Scale Characterization and Correction of Diurnal Cycle Errors in MAPLE

AITOR ATENCIA[a] AND ISZTAR ZAWADZKI

Department of Atmospheric and Oceanic Sciences, McGill University, Montreal, Quebec, Canada

MARC BERENGUER

Centre de Recerca Aplicada en Hidrometeorologia, Universitat Politècnica de Catalunya, Barcelona, Spain

(Manuscript received 17 October 2016, in final form 6 July 2017)

ABSTRACT

The most widely used technique for nowcasting of quantitative precipitation in operational and research centers is the Lagrangian extrapolation of the latest radar observations. However, this technique has a limited forecast skill because of the assumption made on its formulation, such as the fact that the motion vectors do not change and, even more important for convective events, neglect any growth or decay in the precipitation field. In this work, the McGill Algorithm for Precipitation Nowcasting by Lagrangian Extrapolation (MAPLE) errors have been computed for 10 yr of radar composite data over the continental United States. The study of these errors shows systematic bias depending on the time of day. This effect is related to the solar cycle, whose heating energy results in an increase in the average rainfall in the afternoon. This external forcing interacts with the atmospheric system, creating local initiation and dissipation of convection depending on orography, land use, cloud coverage, etc. The signal of the diurnal cycle in MAPLE precipitation forecast has been studied in different locations and spatial scales as a function of lead time in order to recognize where, when, and for which spatial scales the signal is significant. This information has been used in the development of a scaling correction scheme where the mean errors due to the diurnal cycle are adjusted. The results show that the developed methodology improves the forecast for the spatial scales and locations where the diurnal cycle signal is significant.

1. Introduction

Lagrangian extrapolation of the latest radar precipitation (e.g., Germann and Zawadzki 2002; Seed 2003; Berenguer et al. 2005) has shown skill in forecasting the precipitation field with high resolution for lead times up to 6 h (known as nowcasting) or even for longer lead times up to 12 h [known as very short-range forecasting (VSRF)]. This technique estimates the motion of the precipitation field with a tracking algorithm and generates the forecasts by using the estimated motion field to advect the last available radar precipitation observations.

Tracking algorithms capture the motion resulting from the combined effect of (i) the movement of the precipitation transported by the steering-level winds and (ii) the apparent displacement due to systematic growth and decay at the edges of the precipitation systems (Germann et al. 2006).

These authors and, at the same time, Bowler et al. (2006) found that the largest errors in nowcasting techniques based on the advection of the latest observations are due to the evolution of storms, including the development of new ones. These processes are not explained by the apparent motion field.

Initiation and decay of new thunderstorms linked to the diurnal cycle of radiation have a dominant role in the precipitation activity of the warm season in the central-east United States. This diurnal cycle of precipitation shows a maximum in the afternoon across the continent because of the initiation and dissipation of convection (e.g., Wallace 1975; Dai et al. 1999; Carbone et al. 2002; Janowiak et al. 2005; Parker and Ahijevych 2007; Carbone and Tuttle 2008; Surcel et al. 2010) and has a secondary maximum in the central Plains in the late evening caused by the arrival of long-lived mesoscale convective systems traveling from the Rocky Mountains

[a] Current affiliation: Zentralanstalt für Meteorologie und Geodynamik, Vienna, Austria.

Corresponding author: Aitor Atencia, aitor.atencia@zamg.ac.at

DOI: 10.1175/JAMC-D-16-0344.1

(and, thus, showing a semidiurnal cycle; e.g., Carbone et al. 2002; Carbone and Tuttle 2008; Surcel et al. 2010). Berenguer et al. (2012) studied the mean performance of the McGill Algorithm for Precipitation Nowcasting based on Lagrangian Extrapolation (MAPLE; Germann and Zawadzki 2002) and its dependence on the time of day. They found a relation between MAPLE's skill and the diurnal cycle of precipitation: MAPLE is unable to forecast the initiation of new convection, resulting in its worst performance around 1900 UTC, coinciding with the beginning of thunderstorm development in the foothills of the Rockies; the best performance is achieved at 0100 UTC, when precipitation systems show greater organization and their evolution can be better explained by their motion.

The main goal of this paper is to adaptively correct the nowcasts and VSRFs obtained with a Lagrangian persistence algorithm for the systematic errors introduced by the diurnal cycle of precipitation during the warm season in the east-central United States. The correction has to be adaptive because of the geographical and temporal variability of the diurnal cycle. The main difficulty of the correction is precisely to compensate for the variability in space and time of this diurnal cycle. The study is carried out with 10 yr of radar rainfall maps (section 2), which have been used to characterize the dependence of MAPLE errors on time of day, lead time, spatial scale, and location (section 3). The approach used to correct the biases introduced by the diurnal cycle is presented and tested in section 4; finally, the main results are summarized in section 5.

2. Precipitation data and importance of the diurnal cycle

In this section, detailed information about the dataset used in this study to characterize the diurnal cycle of precipitation is provided (section 2a). The importance of the diurnal cycle has been determined by means of the signal (mean error)-to-noise (standard deviation of errors) ratio. Section 2b examines the significance/robustness of these statistics and their dependence on the length of the dataset.

a. Precipitation observations and nowcasts

A requirement for computing significant and robust statistics from the errors of nowcasting techniques is the use of a large dataset. With this purpose, a dataset previously created and employed in previous works such as Atencia and Zawadzki (2014, 2015) is also used in the present study. This dataset is built by merging two radar-based precipitation composite products over North America; the first one is the National Operational Weather Radar (NOWrad) mosaic produced by the Weather Services (WSI; available from 1998 to 2007), and the second dataset is produced by Weather Decision Technologies (WDT; available from 2004 to 2007).[1] Details about these two products, and the corrections applied to the radar data such as ground clutter removal, reflectivity discretization (5 and 2 dBZ, respectively), etc., can be found in Atencia and Zawadzki (2014) or in the original paper by Zhang et al. (2015).

A common grid has been defined for the two datasets used in this study (Fig. 1). The 512 × 512 points grid has a 4-km mesh length[2] and avoids the mountainous western United States to prevent bad coverage in the study domain. The procedure for resampling the two reflectivity products into the common grid can be found in Atencia and Zawadzki (2014). The temporal resolution of the dataset is 15 min with a length of 10 yr from 1998 to 2007. The month of July has been selected to study the effect of the diurnal cycle because of the stronger influence of solar heating on the diurnal cycle of precipitation in comparison with other months.

The nowcasting algorithm used in this study is MAPLE. This algorithm computes the advection motion field by applying the variational echo tracking (VET) technique developed by Laroche and Zawadzki (1994). The latest available radar image is advected by using a modified semi-Lagrangian backward scheme to keep a quasi-constant power spectrum of the extrapolated images as shown by Germann and Zawadzki (2002). The algorithm has been run simulating operational conditions (i.e., no future data have been used in the computation of the motion field or the reflectivity forecast) to recreate the results that would have been obtained in real time. Nowcasts have been generated every 15 min for a maximum lead time of 10 h and with a temporal resolution of 15 min for the same time period as observations are available (a total of 26 957 valid Lagrangian forecasts, or around 2700 forecasts per month, were obtained).

b. Statistical significance

We have first assessed the importance of the diurnal cycle of precipitation present in the radar observations and in the errors of MAPLE forecasts. To quantify the importance of the diurnal cycle, the mean error as a function of the time of day is compared with the

[1] The data have been obtained directly from these two private companies.

[2] The resampling is carried out in rainfall rate units obtained by using the relation $Z = 300 \times R^{1.5}$ by Joss and Waldvogel (1969) because it produces a bit more of light precipitation and with a minimum reflectivity threshold of 10 dBZ.

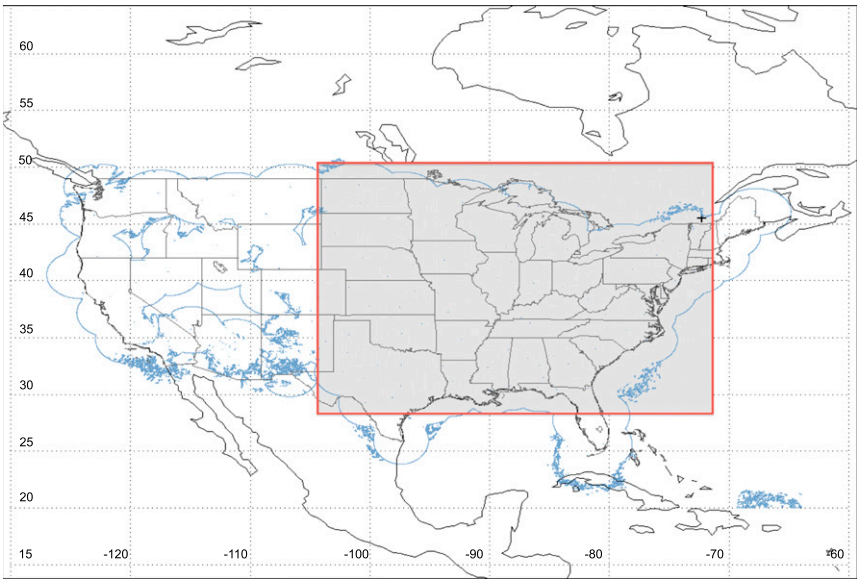
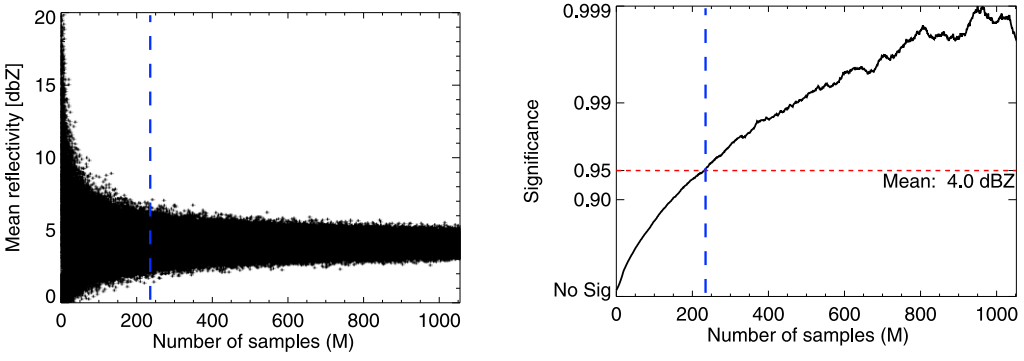


FIG. 1. Location domain. The red rectangle corresponds to the domain on which all the reflectivity fields are smoothed. The blue contours represent the coverage of the reflectivity mosaics.

variability of these errors. For this reason, it is important to test the robustness of the computed moments.

