the measurement is not 100% accurate, the spatial correlation coefficient provided by such a measurement is also imperfect. The measurement error thus affects the temporal variability as well. Therefore, correlation coefficients are corrected to compensate for measurement inaccuracy. Rather than looking for the propagation of the \( \mathbf{X} \) and \( \mathbf{Y} \) vectors’ errors in Eq. (2), the Monte Carlo method is used to define the uncertainty of the spatial correlation coefficient, by comparing accurate fields to their uncertain duplicates. OPERA data, with simulated errors (as described in section 2), are used to produce those uncertain duplicates for 10 different uncertainty levels for all combinations of grid sizes and pixel resolutions (an example is given in Table 4). Results are given as corrections that have to be added to uncertain spatial correlation coefficients.

1) SIMULATED MEASUREMENTS

To depict the characteristics of the spatiotemporal correlation technique, two cases are examined using synthetic data. In the first case, the time separation (the time distance between measurements) of the two measurements is kept constant, while in the second case, the time separations are randomly distributed over the accumulation period. The first case provides more insight into the improvements that are dependent on data accuracy and temporal decorrelation, while the latter simulates more realistic scenarios, but is more difficult to interpret.

The results are presented through the improvements made in absolute and root-mean-square (RMS) errors
when the spatiotemporal correlation technique is compared to the simple average technique. Techniques are applied to approximately 15,000 raining events over the entire OPERA space–time domain and averaged results are shown.

(i) Fixed time separation measurements

Table 5 gives the absolute and the RMS error improvements for the grid size of 248 km with 12-km pixels for 15-min time steps and various temporal separations (centered within an accumulation interval). Measurement errors are set to be 0% for both of the instruments.

Both the absolute and the RMS error improvements behave in the same way. Since both instruments are assumed perfect (i.e., 0% measurement error), weighting is based only on the temporal variability. The maximum improvement occurs when the correlation times (related to spatial correlation) of the two measurements cover as much of the accumulation period as possible. Typically, the length of the correlation time (the time that the information from a single measurement is still useful) is approximately 60 min. If measurements are too close to each other, then their correlation times could overlap, potentially leaving portions of the accumulation period uncovered. This lowers the impact of improvement. The time separations of 0 and 60 min in Table 5 correspond to overlapping correlation time lengths. The ideal scenario is to have measurements that are about twice the mean temporal correlation length apart. A separation of 120 min is an example of this scenario, clearly having the greatest improvement among all of the time separations. It should be noted that the optimal result at 120-min separation is likely localized to Europe and its specific meteorological regimes.

In Tables 6 and 7 error improvements are given for combinations of three different sensors’ accuracies while the time separation is fixed at 120 min to provide simpler analysis.

It is clear from the tables that the best results are obtained if two measurements have the same accuracy. This is expected, since any inequality in the measurement error will lead to more weight being given to a single sensor. This is the case even if the instantaneous grid average rainfall estimate corresponds to the times beyond the correlation time of the more accurate measurement. The temporal variability corresponding to extended time separations of accurate measurements is larger than those corresponding to short time separations of less accurate measurements. It is important to note that this does not mean that a pair of two instruments of equal accuracies produces a better instantaneous grid average rainfall estimate than the same pair after the accuracy of one of them is increased. It simply means that, in this case, the improvement relative to the simple average technique is lower.

(ii) Random time separation measurements

Table 8 presents results when the measurements are randomly distributed over 3-h accumulation periods. The distribution is uniform, allowing any time separation within the 3-h interval to occur, including the case of no separation (i.e., measurements taken at the same time). Measurement errors are set equal for both instruments.

In general, Table 8 indicates that a decrease in sensor accuracy leads to a decrease in error improvements. It is worth mentioning that the spatiotemporal correlation technique applied to measurements with 90% uncertainty still provides an absolute error improvement of 15% and an RMS error improvement of more than 10%. This implies that the spatiotemporal correlation technique has skill even when applied to unreliable instrument measurements. Even more important is the fact that the technique can deal effectively with inaccurate measurements.

