Combining CMORPH and Rain Gauges Observations over the Rio Negro Basin

ALEJANDRA DE VERA AND RAFAEL TERRA

Instituto de Mecánica de los Fluidos e Ingeniería Ambiental, Facultad de Ingeniería, Universidad de la República, Montevideo, Uruguay

(Manuscript received 30 December 2011, in final form 22 May 2012)

ABSTRACT

Several algorithms that combine daily precipitation surface data and satellite Climate Prediction Center Morphing Technique (CMORPH) estimations were implemented and tested for the Rio Negro basin in northeastern Uruguay. Bias removal of satellite data through quantile matching—which requires historical data on nearby rain gauges—produces an unbiased estimate whose skill, as measured by the probability of detection (POD), is better than that obtained from surface observations for distances larger than approximately 50 km, which is twice the network characteristic distance between gauges of 23 km. Adjustment of satellite estimate using spatial interpolation of CMORPH deviations evaluated at nearby points—which requires simultaneous neighboring surface observations—eliminates biases to a large degree. Moreover, it shows higher POD skill than using only surface data for the entire range of distances and daily precipitation thresholds and for both seasons (cold and warm). The skill improvement attained, though, is small when the network density is as high as in the present study. However, these results suggest a promising scenario for the combined use of surface data and satellite retrievals as the latter continues to improve over time, both in resolution—spatial and temporal—and skill.

1. Introduction

The Rio Negro basin, in northeastern Uruguay, is a scenario of growing pressure for water by activities of great economic and strategic importance for the country, mainly in the agriculture and energy sectors. Water resource management in the context of a high spatial and temporal variability of rainfall poses a stringent requirement to precipitation monitoring. This situation contrasts with the endemic difficulty of maintaining a rain gauge network of sufficient quality and density. Remote sensing from satellite platforms offers an alternative, or rather complementary, opportunity. In recent years, several satellite-based precipitation estimates have become available in near–real time. In particular, the Climate Prediction Center Morphing Technique (CMORPH; Joyce et al. 2004), which is a combination of geostationary images from infrared and microwave sensors, provides precipitation estimates with different spatial and temporal resolutions since December 2002. All satellite estimates are indirect measures of precipitation and thus require validation, and eventually calibration, against surface observations, which will always be needed in the daunting task of determining the precipitation field.

Researchers have increasingly moved toward a combined use of satellite and gauge data to improve accuracy, coverage, and resolution. Several approaches have been proposed for merging rainfall satellite estimates and surface data into a single, best-estimate, dataset (Huffman et al. 1997; Adler et al. 2003; Xie et al. 2003). A recent review by Vila et al. (2009) presents a comprehensive assessment of a high-resolution, gauge–satellite-based analysis of daily precipitation over continental South America during 2004. Intercomparisons and cross-validation tests were carried out between independent rain gauges and different merging techniques, including the control algorithm [Tropical Rainfall Measuring Mission (TRMM) Multisatellite Precipitation Analysis real-time algorithm]. The methodologies were tested for different months and seasons and for different network densities. All the merging schemes produce better results than the control algorithm, and when finer temporal- (daily) and spatial-scale (regional networks) gauge datasets are included in the analysis, the improvement is remarkable. Moreover, Ruiz (2009) describes the

Corresponding author address: Alejandra De Vera, IMFIA, Facultad de Ingeniería, Universidad de la República, J. Herrera y Reissig 565, C.P. 11300, Montevideo, Uruguay.
E-mail: adevera@fing.edu.uy

DOI: 10.1175/JHM-D-12-010.1

application of several methods for the calibration of precipitation estimates generated from passive microwave sensors using rain gauge observations over southeastern South America. Results show that the proposed calibration algorithm can effectively reduce the systematic errors in the precipitation estimates, leading to an improvement on its reliability over the study region. Also, Rozante et al. (2010) proposed a new technique (called MERGE) to combine TRMM satellite precipitation estimates with surface observations over the South American continent. Its performance is evaluated for the 2007 summer and winter seasons. Results show that over areas with a high density of observations, the MERGE technique's performance is equivalent to that of simply averaging the stations within the grid boxes. However, over areas with sparse observations, MERGE shows superior results.

In addition, Xie and Xiong (2011) developed a conceptual model to construct high-quality, high-resolution precipitation analyses over land by merging information from gauge observations and CMORPH satellite estimates using data over China for a 5-month period. A two-step strategy was adopted. First, bias correction is performed for the CMORPH estimates by matching the probability density function (PDF) of the satellite data with that of the daily gauge analysis. Then, the bias-corrected CMORPH precipitation estimates are combined with the gauge analysis through the optimal interpolation (OI) technique. Validation against independent gauge observations demonstrates feasibility and effectiveness of the conceptual algorithm, with the merged precipitation analysis showing substantially smaller bias and significantly improved pattern agreements compared to both the input gauge and the satellite data alone. Similar two-step approaches have been addressed in Krajewski (1987) and Seo and Breidenbach (2002).

Existing analyses over South America are continental or regional in scale, having at most a dozen observation stations within Uruguay, while all of the referenced studies cover a time period that ranges from a few months to a year. The present study describes and evaluates several methodologies for merging CMORPH satellite precipitation estimate and daily gauge data at a basin scale (the Rio Negro basin in northeastern Uruguay) with a high density of surface observations and a 7-yr-long record for comparison. Our objective is not to evaluate the CMORPH estimate, but to compare several ways of using CMORPH to improve estimates of precipitation when used in combination with gauge data. We evaluate the combined product (with the different methodologies) at points rather than area averages because point precipitation is known with more certainty, which makes the determination of skill more robust. The long period of study allows for both a better sampling of larger events (up to 100 mm day^{-1}), while the high-density network enables detailed analyses of the added value of the satellite estimate as a function of distance between surface observations.

The paper is organized as follows: Section 2 describes the datasets used while section 3 presents the skill scores used and the ability of the existing rain gauge network to estimate the precipitation at a generic point. Section 4 identifies CMORPH bias and proposed different methodologies to remove it. Next, in section 5, the evaluation of the incremental skill of combined precipitation estimates is discussed. Finally, conclusions are presented in section 6.