The robustness has been studied by resampling the entire dataset into different samples with different numbers of points by random bootstrapping (with replacement). From this bootstrapped dataset, the mean reflectivity is computed; 1000 different means are computed from 1000 different bootstrapped datasets as a function of the number of points in the range between 5 and 1240 (4 values at a given hour $\times$ 31 days in July $\times$ 10 yr). Figure 2a shows the value of the mean reflectivity obtained (at a random location and hour) as a function of the number of samples used in its computation. For each given sample size, a statistical significance test of the mean is done by comparing the mean computed with the whole dataset with the distribution of means obtained from the different subsamples (Hope 1968). The significance value as a function of the total number of points in the bootstrapped dataset is plotted in Fig. 2b. In this



a) Mean reflectivity for each random run in the MC test. b) Significance of the empirical pdf (MC test) of mean reflectivities.

FIG. 2. Results of the Monte Carlo analysis of the significance of the mean reflectivity obtained in a given point of the study domain as a function of the number of samples used in its computation. The original dataset is all the reflectivity values of July from 1998 to 2007 at a random point ($4 \times 4\,\text{km}^2$) in the domain with a temporal resolution of 15 min. (a) Mean reflectivity computed based on bootstrapping ($N = 1000$ realizations) as a function of the number of samples used (M). From these 1000 realizations, an empirical probability distribution function (pdf) of the mean reflectivity can be obtained for each number of samples. These pdfs are compared with the pdf of the mean reflectivity obtained with $M = 1240$. (b) The significance (Hope 1968) of the two means (number of samples M vs the whole dataset) belonging to the same process as a function of the number of samples used to compute the first mean.

study, a 95% value of significance is chosen to determine the number of points needed for computing the mean. The H_0 hypothesis is that the mean (or standard deviation) of a subset with M samples is equal to the mean (or standard deviation) obtained with the whole dataset. Contrarily, the H_1 hypothesis corresponds to the hypothesis that these two means are different.

By repeating this exercise at different locations, it has been observed that the number of points required to obtain a significance at the 95% level depends on the value of the computed mean. Several locations, and also several times of the day, have been used to compute the minimum number of samples needed for a 95% significant mean. Figure 3 shows the relation between the number of points needed to obtain a statistically significant mean and the computed mean. It can be seen that the minimum number of points needed to obtain a statistically significant mean independently of the value of the mean is 264 (dashed line in Fig. 3). In this study, 10 yr of data for the whole month of July have been used; this corresponds to a total of 1240 values, which exceeds the number of points required for significance. Thus, the mean and standard deviation obtained in this work are statistically significant and can be used to assess the importance of the diurnal cycle.

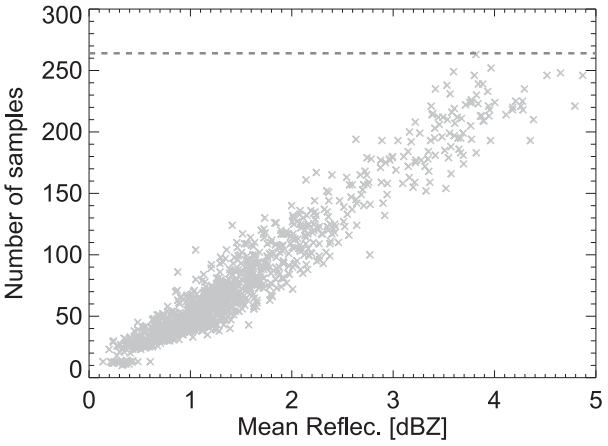


FIG. 3. Number of samples needed to estimate the mean with a significance of 95% vs the value of the mean. These points correspond to a total of 250 random points in random locations and times of the day. A strong correlation can be observed between the number of samples and the mean reflectivity in dBZ. The maximum number of samples needed for significance is 264.

3. Diurnal cycle and errors in MAPLE

For every time step t, MAPLE nowcasted reflectivity fields $\mathbf{x}_f(t, \mathrm{lt})[i, j]$, where $[i, j]$ stands for the spatial location, have been generated with the same resolution as radar observations and for lead times lt up to 10 h. The error in MAPLE forecasts [dB(Z)] is defined[3] as

$$\varepsilon(t, \mathrm{lt})[i, j] = \mathbf{x}_o(t)[i, j] - \mathbf{x}_f(t, \mathrm{lt})[i, j], \qquad (1)$$

where t is the time of day and lt is the lead time of the MAPLE reflectivity forecasts $\{\mathbf{x}_f(t, \mathrm{lt})[i, j]\}$, and the errors are defined by comparison with the observed radar reflectivity field $\{\mathbf{x}_o(t)[i, j]\}$.

These errors have been analyzed as a function of the lead time (lt) by pooling all the times of the day (t) together. Figure 4 shows the bias, defined as the mean (for the dataset) of the errors (averaged over the whole domain

$$\varepsilon(t, \mathrm{lt}) = \sum_i \sum_j \varepsilon(t, \mathrm{lt})[i, j] / (N_i \times N_j)$$

for $i = 1, \ldots,$ to $N_i = 512$ and $j = 1, \ldots,$ to $N_j = 512$), and the root-mean-square error (RMSE). It can be observed that, even though the RMSE increases as a function of the lead time, the bias is very close to 0 for all lead times.

The results are completely different when MAPLE forecast errors are analyzed also as a function of the time of day. The mean error (additive bias as a function of the lead time and time of day) of the whole dataset (Ω) can be defined as

$$\langle \varepsilon(t, \mathrm{lt}) \rangle_\Omega = \langle b(t, \mathrm{lt}) \rangle_\Omega = b(t, \mathrm{lt}). \qquad (2)$$

Then the errors can be decomposed into two terms:

$$\varepsilon(t, \mathrm{lt}) = \hat{\varepsilon}(t, \mathrm{lt}) + b(t, \mathrm{lt}) \quad \text{where} \quad \langle \hat{\varepsilon}(t, \mathrm{lt}) \rangle_\Omega = 0, \qquad (3)$$

where $b(t, \mathrm{lt})$ stands for the additive bias (or mean error) of the forecast and $\hat{\varepsilon}(t, \mathrm{lt})$ is the unbiased error in MAPLE forecasts.

Figure 5 shows the dependence of the mean observed reflectivity, MAPLE bias [$b(t, \mathrm{lt})$], and the standard deviation of the unbiased errors [$\sigma^\Omega_{\hat{\varepsilon}}(t, \mathrm{lt})$] with the time of day. The most noticeable (and also expected) fact is the strong correlation between the MAPLE forecasts' bias (colored solid lines) and the mean observed reflectivity for different times of the day (gray solid line). The timing of the maximum bias occurs at the same time as the maximum mean precipitation in the domain, which is due to the solar heating forcing. Keeping in mind our goal of correcting these biases associated with the diurnal cycle, different questions can arise: Are these averaged errors

[3] Error is usually defined as forecast minus observation, but this definition has been adopted to have a positive correlation between the mean error and the diurnal cycle of precipitation.

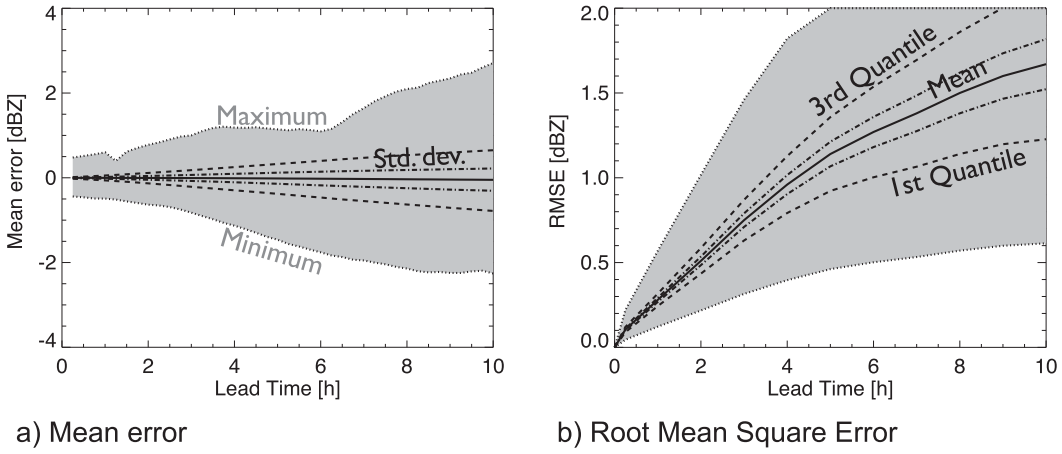


FIG. 4. Errors in MAPLE nowcasts as a function only of the lead time for the whole domain. (a) Mean error over the 10 yr of data (black solid line); the maximum and the minimum are shaded. (b) RMSE as a function of the lead time for the same period over the whole domain. The mean (solid line), standard deviation (dashed–dotted line), and interquartile range (dashed line) are plotted.