Table 5. Error improvement (%) for various time separations for grid size of 248 km and pixel resolution of 12 km. Best improvements, when compared with the simple averaging method, are seen for time separation of 120 min.

<table>
<thead>
<tr>
<th>Separation time (min)</th>
<th>Error improvement (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Absolute</td>
</tr>
<tr>
<td>0</td>
<td>0.00</td>
</tr>
<tr>
<td>60</td>
<td>33.03</td>
</tr>
<tr>
<td>120</td>
<td>47.54</td>
</tr>
<tr>
<td>180</td>
<td>36.00</td>
</tr>
</tbody>
</table>

Table 6. Absolute error improvement (%) for 120-min separation time for grid size of 248 km and pixel resolution of 12 km. Combinations of sensors of equal quality bring the highest improvement relative to the simple averaging method.

<table>
<thead>
<tr>
<th>Uncertainty of sensor 2 (%)</th>
<th>Uncertainty of sensor 1 (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>10</td>
</tr>
<tr>
<td>10</td>
<td>46.62</td>
</tr>
<tr>
<td>30</td>
<td>42.47</td>
</tr>
<tr>
<td>60</td>
<td>25.16</td>
</tr>
</tbody>
</table>

Table 7. RMS error improvement (%) for 120-min separation time, grid size of 248 km, and pixel resolution of 12 km. Combinations of sensors of equal quality bring the highest improvement relative to the simple averaging method.

<table>
<thead>
<tr>
<th>Uncertainty of sensor 2 (%)</th>
<th>Uncertainty of sensor 1 (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>10</td>
</tr>
<tr>
<td>10</td>
<td>45.24</td>
</tr>
<tr>
<td>30</td>
<td>41.75</td>
</tr>
<tr>
<td>60</td>
<td>20.32</td>
</tr>
</tbody>
</table>
In Tables 9 and 10, error improvements are given for three combinations of different sensors’ accuracies. These results mirror those seen in the fixed time separation scenario, except that the improvements are lower. This is expected, since time separation is not set to be at its optimum value (120 min in Table 6).

Finally, Table 11 shows how the absolute error improvement is dependent on the grid size and pixel resolution. Both of the measurement errors are equal to 30%.

No particular pattern is evident from Table 11. The same is the case for other combinations of measurement errors as well as for RMS error. Thus, the conclusion is that improvements do not show well-defined dependency on the grid size or the pixel resolution. However, they do remain reasonably high, keeping their values in the range of 15%–20% for absolute error improvements.

2) REAL DATA APPLICATION

Measurements from five satellite sensors are used to verify the results of the spatiotemporal correlation technique on real satellite data. Sensors’ biases are calculated by comparing satellite instantaneous grid average rainfall measurements to the ground-based instantaneous grid average rainfalls for the 10-day period. After removing the bias, instrument errors are determined by comparing satellite instantaneous pixel rainfall estimates to the NEXRAD stage IV estimates, for each sensor over the study area. Measurement errors are presented in Table 12. The satellite pixel resolution was approximately 0.25 km, currently among the highest resolutions available. The grid size is defined by 4 × 4 pixels resulting in grids of about 100 km in size. Across the 10-day period, two or more instruments have detected 42 raining events having a time separation shorter than 3 h.

The true total rainfall accumulation is estimated using the NEXRAD stage IV data. As described in section 2, these data are hourly accumulation estimates. To calculate the improvement of the spatiotemporal correlation technique over simple averages, 15-min estimates made from the spatiotemporal correlation technique are aggregated to 1-h periods and are compared with NEXRAD’s accumulation. The results and basic information regarding the case study are given in Table 12. The total rainfall accumulations shown in Table 12 correspond only to the area where rain has been detected by satellites. Because of a number of problems related to real data usage, such as mismatching of a 3-h satellite accumulation interval with hourly sampling of NEXRAD stage IV data, the results are not as good as those predicted with synthetic data (shown in brackets).