2. Datasets

a. Rain gauge database

Precipitation data comes from the Administración Nacional de Usinas y Trasmisiones Eléctricas (UTE; public electric utility) network, which is composed of 133 rain gauges within the Rio Negro basin. Daily rainfall totals are taken at 1000 UTC. The period analyzed is from 2003 to 2009, during which a CMORPH satellite estimate is also available. Warm season (October–March) and cold season (April–September) were analyzed separately to assess the seasonality of skill for the different precipitation estimates. The amount of data available is not enough to further disaggregate the annual cycle without compromising the statistical significance of the results.

The Rio Negro basin, with an area of 71 193 km^2, is one of the largest watersheds in Uruguay and plays a key role in the hydroenergy production of the country. For this reason, the precipitation station density is relatively high with a characteristic separation between gauges of 23 km.

b. Rainfall satellite estimate, CMORPH

The satellite-based algorithm analyzed is CMORPH, which combines the superior retrieval accuracy of passive microwave estimates and the higher temporal and spatial resolution of infrared data, with no ground-based information (Joyce et al. 2004). CMORPH estimates are accessible online at different spatial and temporal resolutions, with a maximum spatial resolution of 0.0727° × 0.0727° and a maximum frequency of half an hour (information available in rotating files that contain the most recent 31 days). The entire CMORPH record, uninterrupted since December 2002, is available at a 3-h frequency and a spatial resolution of 0.25° × 0.25°. Since the quantile-matching technique requires an accurate representation of the climatological PDF of precipitation

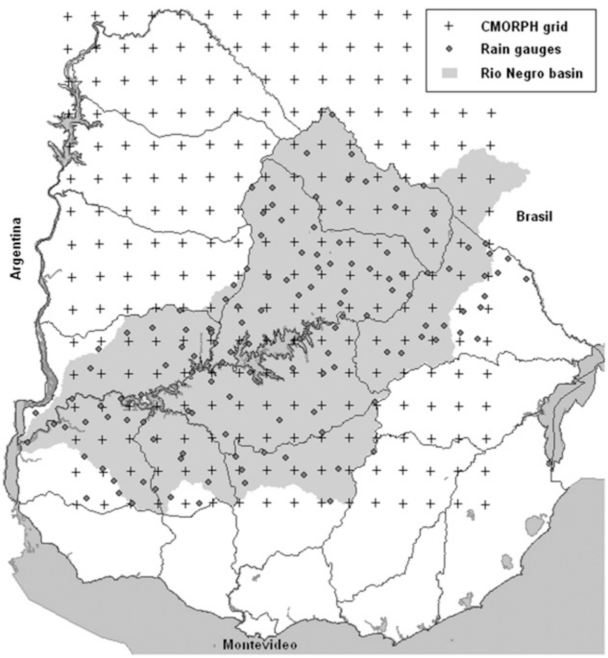

FIG. 1. Rain gauge network and the CMORPH grid in the Rio Negro basin.

intensity, the latter dataset was selected, prioritizing length of record over spatial and temporal resolution. However, the present study is limited to daily precipitation totals, since this is the information available at the rain gauges. The daily accumulation is obtained adding the individual 3-h amounts from 0900 UTC of one day to 0900 UTC of the next. There is, therefore, an inevitable 1-h lag between the satellite estimate and the rain gauge records.

Figure 1 shows the spatial distribution of the rain gauge network and the CMORPH grid, which has 108 points within the Rio Negro basin.

3. Skill scores

We determine the performance of the combined precipitation estimates based on the bias score (BIAS), the probability of detection (POD), the false alarm ratio (FAR), and the equitable threat score (ETS) for different precipitation thresholds, based on a 2×2 contingency table (Table 1) containing all four possible combinations of forecast–event pairs (Wilks 2006).

The total number of forecasts, $N = A + B + C + D$, corresponds to the number of days in the period of study. Notably, in our case, D is far larger than A, B, and C because it includes the days without rain that were correctly forecasted.

The BIAS $= (A + B)/(A + C)$ compares the number of events forecasted against the ones observed.

TABLE 1. Two-by-two contingency table.

	CMORPH $> X$ mm	CMORPH $\leq X$ mm
Rain gauge $> X$ mm	A (hits)	C (misses)
Rain gauge $\leq X$ mm	B (false alarms)	D (correct negatives)

Unbiased forecasts exhibit BIAS = 1, indicating that the event was forecasted the same number of times that it was observed. However, it provides no information about the correspondence between the forecasts and observations. BIAS > 1 indicates overforecasting and BIAS < 1 underforecasting.

The POD $= A/(A + C)$ gives the fraction of events that were correctly detected. It therefore ranges from 0 to 1 with a perfect score of 1.

The FAR $= B/(A + B)$ is the fraction of yes forecasts that turn out to be wrong—that proportion of the forecast events that fail to materialize. It thus ranges from 0—best possible scenario—to 1.

In the case of an unbiased estimator, $A + B = A + C$, and therefore POD $= 1 -$ FAR. In these cases, POD and FAR give essentially the same information.

The ETS $= (A + A_{\mathrm{ref}})/(A + B + C - A_{\mathrm{ref}})$ where $A_{\mathrm{ref}} = (A + B)(A + C)/N$ measures the fraction of observed events that were correctly estimated, adjusted for hits associated with random chance. It therefore ranges from 0 to 1 with a perfect score of 1 (a value of 0 represents an estimate where the number of hits is similar to that obtained by chance expressed by the amount A_{ref}).

Each score was computed for daily precipitation thresholds ranging from 0 to 100 mm every 5 mm.

Figure 2 shows the total number of cases as a function of daily rainfall totals throughout the period analyzed (2003–09) for the warm and cold season separately. It shows that the number of cases sampled is quite low for the highest thresholds. Anyway, the analyses were performed up to 100 mm, aware of the limitations posed by the limited sampling.

In addition to the skill scores based on the contingency table, we computed other parameters to measure the ability of the different schemes, such as the root-mean-square error (RMSE) and the linear correlation coefficient.

a. Statistical significance

To assess the statistical significance of results presented later, it is relevant to compute the POD of a random (but unbiased) precipitation estimate. Consider N total number of days in the sample; in our case and for each season, we have 7 years of data that amount to approximately $N = 1250$ days. There are M rainy days for each precipitation threshold; from Fig. 2 we see than M ranges from 2 (100 mm) to almost 200 (5 mm).