significant? Is the standard deviation larger than the range of variation of the bias for short lead times? Do these results show a spatial-scale dependence? Does the importance of the bias depend on the location? In this section, we analyze the dependence of MAPLE errors with spatial scale, location, lead time, and time of day.

a. Spatial scale and lead time

This section analyzes the question of when and for what spatial scales the mean signal of MAPLE errors is larger than its variability. In other words, the standard deviation of the reflectivity errors of MAPLE forecasts for a given scale and lead time measures the variability of the process not associated with the mean diurnal cycle,[4] whereas the amplitude of the mean error (or bias) represents the influence of the diurnal cycle on these errors. A way to measure the strength of the signal is by the signal-to-noise statistic (D_{sn}; Pomeroy et al. 2002), which is defined as the ratio between the difference of the maximum (μ_{max}) and the minimum (μ_{min}) value of the signal and the sum of the standard deviations (σ) at the time the maximum and minimum values of the signal are registered. In our analysis of the MAPLE errors, the signal corresponds to the mean error or bias associated to the diurnal cycle [$b(t, \mathrm{lt})$] and the noise is the unbiased errors [$\hat{\varepsilon}(t, \mathrm{lt})$]. Consequently, signal-to-noise statistic can be written as

$$D_{sn}(\mathrm{lt}) = \frac{\mu_{max} - \mu_{min}}{\sigma_{max} + \sigma_{min}} = \frac{\max_t[b(t,\mathrm{lt})] - \min_t[b(t,\mathrm{lt})]}{\sigma^{\Omega}_{\hat{\varepsilon}}(t = t_{max}, \mathrm{lt}) + \sigma^{\Omega}_{\hat{\varepsilon}}(t = t_{min}, \mathrm{lt})}, \quad (4)$$

where $\max_t[b(t, \mathrm{lt})]$ is the maximum value of the mean error within the time series of a day and $\sigma^{\Omega}_{\hat{\varepsilon}}(t = t_{max}, \mathrm{lt})$ corresponds to the standard deviation of the unbiased errors for the whole dataset at the time of the maximum value of $b(t, \mathrm{lt})$.

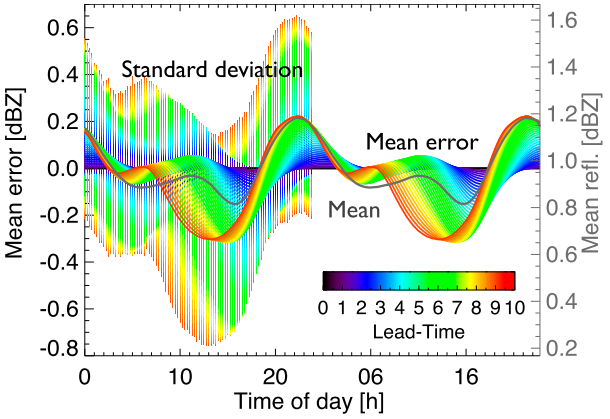


FIG. 5. Mean MAPLE error (averaged over the whole domain) and standard deviation as a function of the time of day. The x axis is doubled (48 h) for an easy visualization. The solid colored lines show the mean errors (different colors represent the lead times from 15 min to 10 h). The time of day represents when the forecast is valid; so a 5-h lead-time forecast at 2000 UTC was initiated at 1500 UTC. The solid dark gray curve represents the mean reflectivity as a function of the time of day. The standard deviations of the unbiased MAPLE errors are plotted as vertical lines for the left half of the figure.

[4] Local effects of the diurnal cycle could be accounted for in the standard deviation.

When D_{sn} is larger than 1, it is worthwhile correcting the bias error associated with the diurnal cycle. On the other hand, if D_{sn} is smaller than 1, the correction of the bias has little influence because the errors of the MAPLE forecast are dominated by other factors. Consequently, the relevance or importance of correcting the bias is associated to the signal-to-noise statistic.

The next step is to analyze for what lead times and for what spatial scales it is relevant to correct for the biases related to the diurnal cycle of precipitation. For this, a relevance analysis has been carried out at pixel level for all the points of the domain $\langle \varepsilon(t, \text{lt})[i, j] \rangle_\Omega$ and for different lead times. The analysis has been done for different spatial scales by low-pass filtering [as in Denis et al. (2002), a squared cosine smooth function is chosen] the observed and forecasted reflectivity fields in the Fourier domain. The reflectivity error field at each lead time is transformed to the frequency space using a discrete cosine transformation (DCT; Denis et al. 2002). The local biases and standard deviations in time have been obtained from the resulting filtered error fields $\{\varepsilon^{f_{\text{FFT}}}(t, \text{lt})[i, j]\}$ to carry out the same analysis as the one described above (beginning of this section) for the unfiltered error fields.

Figure 6 shows D_{sn} associated with the diurnal cycle as a function of cutoff scale and lead time for the whole domain. It can be observed that the importance of the diurnal cycle increases with cutoff scale. However, for the shortest lead time (15 min), the errors do not show significant diurnal cycle signal for any of the scales. For spatial scales around 150 km, the averaged error starts to have a range that is larger than the random component. Consequently, correcting for this bias would improve the root-mean-square error of MAPLE forecasts. For the 30-min lead time, the scale where the diurnal cycle signal is relevant is larger than 256 km, while for the 10-h lead time, the relevance appears at scales around 100 km. Hence, the correction for the diurnal cycle error for scales smaller than 64 km has a small impact and the errors are dominated by the random component, even for lead times of 10 h.

b. Spatial scale and location

The previous results do not take into account the dependence on location of the bias. Significant spatial variability of the diurnal cycle of precipitation was described by Carbone and Tuttle (2008) and Surcel et al. (2010). MAPLE skill depends on spatial scale and location, as was shown by Germann et al. (2006). Here, only the location and scale dependence are shown for the 10-h lead time.

Since our interest is the dependence on location, we will now use wavelet scale decomposition instead

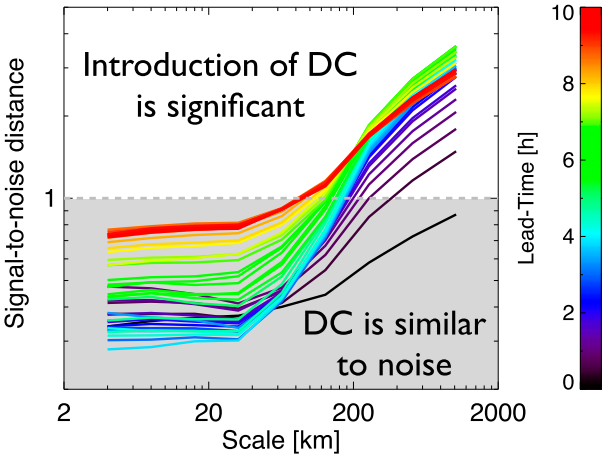
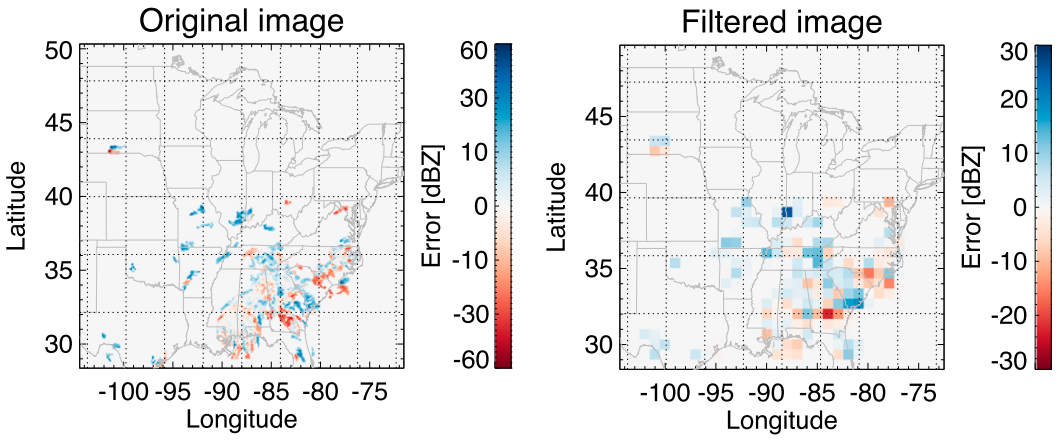


FIG. 6. Analysis of the importance of the introduction of the bias as a function of the spatial scale and lead time. The gray area corresponds to the values of the signal-to-noise statistic [Eq. (4)] < 1. The upper region (white background) represents the value where the signal is larger than the variability so that the correction of the bias (related to the diurnal cycle) is relevant. The different colors of the curves stand for the lead time. The *x* axis is the scale selected in the cutoff filter applied to the original error field.

of Fourier scale analysis. The main difference in general is that wavelets are localized in both location and frequency (wavelength), whereas the standard Fourier transform is only localized in frequency (Addison 2002). The nonredundant Haar wavelet has been selected for this study (Haar 1910). This wavelet in a 2D field is composed by three mother functions (not analyzed) and one scaling function, which is the one applied to perform a low-pass filtering of the field $\{\varepsilon^{f_{\text{WAV}}}(t, \text{lt})[i', j']\}$. Figure 7 shows an example of the original error field $\{\varepsilon\ (t, \text{lt})[i, j]\}$ and the wavelet low-pass filtered field $\{\varepsilon^{f_{\text{WAV}=64\,\text{km}}}(t, \text{lt})[i', j']\}$ for the cutoff scale of 64 km. New coordinates (i' and j') are obtained because the wavelet filtering has been applied in a nonredundant framework to avoid the introduction of small-scale features during the filtering process (Addison 2002). The scaling process carried out using this procedure divides the domain in 1 box centered for the 2048-km scale, 4 boxes for the 1024-km scale, and so on, to a total of 65 536 boxes for the smallest scale. A total of nine levels of this scaling process are obtained (these are equivalent to the scales of 2048, 1024, 512, 256, 128, 64, 32, 16, and 8 km) for each error field of the 10-h lead-time MAPLE forecast.