To show that data application results support our synthetic data findings, details of two selected rain events are shown in Tables 13 and 14. Single snapshots of events are shown in Figs. 8 and 9. Two events are used to present two different raining types. One of the events has fields with low spatial correlation coefficients, indicating convective rain (Table 13 and Fig. 8), while the other event contains high correlation coefficients, implying more stratiform-like raining fields (Table 14 and Fig. 9).

A time separation happened to be at the minimum in the convective type event. Table 13 shows small Δt. This 3-h accumulation period started on 20 April 2006 at 2145 UTC. Two satellite instantaneous grid average rainfall measurements, made by SSM/I13 and SSM/I14 sensors, at approximately 2330 and 2345 UTC, respectively, are 6.060 and 6.236 mm h⁻¹. The times of measurements are marked as 105 and 120 min (relative to the beginning of the accumulation interval) in Table 13.

### Table 8. Error improvement (%) as a function of measurement error for random separation times for grid size of 248 km and pixel resolution of 12 km. Error improvements decrease as measurements become less precise.

<table>
<thead>
<tr>
<th>Measurement error (%)</th>
<th>0</th>
<th>10</th>
<th>20</th>
<th>30</th>
<th>40</th>
<th>50</th>
<th>60</th>
<th>70</th>
<th>80</th>
<th>90</th>
</tr>
</thead>
<tbody>
<tr>
<td>Absolute</td>
<td>22.94</td>
<td>21.74</td>
<td>19.81</td>
<td>18.49</td>
<td>17.13</td>
<td>16.64</td>
<td>15.93</td>
<td>15.86</td>
<td>15.14</td>
<td>15.41</td>
</tr>
</tbody>
</table>

### Table 9. Absolute error improvement (%) for random separation time, grid size of 248 km, and pixel resolution of 12 km.

<table>
<thead>
<tr>
<th>Uncertainty of sensor 2 (%)</th>
<th>Uncertainty of sensor 1 (%)</th>
<th>10</th>
<th>30</th>
<th>60</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>21.74</td>
<td>18.16</td>
<td>9.57</td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>18.16</td>
<td>18.49</td>
<td>12.94</td>
<td></td>
</tr>
<tr>
<td>60</td>
<td>9.57</td>
<td>12.94</td>
<td>15.93</td>
<td></td>
</tr>
</tbody>
</table>

### Table 10. RMS error improvement (%) for random separation time, grid size of 248 km, and pixel resolution of 12 km.

<table>
<thead>
<tr>
<th>Uncertainty of sensor 2 (%)</th>
<th>Uncertainty of sensor 1 (%)</th>
<th>10</th>
<th>30</th>
<th>60</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>15.15</td>
<td>13.42</td>
<td>5.93</td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>13.42</td>
<td>14.16</td>
<td>9.59</td>
<td></td>
</tr>
<tr>
<td>60</td>
<td>5.93</td>
<td>9.59</td>
<td>13.33</td>
<td></td>
</tr>
</tbody>
</table>
A spatial correlation coefficient of the field measured by SSMI13 was 0.181, while 15 min later and with the same field, the SSMI14 had estimated a spatial correlation coefficient of 0.258. Both of the techniques have overestimated the radar accumulation. Also, the simple averaging technique has a slightly better result than the spatiotemporal correlation technique. The absolute error improvement made by the spatiotemporal correlation technique is negative and equal to $-0.9\%$. In other words, the spatiotemporal correlation technique has made a 0.9% larger absolute error than the simple averaging technique. Table 13 supports findings presented in the previous section that low values of spatial correlation coefficients and/or short time separations result in small improvements, or even in slight decreases in skill.