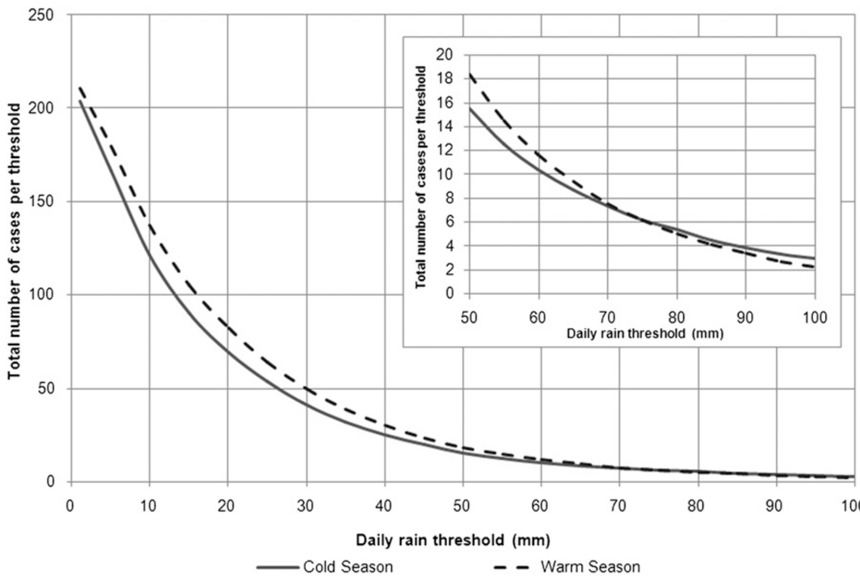

FIG. 2. Total number of events as a function of daily precipitation threshold for each season. The inset gives more detail at the high threshold values.

An unbiased estimator will have M rainy days distributed among the N possible spots, with S hits ($0 \leq S \leq M$). If the estimator is random, the probability of S hits can be computed from combinatory theory:

$$\text{prob}_{\text{hits}=S} = \frac{C_M^S C_{N-M}^{M-S}}{C_N^M} \tag{1}$$

$$\text{where} \quad C_k^n = \frac{n!}{k!(n-k)!}. \tag{2}$$

The POD significant at a 95% confidence level is such that the number of associated hits S^* has a 0.05 chance of being topped randomly, thus

$$\sum_{S=0}^{S^*} \text{prob}_{\text{hits}=S} = 0.95. \tag{3}$$

For the given N and each M, S^* can be computed from the previous expressions. POD at 95% significance level is then S^*/M.

For small M (large precipitation threshold), even if the chance of randomly getting a hit is low, the impact in POD score is very large, resulting in a sharp increase in the level of the statistically significant POD as the number of events runs short. Conversely, as M—or rather M/N—becomes large (low precipitation threshold) the chance of getting random hits grows and so does the significant POD. In between, Fig. 3 shows an ample range of M (between 10 and 60) for which the statistically significant POD is about 0.1. For the vast majority of cases considered below, the significant POD is less than 0.15.

b. Ability of the rain gauge network

As a reference, we first compute the ability of the existing rain gauge network to estimate the precipitation at a generic point by determining the unbiased probability of detection of daily precipitation as a function of distance and rainfall amounts.

For each pair (8778 in all) of stations, the unbiased relative POD was computed; bias was removed following the quantile-matching technique presented in section 4b(1). POD scores were sorted by the distance among the stations involved and then averaged in 10-km bins to construct the contour plots shown in Fig. 4.

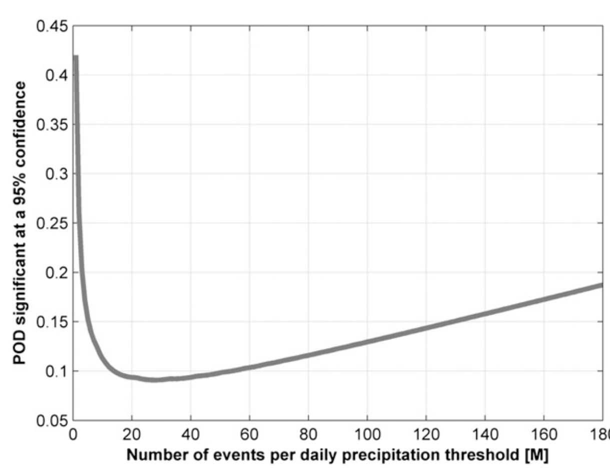

FIG. 3. POD at the 95% significance level as a function of number events M, for a sample with N = 1250 days.

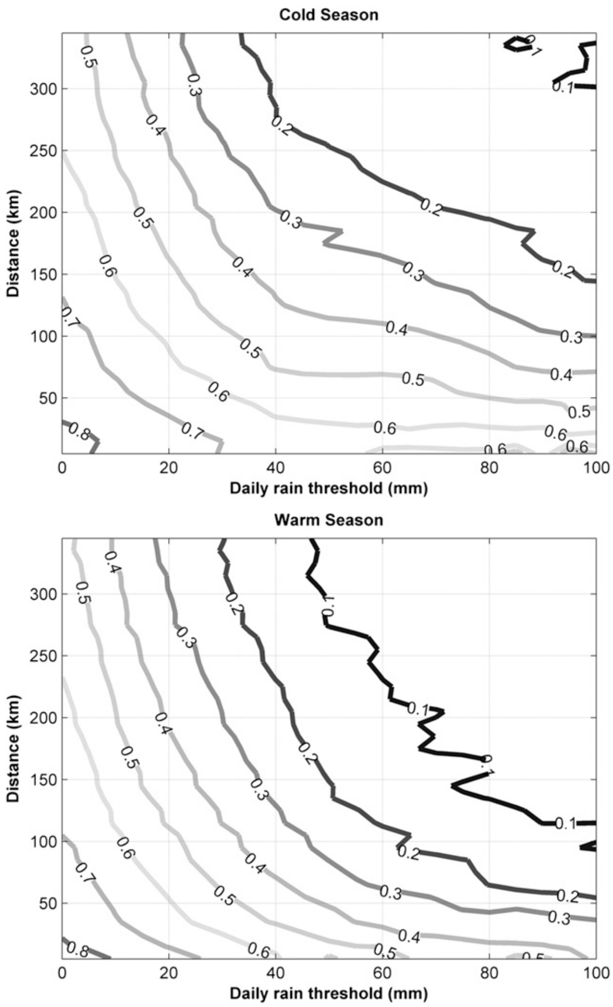

FIG. 4. POD as a function of distance and daily rain threshold for (top) cold and (bottom) warm seasons—rain gauge network.