The importance of the diurnal cycle signal has been analyzed. At each wavelet scaling component, the bias has been computed as a function of the time of day and lead time. From the filtered MAPLE error fields, $D_{sn}^{f_{\text{WAV}}}[i', j']$ has been obtained as a function of



a) Error field (original resolution) for the 1st July 2003 at 1:45. b) Error field (64 km cut-off) for the 1st July 2003 at 1:45.

FIG. 7. Examples of the error (2-h lead time) for 0145 UTC 1 Jul 2003 for the (a) original-resolution (4 km) field, $\varepsilon(t, \text{lt})[i, j]$, and (b) wavelet low-pass filtered (cutoff scale of 64 km) field, $\varepsilon^{f_{\text{WAV}}=64\,\text{km}}(t, \text{lt})[i', j']$.

location (i' and j') and scale (f_{WAV}). The wavelet scaling components where the signal-to-noise statistic is larger than 1 are selected as locations with relevant bias associated with the diurnal cycle for that spatial scale.

Figure 8 shows the result of this analysis for the selected locations at seven levels of spatial scales (the study was carried out at nine levels, but only the seven levels with $D_{\text{sn}}^{f_{\text{WAV}}} > 1$ are plotted in the figure). The first result is that the diurnal cycle for the largest scales (2048 and 1024 km) are significant. These findings are consistent with the scale–lead time results shown in Fig. 6 based on Fourier scale analysis. It can be observed that only the area with a poor radar coverage (around the Great Lakes) shows no significance for the 512-km scale. The southeastern coast shows a signal for the scale of 256 km. The same is observed in the east area of the Rocky Mountains, where there is a location that relevance is found at the 128-km scale (possibly due to thunderstorm initiation in the mountainous area). Finally, Florida and some spots in the Gulf of Mexico show $D_{\text{sn}} > 1$ for a cutoff scale of 64 km (mostly driven by land–sea breeze effects). The dependence with lead time was similar to the observed in the previous section (not shown).

4. Incorporation of the diurnal cycle in MAPLE

To introduce the diurnal cycle in MAPLE, we need to correct for the bias analyzed above. In this section, two approaches are tested. In section 4a, we focus on the potential for error correction associated with the diurnal cycle when the effect is detected a priori. In section 4b, we introduce a technique to predict and correct the presence of the diurnal cycle.

a. Bias correction

To illustrate the potential for correcting the diurnal cycle errors on MAPLE forecasts, a day with a strong diurnal cycle influence has been selected. The scales larger than 512 km for the region of Florida (solid black square located around 30°N and 85°W in Fig. 8) have been selected for carrying out this experiment. The spatially low-pass (wavelet) filtered reflectivity for 25 July 1998 and the low-pass filtered (wavelet) MAPLE forecasts are shown in Fig. 9. It can be observed that the actual radar reflectivity shows minimum precipitation

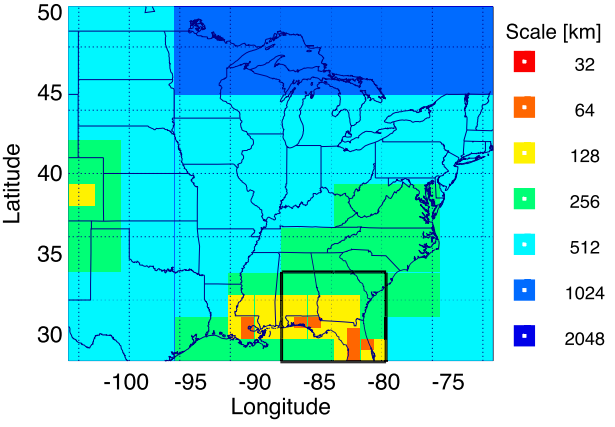


FIG. 8. Analysis of the importance of the introduction of the bias as a function of the scale and location for the 10-h lead time. The shaded rectangles correspond to those scaling components whose values of $D_{\text{sn}}^{f_{\text{WAV}}}[i', j']$ are >1. The different colors of the rectangles stand for the scale filtered by the Haar wavelet scaling function. The relevant locations are represented by the position of the black rectangle over the map. The four boxes of the 1024-km scale have a significant diurnal cycle signal, and for this reason the color of the 2048-km scale cannot be seen, although it is relevant too.

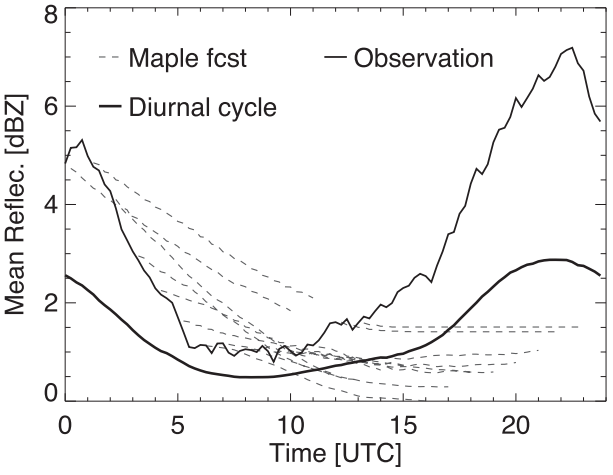


FIG. 9. Example of filtered reflectivity (scale of 512 km) for the event of 25 Jul 1998. The observations (gray thin line) show a strong diurnal cycle when compared with the mean reflectivity for July 1998–2007 averaged as a function of the time of day (black thick line). The MAPLE forecasts are plotted as dashed gray thin lines every hour. An important error in the MAPLE is observed from 1200 UTC until the end of the day. This error is related to the increase of the observed rainfall intensity that is caused by the solar heating.

activity around 1000 UTC (corresponding to the morning) and the precipitation starts developing around noon local time (1700 UTC) and peaks around 2200 UTC. In Fig. 9, it is clear that the two solid lines follow a similar evolution, indicating the diurnal cycle influence. The amplitude is different because the lower solid line is the average for the whole dataset, while the upper one is the reflectivity for a particular case. This event can thus be considered a day with an important influence by the diurnal cycle. MAPLE forecasts do not have any source or sink term[5] in its formulation, so the nowcasted reflectivity stays constant over the domain (the variations shown in Fig. 9 can be attributed to some precipitation leaving the subdomain). In the afternoon, the storms developed are quasi stationary. The VET velocities are small, and consequently the forecasted reflectivity is almost constant. The development of convective storms in the afternoon associated to the solar forcing causes large errors in the MAPLE forecasts.

The errors in MAPLE forecasts can be seen in the bottom left panel of Fig. 10. These errors show some agreement with the average behavior observed in Fig. 5; in other words, the longer the lead time, the larger the spread of the forecast error. This behavior lasts until the second part of the day as well.

The goal of this section is to correct the error caused by the diurnal cycle. The statistical mean of diurnal cycle error [the bias in Eq. (2)] can directly be corrected from the forecasts according to the following equation:

$$x_{\mathrm{cf}}^{f_{\mathrm{WAV}}=512\,\mathrm{km}}(t,\mathrm{lt})[i'=3,j'=1] = x_{f}^{f_{\mathrm{WAV}}=512\,\mathrm{km}}(t,\mathrm{lt})[i'=3,j'=1] + b^{f_{\mathrm{WAV}}=512\,\mathrm{km}}(t,lt)[i'=3,j'=1], \quad (5)$$

where $x_{f}^{f_{\mathrm{WAV}}=512\,\mathrm{km}}(t,\mathrm{lt})[i'=3,j'=1]$ is the MAPLE forecasts over the region of Florida $[i'=3,j'=1]$ filtered to the cutoff scale of 512 km by using the wavelet scaling function; $b^{f_{\mathrm{WAV}}=512\,\mathrm{km}}(t,\mathrm{lt})[i'=3,j'=1]$ represents the MAPLE bias associated to the diurnal cycle for the same location and scale as the MAPLE forecast; and $x_{\mathrm{cf}}^{f_{\mathrm{WAV}}=512\,\mathrm{km}}(t,\mathrm{lt})[i'=3,j'=1]$ stands for the corrected forecasts.

By construction, the formulation of Eq. (5) would suppress the bias when applied over the entire dataset. However, the error of the new forecast for a specific event could increase. To avoid this, an alternative formulation has been applied by introducing an adaptive parameter or weight, $W^{f_{\mathrm{WAV}}=512\,\mathrm{km}}(t,\mathrm{lt})[i'=3,j'=1]$:

$$x_{\mathrm{cf}}^{f_{\mathrm{WAV}}=512\,\mathrm{km}}(t,\mathrm{lt})[i'=3,j'=1] = x_{f}^{f_{\mathrm{WAV}}=512\,\mathrm{km}}(t,\mathrm{lt})[i'=3,j'=1] + W^{f_{\mathrm{WAV}}=512\,\mathrm{km}}(t,\mathrm{lt})[i'=3,j'=1] \times b^{f_{\mathrm{WAV}}=512\,\mathrm{km}}(t,\mathrm{lt})[i'=3,j'=1]. \quad (6)$$

To demonstrate the utility of this approach, the weights have been computed with Eq. (6) assuming the perfect forecast, i.e.

$$x_{\mathrm{cf}}^{f_{\mathrm{WAV}}=512\,\mathrm{km}}(t,\mathrm{lt})[i'=3,j'=1] = x_{o}^{f_{\mathrm{WAV}}=512\,\mathrm{km}}(t)[i'=3,j'=1].$$

The optimal weights obtained for the event of 25 July 1998 are shown in the right panels of Fig. 10. The weights have a larger variability for the longer lead times than for the shorter ones. Note that the variation with time of day disappears in the afternoon. When the solar heating is driving the increase of rainfall intensity, the weights do not depend on the time of day, which explains the overlapping lines in the upper-right panel in Fig. 10. Thus, previously computed weights can be used to correct future MAPLE forecasts. In other words, if in Eq. (6) the time dependence of the weight is removed, that is,

[5] Except for the consistent growth and decay captured by the motion of the precipitation field.