Another raining event that occurred on 30 April 2006 was detected by the TMI sensor. The 3-h rainfall accumulation period, starting at 1800 UTC, with satellite measurements of the instantaneous grid average rainfall at approximately 1900 and 2045 UTC is shown in Table 14. Instantaneous grid average rainfall measurements had values of 1.992 and 0.144 mm h$^{-1}$. Two measured fields had spatial correlation coefficients of 0.422 and 0.239, respectively, implying a higher uniformity than the fields presented in Table 13. Here, both techniques have underestimated the radar accumulation, except that the spatiotemporal correlation technique had a 0.847 mm h$^{-1}$ smaller absolute error than the simple averaging technique, providing an improvement of 41%.

Clearly, higher spatial correlation coefficients and longer time separations resulted in an increased skill of the spatiotemporal correlation technique.

5. Conclusions

The spatiotemporal correlation technique has been developed to improve estimations of 3-h satellite rainfall accumulations. Statistical properties of rain and technical
properties of instruments have been combined to assign
the weights to infrequent and limited instantaneous
rainfall measurements. When weighted, they were first
combined to form the estimates of the rainfall at the
times between the measurements, and then used to es-
timate a 3-h total rainfall accumulation. The results are
then compared to the simple averaging technique,
which takes a simple mean of the measurements as a
constant rainfall rate over the entire accumulation period.
The comparison is presented as improvements of the ab-
solute and the RMS errors. The results imply a potential
improvement of the total rainfall accumulations of cur-
rently used accumulation methods, such as CMORPH
and TMPA, if the technique would be implemented into
their estimates.

This new technique has shown skill in combining both
inaccurate and infrequent measurements. Rainfall fields
have characteristic temporal and spatial correlations,
which if used, can assist in making estimates between
satellite overpasses. It was shown that more uniform
rainfall fields have longer time correlations than less
uniform fields. It was also shown that the temporal
correlation length, or the time that the information from
a single measurement is still useful, is approximately
60 min. This greatly overcomes the temporal sparseness
of the measurements, provided that the measurements
are properly spaced over the accumulation period.

The best results are seen when the measurements are
120 min apart, and of equal levels of accuracy. This
corresponds to having the two measurements optimally
covering the entire accumulation period.

In addition to time separation dependence, the im-
provements from the spatiotemporal correlation technique
depend on the accuracy of the measurements. This tech-
technique has shown the capability of combining the mea-
surements of different accuracy, although the most
valuable results tend to occur if the measurements are
of similar accuracies. Results indicate that when the

![FIG. 8. A convective-like rain event occurred at 2200 UTC 20 Apr 2006 (NEXRAD stage IV hourly rainfall accumulations).]

![FIG. 9. A stratiform-like raining event occurred at 1800 UTC 30 Apr 2006 (NEXRAD stage IV hourly rainfall accumulations).]
spatiotemporal correlation technique is compared to the simple averaging technique, it creates improvement of 0%–50% in absolute error, and 0%–40% in RMS error, depending on the time separation and the measurement accuracy for the simulation study. Additionally, the spatiotemporal correlation technique rarely results in a deterioration of quality.

It is certain that the time separation between the measurements cannot be chosen or forced to its ideal length (120 min in this study). However, the results obtained by simulating the realistic scenario of having the length of time separation randomly distributed between 0 and 180 min are promising. They imply that the spatiotemporal correlation technique is capable of making up to a 25% improvement in absolute error and up to a 15% improvement in RMS error when compared to the simple averaging technique. This is nonnegligible, especially since the spatiotemporal correlation technique is computationally inexpensive.

Better synthetic data simulation that would account for more realistic satellite error distributions [as described in Villarini et al. (2009)] and more carefully performed real-data application, including longer data sets, should be the object of future studies.

Acknowledgments. Gratitude is expressed to Dr. Iwan Holleman from the Royal Netherlands Meteorological Institute for providing OPERA data support. This work was supported by Project NNX10AG75G.

REFERENCES