Separate analyses were performed for the April–September semester (cold season) and October–March semester (warm season), considering that characteristic precipitating systems may differ and therefore the spatial correlation of precipitation could potentially change. We indeed observe slightly higher levels of detection for a given distance in the cold season, when frontal systems are dominant, than in warm season, when smaller-scale convective systems are more frequent. This is most evident for large thresholds, as expected. For instance, for an intermediate threshold of 50 mm and a distance of 25 km (characteristic of the rain gauge network), the probability of detection is 0.47 in the warm season and 0.63 in the cold season, confirming the faster spatial decorrelation in the former case.

These results summarize the present ability of the rain gauge network alone to estimate the precipitation at a generic point in the basin. We next explore the ability of the satellite retrievals to improve the skill when used in combination with gauge data.

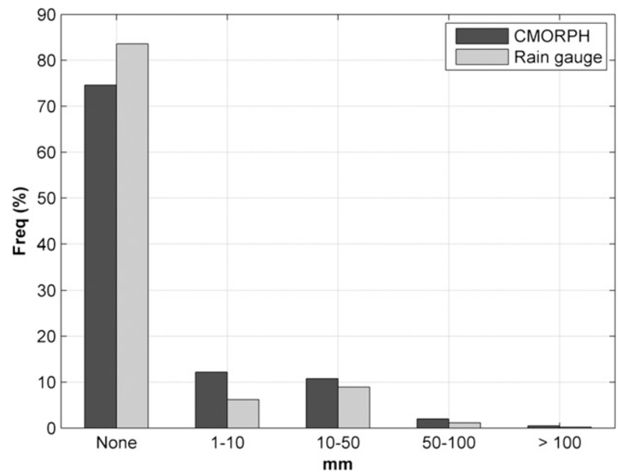

FIG. 5. Global PDF (%) of daily precipitation in the Rio Negro basin for 133 rain gauges and 108 CMORPH grid points—period 2003–09.

4. Methodology

a. *CMORPH bias identification*

To describe CMORPH bias, Fig. 5 displays the global (all stations during the entire period) PDFs of daily precipitation in the Rio Negro basin for both the observational network and CMORPH estimate. It shows that CMORPH overestimates precipitation frequency for all intensities, but especially low-amount events. CMORPH bias should not be interpreted as equivalent to error, since part of the bias is due to the fact that we are comparing an area-averaged estimation against a point observation.

Figure 6 shows the average bias of CMORPH estimates when compared to the nearest gauge and its standard deviation among the 133 stations in the Rio Negro basin. The satellite estimate has a large bias—it overestimates rainfall—for all daily precipitation thresholds, while the standard deviation is relatively small for low thresholds and increases steadily with daily precipitation amounts.

Previous results confirm the need to implement a bias removal scheme to the satellite estimator before combining it with rain gauge data.

b. *Bias removal schemes*

1) HISTORICAL QUANTILE MATCHING

One way to remove the CMORPH bias, when historical information of both the estimator and observed data in a nearby station is available, is through quantile matching (Panofsky and Brier 1958; Déqué 2007). Historical

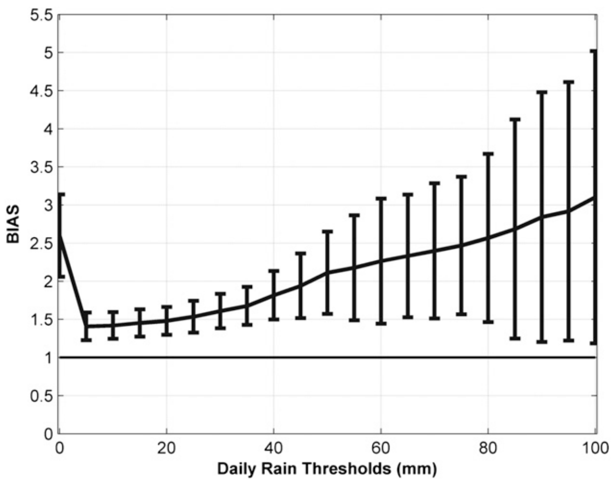

FIG. 6. Average and standard deviation of BIAS of CMORPH estimate with respect to the nearest-neighbor stations in the Rio Negro basin.

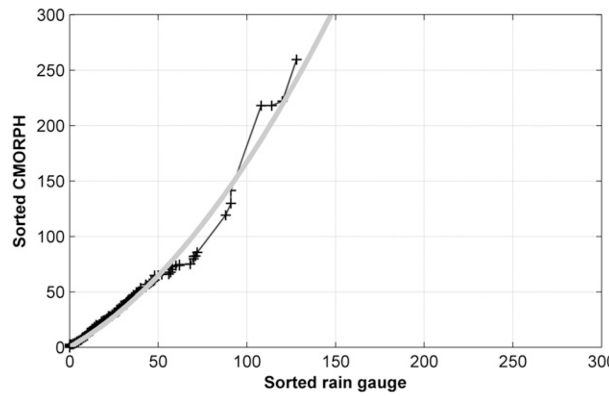

FIG. 7. CMORPH vs rain gauge—bias removal through quantile matching (polynomial fitting is shown for illustration but not used).

records, from both rain gauges and CMORPH, are used to define the function that relates the distributions of both datasets. This can be made parametrically (for which we must assume a particular analytical distribution) or with a nonparametric approach; in this case we chose the latter method.

To determine the piece-linear function that defines the matching, CMORPH estimations are compared against the observed precipitation in the nearest rain gauge, having previously ordered both records separately. Figure 7 shows a particular case; every point on the graph shares the same percentile in their respective time series but not the same date. In the presence of a new value of the estimator, the bias can be removed by calculating the value that shares the same percentile in the distribution of the gauge by interpolation between the closest historical values.