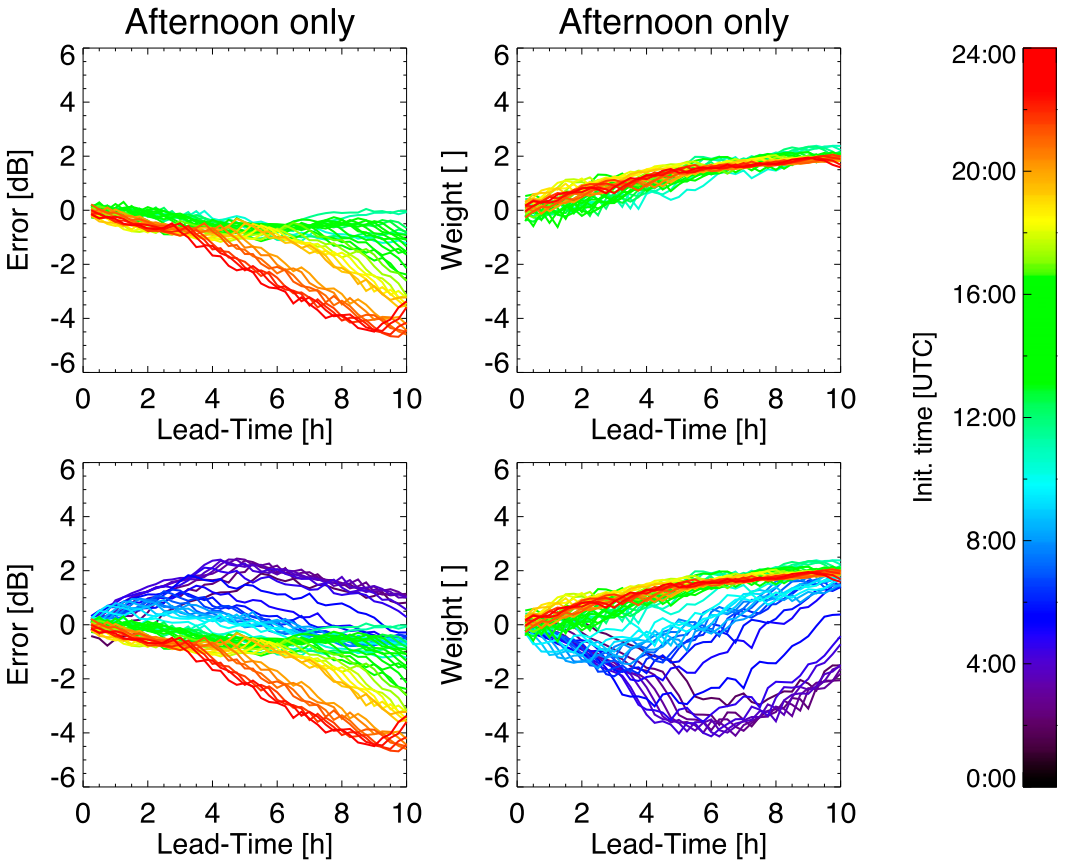


FIG. 10. (bottom left) The errors of the event (25 Jul 1998) plotted in Fig. 9 as function of the lead time for the whole event. The different colors represent the time of initialization of the MAPLE forecast. (bottom right) The weights associated with the correction of this error according to Eq. (6) as a function of the lead time. (top) As in the bottom panels, but for only the afternoon-until-midnight period to emphasize that, even though the errors vary as a function of the initiation time of the MAPLE forecast, the weights are (quasi) constant during the second half of the day.

$$W^{f_{\mathrm{WAV}}}(t,\mathrm{lt})[i',j'] = W^{f_{\mathrm{WAV}}}(\mathrm{lt})[i',j'],$$

then we get an equation to correct future forecasts. It can be formulated in a general way for different scales and location as

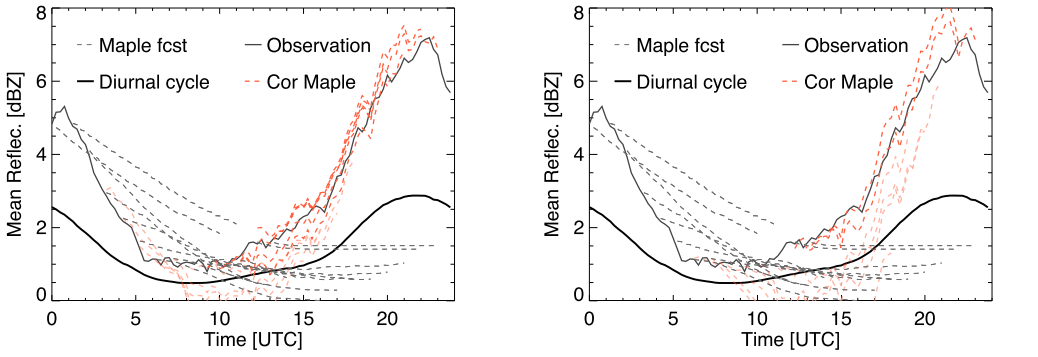
$$x_{\mathrm{cf}}^{f_{\mathrm{WAV}}}(t,\mathrm{lt})[i',j'] = x_{f}^{f_{\mathrm{WAV}}}(t,\mathrm{lt})[i',j'] + W^{f_{\mathrm{WAV}}}(\mathrm{lt})[i',j'] \times b^{f_{\mathrm{WAV}}}(t,\mathrm{lt})[i',j']. \quad (7)$$

Figure 11, left panel, shows the results obtained using the weights computed 3 h before the initialization of MAPLE, and the right one is using the weights obtained 6 h before. The red curves are the corrected forecasts obtained by applying Eq. (7). The improvement of the MAPLE forecasts can clearly be seen in the two panels of Fig. 11 where the corrected forecasts actually follow the observed increase of reflectivity in an event with a strong diurnal heating influence.

These results suggest that it is possible to correct the effect of the diurnal cycle on precipitation nowcasts when and where it is relevant. However, the relevance of this signal is not only dependent on the spatial scale, lead time, and location, as shown in the previous section 3 (and Fig. 8), but it also varies from day to day, as can be observed in Fig. 12. Some days have a strong correlation with the diurnal cycle (maximum correlation around 0.7); other days have a small correlation (lower than 0.1). Only MAPLE forecasts for the days with a strong diurnal cycle component (high correlation) would be improved by removing the bias associated with the diurnal cycle.

b. *Adaptive correction*

In this section, a formulation to adaptively compute the weights in Eq. (7) is proposed. This method has to be



a) Correction with weight computed 3 h. before its application. b) Correction with weight computed 6 h. before its application.

FIG. 11. Mean observed reflectivity (solid thin gray line) as a function of time for the event of 25 Jul 1998. The mean observed reflectivity for the whole dataset (Julys from 1988 to 2008) is plotted as an example of the diurnal cycle of precipitation (solid thick black line) averaged over the Florida box of 512 km × 512 km. The MAPLE forecast is plotted every hour (dashed gray lines). The weights are computed using previous observations. The weights obtained with observation from (a) 3 and (b) 6 h ago are used to correct the MAPLE forecast. The bright/colorful dashed red lines are the corrected forecast for the initialization times where the weights were quasi constant in time (after 1400 UTC); the pale/light dashed red lines are the corrected forecast for the initialization times prior to 1400 UTC.

able to detect when the weights could be used to improve the forecast. As shown above (Fig. 10), this improvement is effective when weights vary little in time. Accordingly, the new weights should account for the temporal variation of the old weights and, at the same time, for the strength of the diurnal cycle. If $W^{f_{\mathrm{WAV}}}(\mathrm{lt})[i',j']$ is consistently large in time, the diurnal cycle is consistent in time, and consequently, the introduction of the correction of the diurnal cycle mean error is improving the forecast. What is the value of $W^{f_{\mathrm{WAV}}}(\mathrm{lt})[i', j']$ that we should retain as a measure of the diurnal cycle? It is a value that has to be at least that of the climatology of the diurnal cycle, and therefore, it depends on the geographical region. Taking into account that the adaptive weight only depends on the lead time $\{W^{f_{\mathrm{WAV}}}(\mathrm{lt})[i',j']\}$, N past observations are used to measure the variability of the weights as an indicator of the robustness of the weights. From Eq. (7), and using N past forecasts for a given lead time (lt), the following system of equations can be obtained:

$$\begin{cases} x_{\mathrm{cf}}^{f_{\mathrm{WAV}}}(t=t_1,\mathrm{lt})[i',j'] = x_f^{f_{\mathrm{WAV}}}(t=t_1,\mathrm{lt})[i',j'] + W^{f_{\mathrm{WAV}}}(\mathrm{lt})[i',j'] \times b^{f_{\mathrm{WAV}}}(t=t_1,\mathrm{lt})[i',j'] \\ x_{\mathrm{cf}}^{f_{\mathrm{WAV}}}(t=t_2,\mathrm{lt})[i',j'] = x_f^{f_{\mathrm{WAV}}}(t=t_2,\mathrm{lt})[i',j'] + W^{f_{\mathrm{WAV}}}(\mathrm{lt})[i',j'] \times b^{f_{\mathrm{WAV}}}(t=t_2,\mathrm{lt})[i',j'] \\ \cdots \\ x_{\mathrm{cf}}^{f_{\mathrm{WAV}}}(t=t_N,\mathrm{lt})[i',j'] = x_f^{f_{\mathrm{WAV}}}(t=t_N,\mathrm{lt})[i',j'] + W^{f_{\mathrm{WAV}}}(\mathrm{lt})[i',j'] \times b^{f_{\mathrm{WAV}}}(t=t_N,\mathrm{lt})[i',j']. \end{cases} \quad (8)$$

To obtain the optimum value of $W^{f_{\mathrm{WAV}}}(\mathrm{lt})[i', j']$, the sum of squared distances, $S\{W^{f_{\mathrm{WAV}}}(\mathrm{lt})[i', j']\}$, between observations, $x_o^{f_{\mathrm{WAV}}}(t = t_k)[i', j']$, and the corrected forecast, $x_{\mathrm{cf}}^{f_{\mathrm{WAV}}}(t = t_k, \mathrm{lt})[i', j']$, is minimized, leaving $W^{f_{\mathrm{WAV}}}(\mathrm{lt})[i', j']$ as the only unknown. This is possible because the corrected forecasts are replaced by the observations in this system of equations. When this replacement is done, future information is used, and it is called analysis mode. This mode would give us the optimal weights and would, therefore, produce the optimal forecast that can be obtained with this technique. The squared distances between observations and corrected forecasts (analysis mode) are computed as

$$S\{W^{f_{\mathrm{WAV}}}(\mathrm{lt})[i',j']\} = \sum_{k=1}^{N} \{x_o^{f_{\mathrm{WAV}}}[i',j'] - x_{\mathrm{cf}}^{f_{\mathrm{WAV}}}(t=t_i,\mathrm{lt})[i',j']\}^2$$

$$= \sum_{k=1}^{N} \left(x_o^{f_{\mathrm{WAV}}}(t=t_k)[i',j'] - \left\{ x_f^{f_{\mathrm{WAV}}}(t=t_k,\mathrm{lt})[i',j'] + W^{f_{\mathrm{WAV}}}(lt)[i',j'] \times b^{f_{\mathrm{WAV}}}(t=t_k,\mathrm{lt})[i',j'] \right\}\right)^2. \quad (9)$$

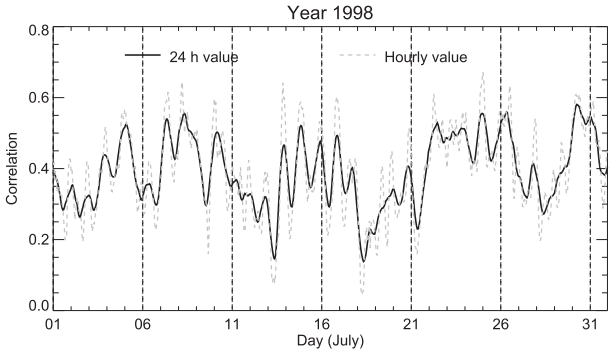


FIG. 12. Correlation between the average rainfall field for each hour (with the 10-yr dataset) associated with the diurnal cycle of precipitation and the observed rainfall field for the whole month of July 1998. To assess the presence of the diurnal cycle signal, this temporal correlation is computed by using 96 (24 h × 4 fields every hour) points of observed rainfall center at every hour for each day vs the diurnal cycle of precipitation. The gray dashed line represents the correlation for a temporal scale of 1 h. To partially avoid the diurnal variation of the correlation, a 24-h filter is applied (black solid line).

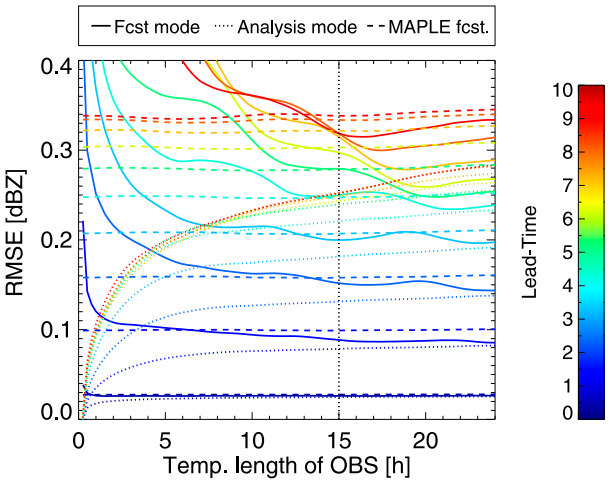


FIG. 13. Root mean $S\{W^{f_{WAV}}(t, \mathrm{lt})[i', j']\}$, Eq. (9), for the whole month of July 1998 as a function of the number of observations N. The dashed lines correspond to $x_{cf}^{f_{WAV}}(t = t_k, \mathrm{lt})[i', j'] = x_f^{f_{WAV}}(t = t_k, \mathrm{lt})[i', j']$, the original MAPLE forecast. The other lines correspond to the corrected forecast in the analysis mode (dotted lines) and the forecast mode (solid lines). The weight obtained from past observations (analysis mode) is used to correct MAPLE forecasts (forecast mode). These lines are averaged for the month of July 1998. The different colors represent the lead times.

The function $S\{W^{f_{WAV}}(\mathrm{lt})[i', j']\}$ is quadratic in $W^{f_{WAV}}(\mathrm{lt})[i', j']$ with a positive-definite Hessian, and therefore, this function possesses a unique global minimum at $W^{f_{WAV}}(\mathrm{lt})[i', j'] = \widehat{W}^{f_{WAV}}(\mathrm{lt})[i', j']$ which can be retrieved with the explicit equation:

$$\widehat{W}^{f_{WAV}}(\mathrm{lt})[i', j'] = \frac{\sum_{k=1}^{N} \left\langle b^{f_{WAV}}(t = t_k, \mathrm{lt})[i', j'] \times \{x_o^{f_{WAV}}(t = t_k)[i', j'] - x_f^{f_{WAV}}(t = t_k, \mathrm{lt})[i', j']\} \right\rangle}{\sum_{k=1}^{N} \{b^{f_{WAV}}(t = t_k, \mathrm{lt})[i', j']\}^2} \tag{10}$$

The value of $\widehat{W}^{f_{WAV}}(\mathrm{lt})[i', j']$, for a given location, lead time, and spatial scale, has a dependence on the number of observations (N) used in its computation. For a single observation ($N = 1$), we obtain the same weight as in the previous section. The introduction of more observations (larger N) provides information about the variability of the weights in the system of equations. This variability can be related to the strength of the diurnal cycle signal for a given day; as seen in Fig. 10, in the presence of a strong diurnal cycle, these weights remained constant. The weight obtained with Eq. (10) minimizes the RMSE in analysis mode. The weight obtained for a given time step (analysis mode) is used for correcting the next forecast. It is possible to compute these forecast in a real-time operational framework, and for this reason, this is called forecasting mode. The difference between the RMSE obtained in analysis and forecasting mode is used to find the number of observations N that optimizes this correction.

Figure 13 shows the RMSE obtained in analysis and forecasting modes and the RMSE of the MAPLE forecast for several lead times over the whole domain (i.e., for a scale of 2048 km, the root-mean $S\{W^{f_{WAV}=2048\,\mathrm{km}}(\mathrm{lt})[i' = 1, j' = 1] = W(\mathrm{lt})\}$). In analysis mode (dotted line), the RMSE increases as more observations from the past are used. However, in the forecast mode, as more observations are used, RMSE decreases. The goal of this test is to determine the number of observations needed for a forecast to be better than MAPLE. At the same time, it introduces a correction weight when the diurnal cycle bias in MAPLE forecasts is important. For this reason, the use of few observations is enough for shorter lead times, while

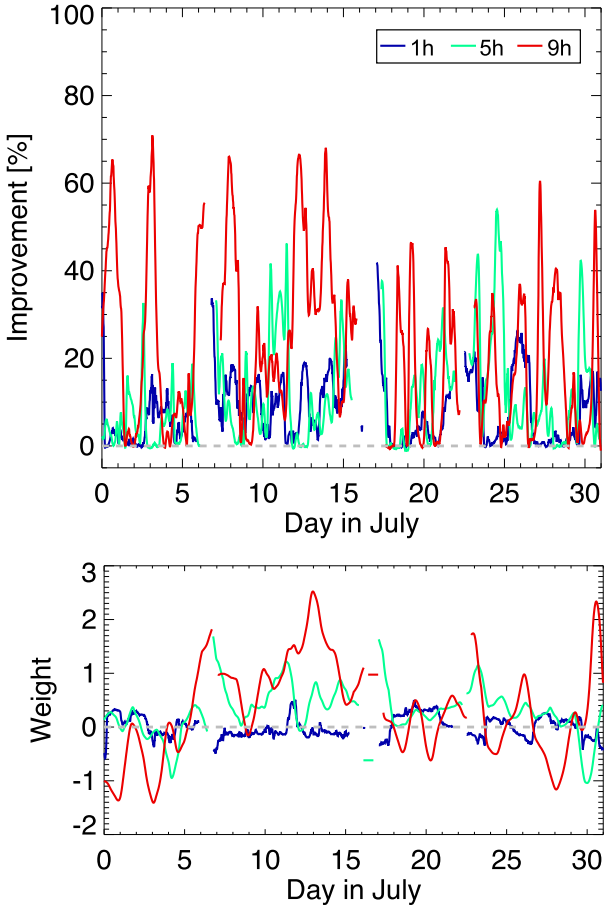


FIG. 14. (top) The skill score of the RMSE (S_{RMSE}) of the corrected forecast as a function of the day of the month of July 1998 for the whole domain (scale: 2048 km). Three lines are plotted corresponding to three different lead times (1 h: blue line; 5 h: light green line; and 9 h: red line). (bottom) The weights obtained from Eq. (10) used to correct the forecast, Eq. (7).