Figure 7 highlights the limitations of the nonparametric quantile-matching approach when dealing with extreme events, where historical sampling is sparse. In particular, in the presence of a new maximum in CMORPH estimation, the historical observed maximum of the nearest rain gauge was used.

The detailed methodology is as follows: First, every CMORPH estimate during the period is quantile matched on its own grid using the historical information from the nearest rain gauge. The matched estimations are then interpolated (either linearly or taking the nearest neighbor) from CMORPH grid to the stations and compared against actual observations to compute scores.

2) ADJUSTMENT BASED ON SIMULTANEOUS NEARBY STATIONS

If simultaneously observed records in neighbor stations are available, these can be used, combined with CMORPH, to assess a spatial description of the differences between the two platforms, which in turn can be used to adjust CMORPH estimation in a generic point. No historical data of any kind are required. Systematic biases are not guaranteed to be eliminated through this adjustment, as was the case with quantile matching. However, later results will show that the CMOPRH estimation adjusted in this way presents very small biases.

The detailed methodology is as follows: First, CMORPH estimates are interpolated (linearly or nearest neighbor) to each rain gauge and the difference is computed at said points by subtracting the simultaneously observed values. Withdrawing one station at a time in a cross-validation approach, these differences are further interpolated (with same method as before) from all remaining gauges and used to adjust the interpolated CMORPH value at the withdrawn station. Finally, the adjusted estimate is compared against the actual observation to compute scores.

3) QUANTILE MATCHING AND ADJUSTMENT

In the presence of both historical information and simultaneous observations in nearby rain gauges, we can combine the quantile matching with the adjustment by interpolation of CMORPH deviations assessed at neighboring stations. This combination makes full use of both satellite and rain gauge information, both historical and simultaneous.

The detailed methodology is as follows: First, every CMORPH estimate is quantile matched on its own grid using the historical information from the nearest station as in 1 above. Second, matched CMORPH estimates are interpolated (linearly or nearest neighbor) to each station and the difference is computed at said points by subtracting the simultaneously observed values. Withdrawing one station at a time, these differences are further interpolated (with same method as before) from all remaining gauges and used to adjust the interpolated matched CMORPH value at the withdrawn station. Finally, the matched-adjusted estimate is compared against actual observation to compute scores.

Note that methodologies 2 and 3 above can generate negative precipitation estimates. Scores based on the contingency table are not affected by negative values but both the RMSE and the correlation coefficient—as most real-life applications—are. In these cases, zero precipitation was assumed.

We end by summarizing the different methodologies presented above. In every case an interpolation is performed we used both linear and nearest neighbor:

1) Quantile-matching at CMORPH grid (CG) using nearest rain gauge (RG) + interpolation from matched CG to RG.
2) Interpolation from CG to RG + compute difference at RG + interpolation of the difference from RG (excluding self-point) to RG.
3) Quantile matching at CG using nearest RG + interpolation from matched CG to RG + compute difference at RG + interpolation of difference from RG (excluding self-point) to RG.

c. *Assessment and selection of methodology*

Methodologies 1, 2, and 3 were implemented for 121 stations internal to Rio Negro basin for the entire period of study; those stations in the border of the basin were left out of the analysis since the spatial interpolation from neighboring stations is ill defined.

Figures 8 and 9 present the average bias for the Rio Negro basin and its standard deviation based on the quantile-matched, adjusted, and matched-adjusted CMORPH, with linear and nearest-neighbor interpolation, respectively.

Figures 10 and 11 show the average and standard deviation of POD in the Rio Negro basin based on the quantile-matched, adjusted, and matched-adjusted CMORPH, with linear and nearest-neighbor interpolation, respectively.

A combined evaluation of Figs. 8–11 shows that the adjustment based on simultaneous nearby stations (methodology 2) greatly improves the probability of detection as compared with quantile-matched alone (methodology 1), without a significant detrimental effect on the bias. Within the adjusted estimates, the impact of a previous quantile matching is extremely small in both the bias and POD scores. Considering the extra requirement in historical data, its use is not justified. However, in the absence of simultaneous nearby observations with which to compute the adjustment, the historical quantile matching is still necessary and useful to remove the bias.

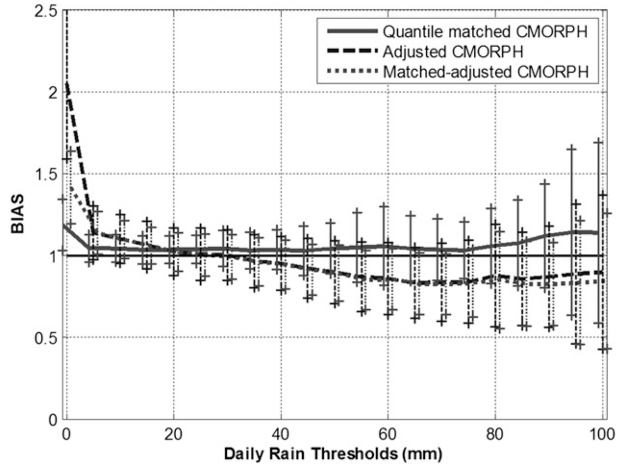

FIG. 8. Average and standard deviation of BIAS in the Rio Negro basin, period 2003–09—quantile-matched, adjusted, and matched-adjusted CMORPH with linear interpolation.

Tables 2 and 3 show the RMSE and the correlation coefficient for the three schemes proposed and with linear and nearest-neighbor interpolation. Linear interpolation has a better performance than the nearest neighbor (lower RMSE and higher correlation coefficient) in all cases. Quantile matching again underperforms compared to the adjusted methodologies. Within the latter, the impact of a previous quantile matching is extremely small in both RMSE and correlation, as was the case with the skill scores.

Overall, linear interpolation introduces somewhat larger bias but also a slightly higher chance of detection than the nearest-neighbor interpolation and has significantly better RMSE and correlation coefficient. Therefore, linear interpolation is selected hereafter in view of the higher skill in the parameters considered most relevant to the purpose of the study.

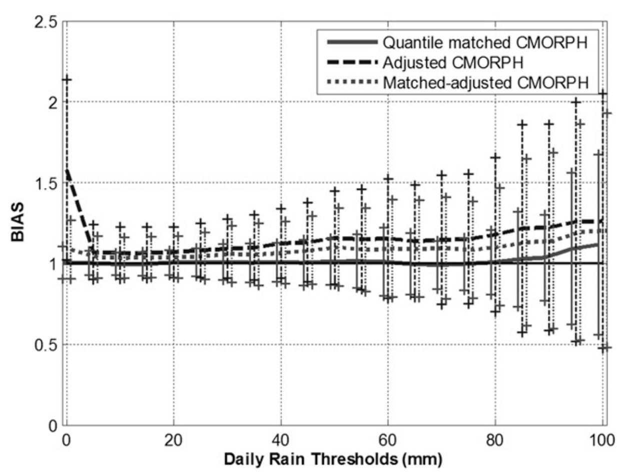

FIG. 9. As in Fig. 8, but with nearest-neighbor interpolation.

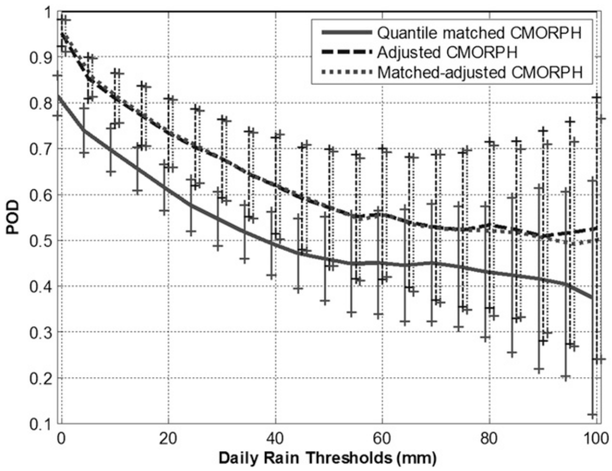

FIG. 10. Average and standard deviation of POD in the Rio Negro basin, period 2003–09—quantile-matched, adjusted, and matched-adjusted CMORPH with linear interpolation.

Figure 12 presents the average ETS for the Rio Negro basin and its standard deviation based on the linear interpolated quantile-matched, adjusted, and matched-adjusted CMORPH. The relative performance among the methodologies is the same as for all other measures of skill presented.

5. Incremental skill of combined precipitation estimates

Next, we assess the incremental skill of the combined precipitation estimates comparing with the situation in which only rain gauge data is available. For this purpose we compute the probability of detection as a function of distance and daily precipitation threshold as in section 3b.

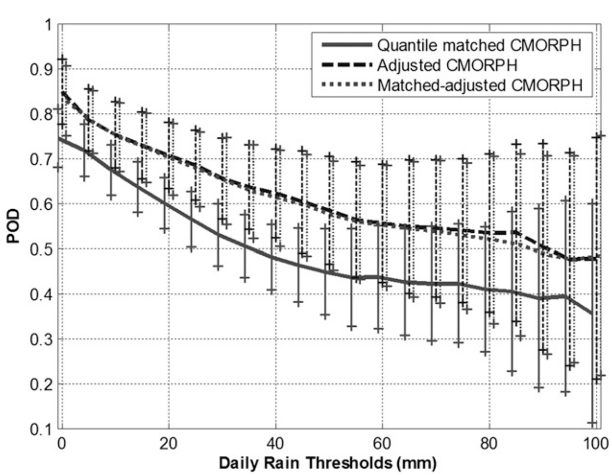

FIG. 11. As in Fig. 10, but with nearest-neighbor interpolation.

TABLE 2. RMSE for the different schemes proposed for both linear and nearest-neighbor interpolation.

Methodology	Linear interpolation (mm)	Nearest-neighbor interpolation (mm)
Quantile matched	8.0	8.2
Adjusted	6.2	7.5
Matched adjusted	6.1	7.4

a. Quantile-matched CMORPH

We showed that, if historical records are available, the removal of systematic biases of CMORPH could be achieved through quantile matching. The matched estimate does not use any simultaneous rain gauge data and therefore the probability of detection does not depend on distance but only on the daily rainfall threshold as shown in Figs. 10 and 11.

Figure 13 shows the incremental POD contour plots for the matched CMORPH minus that with the rain gauge data only. As before, the results are presented separately for the cold and warm season. For distance smaller than approximately 50 km—with variations depending on the season and threshold—contours are negative, indicating that interpolation from rain gauges has greater POD than matched CMORPH estimate. For longer distances, though, the quantile-matched CMORPH provides a more skillful estimate, especially during the cold season.

b. Adjusted CMORPH

We showed that, in the presence of simultaneous observations in nearby stations, adjusting CMORPH estimates based on spatial interpolation of the differences with rain gauge data largely corrects the bias, and that no further skill is gained by previously quantile-matching CMORPH, which requires historical data.

In Fig. 14 we present the contour plots of the POD of linearly interpolated adjusted CMORPH as a function of daily precipitation threshold and distance to the station, which is used to estimate CMORPH deviations. Again, the seasonal differences in POD are evident with larger values during the cold season, especially for large thresholds. A quick comparison with Fig. 4 highlights the much larger skill of the adjusted CMORPH for large distances, as the POD of the gauge network vanishes and

TABLE 3. Correlation coefficient for the different schemes proposed for both linear and nearest-neighbor interpolation.

Methodology	Linear interpolation	Nearest-neighbor interpolation
Quantile matched	0.75	0.74
Adjusted	0.85	0.79
Matched adjusted	0.85	0.79

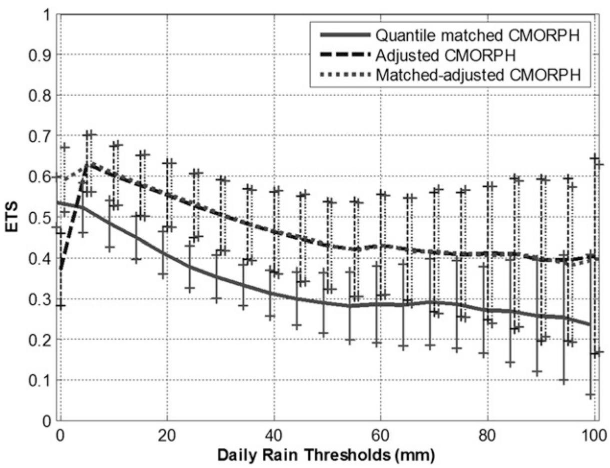

FIG. 12. Average and standard deviation of ETS in the Rio Negro basin, period 2003–09—quantile-matched, adjusted, and matched-adjusted CMORPH with linear interpolation.

the performance of the adjusted estimate is determined by the skill of the local interpolated CMORPH.