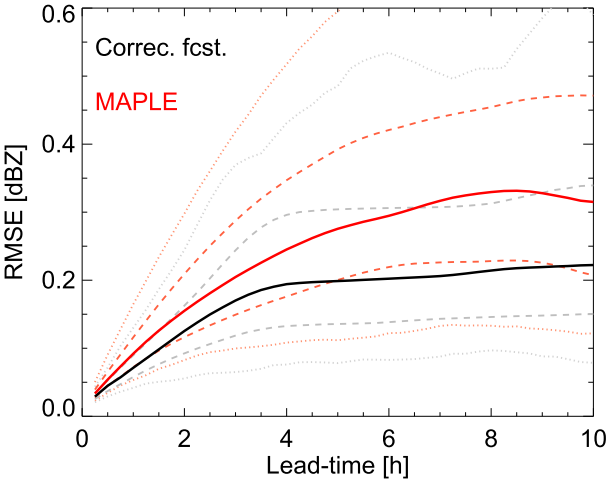


FIG. 15. RMSE as a function of the lead time for the original MAPLE forecast (red lines) and the corrected version (black lines) for the whole domain (scale: 2048 km). The solid lines are for the pooled RMSE during the whole month of July 1998. The dashed lines are the standard deviation, and the dotted lines are ±90% of all of the RMSE obtained during the month.

for longer lead times, more observations are required. Finally, 15 h of data have been selected for computing the weights for all the lead times. This technique will affect the forecast when the diurnal cycle is well above the remaining precipitation noise, and it will not affect the forecast otherwise. With this number of observations, the improvement is noticeable for the events with a strong diurnal cycle (independently of the lead time), and the effect when this diurnal cycle signal is not important is mitigated. Let us introduce the skill RMSE $[S_{\text{RMSE}}(t,\ \text{lt})]$ to compare the original MAPLE forecast with the corrected forecast obtained with Eq. (7) using the weights computed with Eq. (10). This score accounts for the percentage of improvement in the RMSE:

$$S^{f_{\text{WAV}}}_{\text{RMSE}}(t,\text{lt})[i',j'] = 100 \times \left[\frac{\text{RMSE}^{f_{\text{WAV}}}_{f}(t,\text{lt})[i',j'] - \text{RMSE}^{f_{\text{WAV}}}_{\text{cf}}(t,\text{lt})[i',j']}{\text{RMSE}^{f_{\text{WAV}}}_{f}(t,\text{lt})[i',j']}\right](\%), \tag{11}$$

or for the scale of 2048 km that corresponds to the whole domain, the equation can be simplified to

$$S_{\text{RMSE}}(t,\text{lt}) = 100 \times \left[\frac{\text{RMSE}_{f}(t,\text{lt}) - \text{RMSE}_{\text{cf}}(t,\text{lt})}{\text{RMSE}_{f}(t,\text{lt})}\right](\%), \tag{12}$$

where RMSE_f and RMSE_{cf} are the root-mean-square errors of the original forecast and of the corrected forecast (forecast mode), respectively. The adaptive behavior can be observed in Fig. 14, where the skill of the RMSE is plotted as a function of day in July of 1998. The skill score has a strong dependence on the lead time and shows different behavior for different events. It is important to highlight that in the worst case (i.e., when the evolution of the precipitation is not explained by the diurnal cycle), the improvement is around 0%; that is, the MAPLE forecast is not modified (this can be

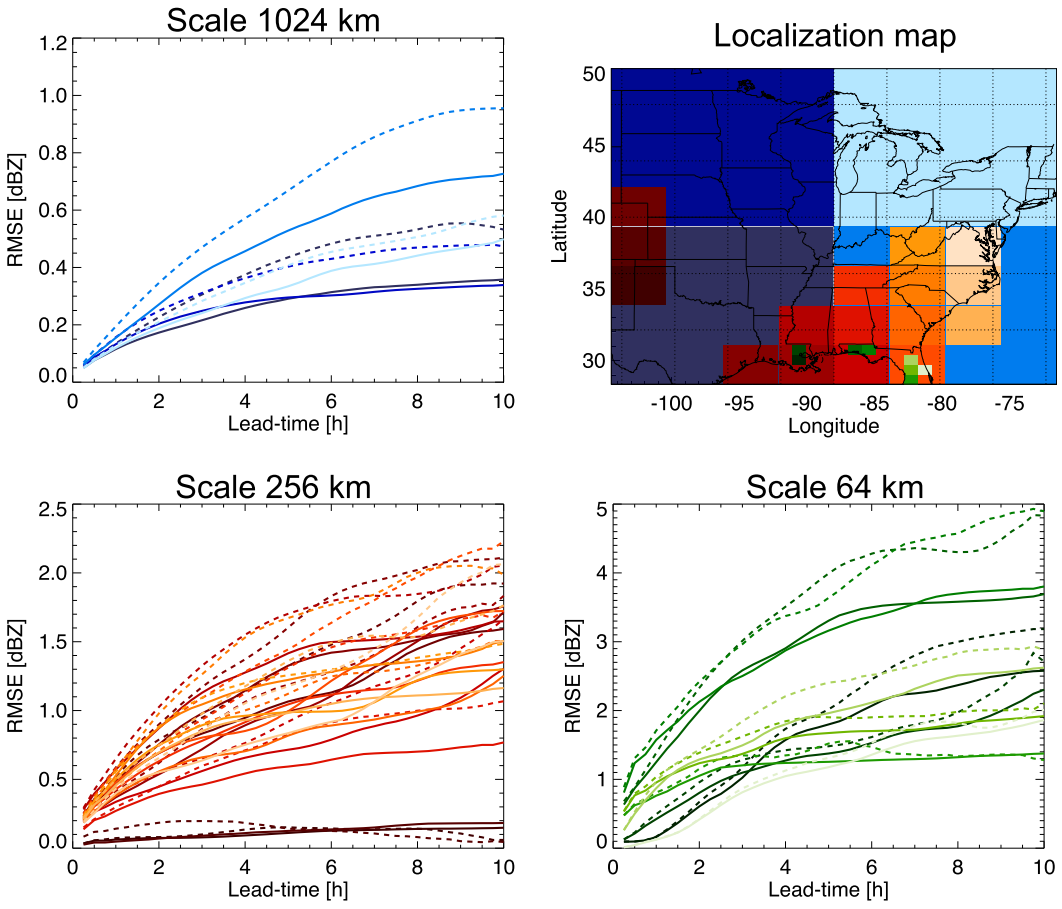


FIG. 16. (top right) The locations over the continental United States of the lines in the figure's other three panels. The different colors stand for different spatial scales, and the different shades of color distinguish the different locations. The RMSE as a function of the lead time for the original MAPLE forecast (dashed lines) and the corrected version (solid lines) for three different scales: (top left) 1024, (bottom left) 256, and (bottom right) 64 km.

observed in the lower panel of Fig. 14 where the evolution of the weight is plotted). This is one of the properties of the minimization process using 15 h of previous observations. The improvement is smaller than that obtained in section 4a for a single event, with a known strong influence of the diurnal cycle or strong correlation with the bias. The adaptive technique improves the forecast only when the diurnal cycle signal is observed in the error in past observations. Figure 15 summarizes the results over the entire domain of MAPLE forecasts for the month of July 1998.

A figure similar to Fig. 15 can be obtained at each location and scale where the diurnal cycle signal is significant (see section 3b). The range of values for $\mathrm{RMSE}^{f_{\mathrm{WAV}}}(t, \mathrm{lt})[i', j']$ is different for each scale but also among the different locations, as can be observed in Fig. 16. An improvement is obtained for almost all the locations and scales (only a few are shown in Fig. 16). It can also be observed as the averaged reflectivities are dependent on the scale analyzed reaching values around 5 dBZ for the smallest significant spatial scale. The upper-left panel shows the pooled RMSE for the four subdomains of 1024 km. The northeast location has the smallest improvement [comparing the corrected forecast (solid line) with the original forecast (dashed line)], which can be explained by the fact that it is an area with a small D_{sn} in comparison with the southeast location with a larger D_{sn}. The other fact to remark from this location-scale analysis is that the improvement is not present for all the locations; for instance, at 256-km cutoff scale, the corrected forecast of the southern subdomain in the Rocky Mountains has a worse performance than the original MAPLE forecast. This can be caused by the importance of the transportation over the growth and decay; this effect has been observed in previous studies (Berenguer et al. 2012).

To summarize these results, the average improvement in the RMSE for each scale and the range of maximum

TABLE 1. A summary of the results observed in Fig. 16. The first column represents the spatial scale. The second column is the average for all the significant locations of the maximum skill score of RMSE $S\{W^{f_{WAV}}_{RMSE}(t, lt)[i', j']\}$. The range of maximum $S^{f_{WAV}}_{RMSE}(t, lt)[i', j']$ is also introduced. The average lead time at which this maximum $S^{f_{WAV}}_{RMSE}(t, lt)[i', j']$ happens is in the fifth column. The minimum skill score for any location at a given scale from the 2-h lead time until the end of the forecasting period (10-h lead time) is in the last column.

Scale (km)	Mean S_{RMSE} (%)	Range S_{RMSE}		Averaged lt (h)	Min S_{RMSE} (%)
		Min	Max		
1024	27.74	17.94	37.13	7.25	5.26
512	29.19	18.87	39.90	7	2.37
256	29.76	16.99	64.86	5.75	−51.32
128	23.29	12.93	32.84	6.75	−7.14
64	19.68	13.25	24.43	6.5	−6.87

improvement among the different locations is introduced in Table 1. The average lead time of the maximum improvement is provided too. The last column shows the worst performance (minimum skill RMSE) for the corrected forecast. This minimum skill RMSE is obtained from the second hour of lead time until the end of the forecast period (10 h) to avoid having low values from the first hours when the Lagrangian persistence forecast is really accurate, and this diurnal cycle correction introduces small corrections. For the higher scales (1024 and 512 km), there is always an improvement and the worst performance happens in the second hour of forecast, which is caused because the Lagrangian skill is still good and the correction has a less important effect. For smaller scales, the minimum skill score is negative (worsens the forecast). This happens in latter lead times (around 9–10 h). The worse performance is at the scale of 256 km, and a negative skill score is obtained only in one location, but it lasts from 6 h of lead time to the end. A negative skill score is obtained in the same location for the scales of 128 and 64 km. This degradation is only happening for the last 45 min of forecast. The rest of the scales and locations are showing always an improvement in the skill RMSE, proving the positive effect of introducing the diurnal cycle mean error correction.