For a more detailed comparison, in Fig. 15 we present the contour plots with the difference between the POD of linearly interpolated adjusted CMORPH estimations minus that of the rain gauge network. It represents the benefit in skill, as measured by POD, of using satellite information on top of the rain gauge data, as a function of the distance and the daily precipitation threshold.

In general, Fig. 15 shows that the addition of CMORPH data to the surface observations generates an improvement of rainfall estimation for virtually the entire range analyzed in both seasons. It has to be noted that, although the improvement in POD for large distances is very significant, for smaller ones (less than approximately 100 km) the impact is marginally significant from a statistical point of view (see section 3a). Seasonality of the differences is much less pronounced than in the absolute skill itself.

6. Summary and conclusions

We first identified the bias of the satellite rainfall estimate CMORPH within the Rio Negro basin in northeastern Uruguay. Next we devised several bias removal schemes from which we developed three methodologies to combine the satellite retrievals with rain gauge data. The following methodologies were implemented and tested:

1) a quantile-matching algorithm that removes the bias in the distribution of daily precipitation intensity forcing the distribution of the estimator to coincide with the historical distribution at a nearby rain gauge;
2) an adjustment that makes use of simultaneous nearby surface observations to estimate CMORPH deviations, which is then interpolated to the target point and removed; and
3) a method that combines the two previous ones: it first applies the quantile matching of CMORPH estimates and then adjusts through interpolation of the remaining differences between CMORPH and the neighboring gauges.

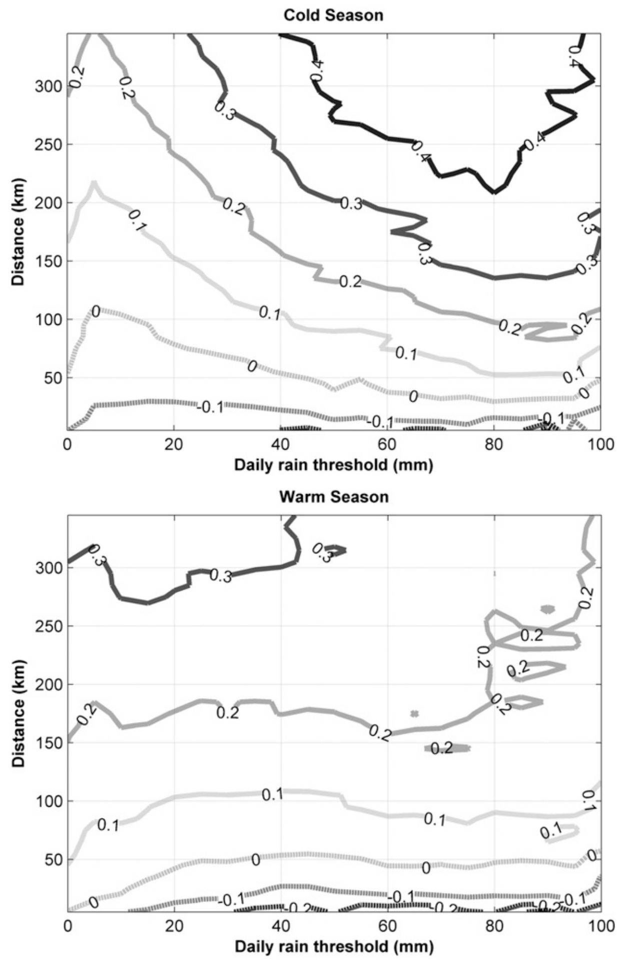

FIG. 13. Differences in POD between the matched CMORPH estimation and the rain gauge network as a function of distance and daily rain threshold for (top) cold and (bottom) warm seasons.

The quantile-matched CMORPH (which requires only historical data on nearby rain gauges) has a higher probability of detection (POD)—compared with that obtained from surface observations only—for distances greater than approximately 50 km, but not in the case of Rio Negro basin where the characteristic distance between gauges is 23 km.

Within the adjusted estimates, the impact of a previous quantile matching is very small in all scores computed. Considering that the matched-adjusted CMORPH 3) requires historical data and simultaneous neighboring

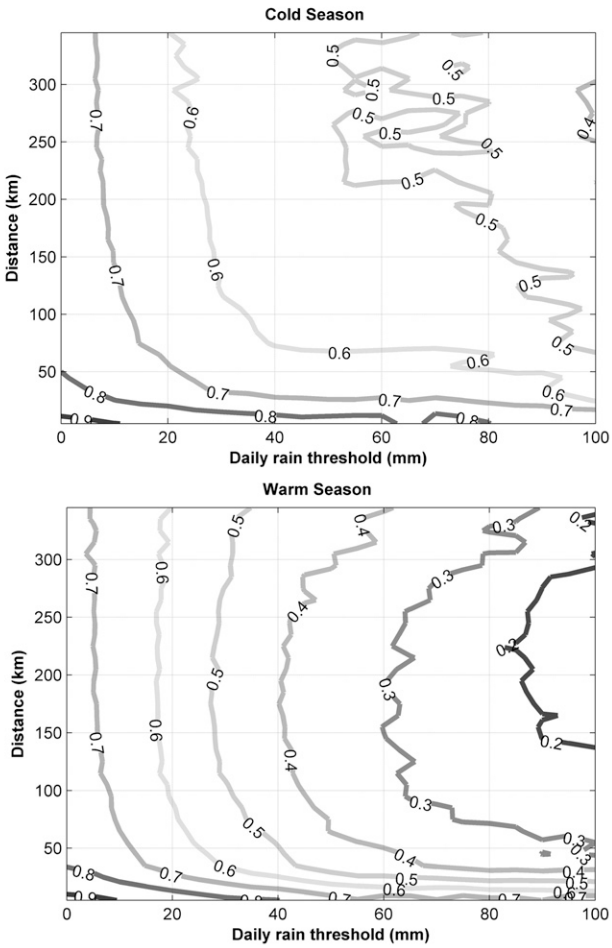

FIG. 14. POD as a function of distance and daily rain threshold for (top) cold and (bottom) warm seasons—linearly interpolated adjusted CMORPH.