5. Conclusions

The errors in MAPLE rainfall forecasts for the month of July between 1998 and 2008 have been studied as a function of the lead time and the time of day. The objective of introducing the time of day in this analysis was to illustrate the impact of the diurnal cycle of precipitation into MAPLE errors. We have developed a methodology to adaptively correct these nowcasts for the effects of the diurnal cycle.

The first part of the paper focuses on the spatial scales and locations for which forecast errors were mainly associated with a diurnal signal. As expected from previous published results, the diurnal cycle of precipitation was significant for scales of 128 km in the Colorado plains and even 64 km for the Florida region. In the rest of the domain, the diurnal signal is still detectable over other possible causes of forecast error (a signal-to-noise ratio larger than 1), but it specially affects the larger scales of the rainfall field.

The proposed approach is an adaptive method that determines the weight of the diurnal cycle correction of MAPLE based on the observations from the previous hours. A small variability of this correction (quasi-constant weights) indicates a strong diurnal cycle, and, as a consequence, MAPLE forecasts need a strong bias correction. On the other hand, a large variability was associated with forecast errors due to other causes, and the diurnal cycle bias correction is minimal, keeping almost intact the MAPLE forecast. The results indicate that the adaptive technique introduces improvement up to 60% of the RMSE for some events, whereas the negative impact, for events where the errors were not associated with the diurnal cycle of precipitation, was no larger than −2%. This adaptive approach has been tested for all locations and spatial scales where the signal-to-noise ratio was larger than 1, and the results are similar to the ones obtained for the entire domain. On average, the corrections improve the RMSE of the forecast up to 30% for lead times up to 8 h during the month of July 1998.

Zhang and Klein (2010) studied the mechanisms that drive the transition from shallow to deep convection. In their study, they found a temporally delayed high correlation between the diurnal cycle of precipitation and the structure (in height) of clouds. This delayed correlation suggests the use of cloud coverage, structure, or other weather variables (such as the moisture at the lower free troposphere) to detect the diurnal cycle ahead of time. The application of these predictors to assess the strength of the diurnal cycle could lead to the use of fewer rainfall observations to compute the weights, which, at the same time, could result in improved forecasts.

However, these corrections were only applied for selected scales and locations. A technique to downscale these corrections to the original image resolution would improve the applicability of this study for other meteorological or hydrological purposes such as storm forecasting or flash-flood forecasting. Previous studies have shown that the stochastic perturbation of some statistical parameters of rainfall images, such as the mean reflectivity (e.g., Berenguer et al. 2011), can be used to generate different ensemble members for a probabilistic nowcasting of precipitation. Our results suggest that the application of a correction instead of a

stochastic perturbation of these parameters could be beneficial for the locations and scales where the diurnal cycle signal is significant.

Acknowledgments. The authors express their gratitude to the Global Hydrology and Climate Center (GHRC) for providing access to the WSI radar composites data. Special thanks go to all our group members for their fruitful discussions during the group meetings and especially to Professor Frédéric Fabry for his suggestions about nowcasting biases and the computation of errors in reflectivity fields.The contribution to this work of Dr. Marc Berenguer was supported in the framework of the Spanish Project FFHazF (CGL2014-60700R).

REFERENCES

Addison, P. S., 2002: *The Illustrated Wavelet Transform Handbook: Introductory Theory and Applications in Science, Engineering, Medicine and Finance.* CRC Press, 368 pp.

Atencia, A., and I. Zawadzki, 2014: A comparison of two techniques for generating nowcasting ensembles. Part I: Lagrangian ensemble technique. *Mon. Wea. Rev.*, **142**, 4036–4052, doi:10.1175/MWR-D-13-00117.1.

——, and ——, 2015: A comparison of two techniques for generating nowcasting ensembles. Part II: Analogs selection and comparison of techniques. *Mon. Wea. Rev.*, **143**, 2890–2908, doi:10.1175/MWR-D-14-00342.1.

Berenguer, M., C. Corral, R. Sánchez-Diezma, and D. Sempere-Torres, 2005: Hydrological validation of a radar-based nowcasting technique. *J. Hydrometeor.*, **6**, 532–549, doi:10.1175/JHM433.1.

——, D. Sempere-Torres, and G. G. Pegram, 2011: SBMcast—An ensemble nowcasting technique to assess the uncertainty in rainfall forecasts by Lagrangian extrapolation. *J. Hydrol.*, **404**, 226–240, doi:10.1016/j.jhydrol.2011.04.033.

——, M. Surcel, I. Zawadzki, M. Xue, and F. Kong, 2012: The diurnal cycle of precipitation from continental radar mosaics and numerical weather prediction models. Part II: Intercomparison among numerical models and with nowcasting. *Mon. Wea. Rev.*, **140**, 2689–2705, doi:10.1175/MWR-D-11-00181.1.

Bowler, N. E., C. E. Pierce, and A. W. Seed, 2006: STEPS: A probabilistic precipitation forecasting scheme which merges an extrapolation nowcast with downscaled NWP. *Quart. J. Roy. Meteor. Soc.*, **132**, 2127–2155, doi:10.1256/qj.04.100.

Carbone, R., and J. Tuttle, 2008: Rainfall occurrence in the U.S. warm season: The diurnal cycle. *J. Climate*, **21**, 4132–4146, doi:10.1175/2008JCLI2275.1.

——, ——, D. Ahijevych, and S. Trier, 2002: Inferences of predictability associated with warm season precipitation episodes. *J. Atmos. Sci.*, **59**, 2033–2056, doi:10.1175/1520-0469(2002)059<2033:IOPAWW>2.0.CO;2.

Dai, A., F. Giorgi, and K. E. Trenberth, 1999: Observed and model-simulated diurnal cycles of precipitation over the contiguous United States. *J. Geophys. Res.*, **104**, 6377–6402, doi:10.1029/98JD02720.

Denis, B., J. Côté, and R. Laprise, 2002: Spectral decomposition of two-dimensional atmospheric fields on limited-area domains using the discrete cosine transform (DCT). *Mon. Wea. Rev.*, **130**, 1812–1829, doi:10.1175/1520-0493(2002)130<1812:SDOTDA>2.0.CO;2.

Germann, U., and I. Zawadzki, 2002: Scale-dependence of the predictability of precipitation from continental radar images. Part I: Description of the methodology. *Mon. Wea. Rev.*, **130**, 2859–2873, doi:10.1175/1520-0493(2002)130<2859:SDOTPO>2.0.CO;2.

——, ——, and B. Turner, 2006: Predictability of precipitation from continental radar images. Part IV: Limits to prediction. *J. Atmos. Sci.*, **63**, 2092–2108, doi:10.1175/JAS3735.1.

Haar, A., 1910: Zur theorie der orthogonalen funktionensysteme. *Math. Ann.*, **69**, 331–371, doi:10.1007/BF01456326.

Hope, A. C., 1968: A simplified Monte Carlo significance test procedure. *J. Roy. Stat. Soc.*, **30B**, 582–598.

Janowiak, J. E., V. E. Kousky, and R. J. Joyce, 2005: Diurnal cycle of precipitation determined from the CMORPH high spatial and temporal resolution global precipitation analyses. *J. Geophys. Res.*, **110**, D23105, doi:10.1029/2005JD006156.

Joss, J., and A. Waldvogel, 1969: Raindrop size distribution and sampling size errors. *J. Atmos. Sci.*, **26**, 566–569, doi:10.1175/1520-0469(1969)026<0566:RSDASS>2.0.CO;2.

Laroche, S., and I. Zawadzki, 1994: A variational analysis method for retrieval of three-dimensional wind field from single-Doppler radar data. *J. Atmos. Sci.*, **51**, 2664–2682, doi:10.1175/1520-0469(1994)051<2664:AVAMFR>2.0.CO;2.

Parker, M. D., and D. A. Ahijevych, 2007: Convective episodes in the east-central United States. *Mon. Wea. Rev.*, **135**, 3707–3727, doi:10.1175/2007MWR2098.1.

Pomeroy, S., and Coauthors, 2002: Prediction of central nervous system embryonal tumour outcome based on gene expression. *Nature*, **415**, 436–442, doi:10.1038/415436a.

Seed, A., 2003: A dynamic and spatial scaling approach to advection forecasting. *J. Appl. Meteor.*, **42**, 381–388, doi:10.1175/1520-0450(2003)042<0381:ADASSA>2.0.CO;2.

Surcel, M., M. Berenguer, and I. Zawadzki, 2010: The diurnal cycle of precipitation from continental radar mosaics and numerical weather prediction models. Part I: Methodology and seasonal comparison. *Mon. Wea. Rev.*, **138**, 3084–3106, doi:10.1175/2010MWR3125.1.

Wallace, J. M., 1975: Diurnal variations in precipitation and thunderstorm frequency over the conterminous United States. *Mon. Wea. Rev.*, **103**, 406–419, doi:10.1175/1520-0493(1975)103<0406:DVIPAT>2.0.CO;2.

Zhang, Y., and S. A. Klein, 2010: Mechanisms affecting the transition from shallow to deep convection over land: Inferences from observations of the diurnal cycle collected at the ARM Southern Great Plains site. *J. Atmos. Sci.*, **67**, 2943–2959, doi:10.1175/2010JAS3366.1.

——, D. Kitzmiller, D.-J. Seo, D. Kim, and R. Cifelli, 2015: Creation of multisensor precipitation products from WSI NOWrad reflectivity data. *J. Hydrol. Eng.*, **22**, E4015001, doi:10.1061/(ASCE)HE.1943-5584.0001216.