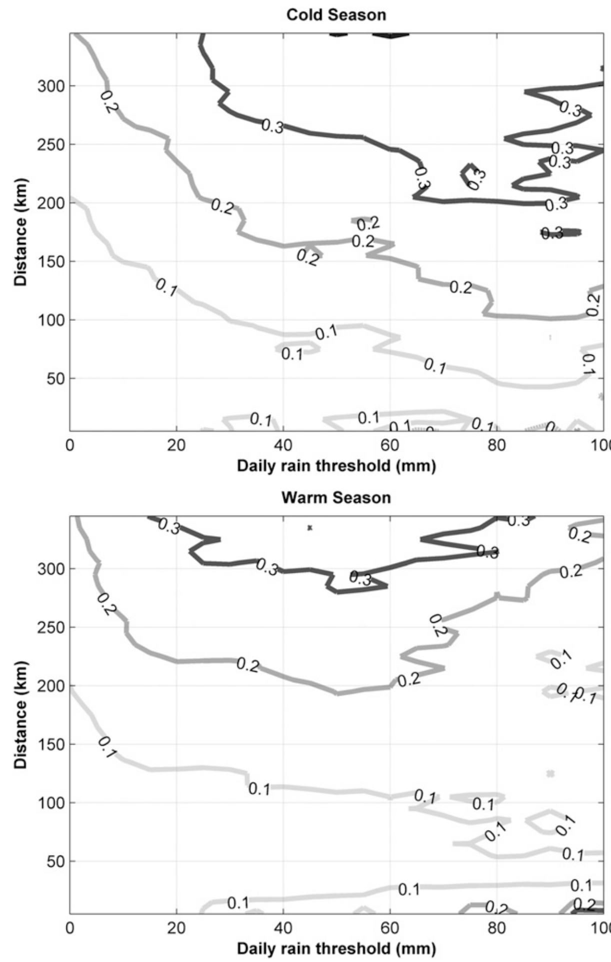

FIG. 15. Differences in POD between the linearly interpolated adjusted CMORPH estimation and the rain gauge network as a function of distance and daily rain threshold for (top) cold and (bottom) warm seasons.

observations, and adjusted CMORPH 2) only needs the simultaneous surface data, the use of the matched-adjusted estimates is not justified. However, in the absence of simultaneous nearby observations with which to compute the adjustment, the historical quantile matching, as in 1) above, is still necessary and useful to remove the bias.

The adjusted estimate 2) performs better, as measured by POD, than using only surface data, for the entire range of distances and daily precipitation thresholds and for both seasons. It has to be noted, though, that the improvement in skill is marginally significant for distances less than approximately 100 km.

The incremental skill attained in the representation of the precipitation field through the appropriate addition of satellite retrievals to the rain gauge data is already measurable, although not very significant when the network density is high as in the present study. However, it is reasonable to expect that the satellite products continue to improve over time (both in resolution—spatial and temporal—and skill), providing an encouraging scenario for the combined use of satellite technology and rain gauges in the determination of rainfall distribution.

Acknowledgments. We very much appreciate the comments and suggestions of three anonymous reviewers whose input greatly helped to improve the manuscript. The first author enjoyed a research initiation fellowship from ANII that greatly contributed to this work. Part of the work was conducted in the framework of a PROHIMET-WMO project. We are grateful to UTE for the availability of the rain gauge data.

REFERENCES

Adler, R. F., and Coauthors, 2003: The Version-2 Global Precipitation Climatology Project (GPCP) monthly precipitation analysis (1979–present). *J. Hydrometeor.,* **4,** 1147–1167.

Déqué, M., 2007: Frequency of precipitation and temperature extremes over France in an anthropogenic scenario: Model results and statistical correction according to observed values. *Global Planet. Change,* **57,** 16–26.

Huffman, G. J., and Coauthors, 1997: The Global Precipitation Climatology Project (GPCP) combined precipitation dataset. *Bull. Amer. Meteor. Soc.,* **78,** 5–20.

Joyce, R. J., J. E. Janowiak, P. A. Arkin, and P. Xie, 2004: CMORPH: A method that produces global precipitation estimates from passive microwave and infrared data at high spatial and temporal resolution. *J. Hydrometeor.,* **5,** 487–503.

Krajewski, W. F., 1987: Cokriging radar-rainfall and rain gage data. *J. Geophys. Res.,* **92** (D8), 9571–9580.

Panofsky, H. A., and G. W. Brier, 1958: *Some Applications of Statistics to Meteorology*. Pennsylvania State University Press, 224 pp.

Rozante, J. R., S. M. Demerval, L. G. G. de Goncalves, and D. A. Vila, 2010: Combining TRMM and surface observations of precipitation: Technique and validation over South America. *Wea. Forecasting,* **25,** 885–894.

Ruiz, J. J., 2009: Evaluación de diferentes metodologías para la calibración de las estimaciones de precipitación CMORPH sobre Sudamérica. *Rev. Bras. Meteor.,* **24,** 473–488.

Seo, D. J., and J. P. Breidenbach, 2002: Real-time correction of spatially nonuniform bias in radar rainfall data using rain gauge measurements. *J. Hydrometeor.,* **3,** 93–111.

Vila, D. A., L. G. G. de Goncalvez, D. L. Toll, and J. R. Rozante, 2009: Statistical evaluation of combined daily gauge observations and rainfall satellite estimates over continental South America. *J. Hydrometeor.,* **10,** 533–543.

Wilks, D. S., 2006: *Statistical Methods in the Atmospheric Sciences.* 2nd ed. Academic Press, 627 pp.

Xie, P., and A.-Y. Xiong, 2011: A conceptual model for constructing high-resolution gauge-satellite merged precipitation analyses. *J. Geophys. Res.,* **116,** D21106, doi:10.1029/2011JD016118.

——, J. E. Janowiak, P. A. Arkin, R. Adler, A. Gruber, R. Ferraro, G. J. Huffman, and S. Curtis, 2003: GPCP pentad precipitation analyses: An experimental dataset based on gauge observations and satellite estimates. *J. Climate,* **16,** 2197–2214.