Using Precipitation Observations in a Mesoscale Short-Range Ensemble Analysis and Forecasting System

TADASHI FUJITA
NOAA/National Severe Storms Laboratory, and Sasaki Institute, University of Oklahoma, Norman, Oklahoma, and Numerical Prediction Division, Japan Meteorological Agency, Tokyo, Japan

DAVID J. STENSRUD
NOAA/National Severe Storms Laboratory, Norman, Oklahoma

DAVID C. DOWELL
Cooperative Institute for Mesoscale Meteorological Studies, University of Oklahoma, and NOAA/National Severe Storms Laboratory, Norman, Oklahoma

(Manuscript received 27 November 2006, in final form 1 November 2007)

ABSTRACT

A simple method to assimilate precipitation data from a synthesis of radar and gauge data is developed to operate alongside an ensemble Kalman filter that assimilates hourly surface observations. The mesoscale ensemble forecast system consists of 25 members with 30-km grid spacing and incorporates variability in both initial and boundary conditions and model physical process schemes. The precipitation assimilation method only incorporates information on when and where rainfall is observed. Model temperature and water vapor mixing ratio profiles at each grid point are modified if rainfall is observed but not predicted, or if rainfall is predicted but not observed. These modifications act to either increase or decrease, respectively, the likelihood that precipitation develops at that grid point.

Two cases are examined in which this technique is applied to assimilate precipitation data every 15 min from 1200 to 1800 UTC, while hourly surface observations are also assimilated at the same time using the more sophisticated ensemble Kalman filter approach. Results show that the simple method for assimilating precipitation data helps the model develop precipitation where it is observed, resulting in the precipitation area being reproduced more accurately than in the run without precipitation-data assimilation, while not negatively influencing the positive results from the surface data assimilation. Improvement is also seen in the reliability of precipitation probabilities for a 1 mm h⁻¹ threshold after the assimilation period, indicating that assimilating precipitation data may provide improved forecasts of the mesoscale environment for a few hours.

1. Introduction

In recent years, increasing attention has turned toward improving quantitative precipitation forecasts (QPFs). These forecasts are consistently the poorest performance area of numerical weather prediction (Fritsch and Carbone 2004), yet they have a direct impact on the risk management of a large section of society. Fritsch and Carbone (2004) strongly suggest that the probabilistic prediction of QPF is the most likely approach to improving precipitation forecasts and that ensemble forecasts are one method of obtaining this probability information. It also has been recognized that in order to obtain an improved QPF, models need to be initialized with the most accurate possible description of the atmosphere. Thus, incorporating information from observations directly related to ongoing precipitation into the model initial condition is vital to forecasting precipitation, because without this information precipitation formation often is delayed and forecast utility lessened (Wang and Warner 1988).
Because mesoscale and global-scale models employ a convective parameterization scheme, several approaches have been proposed to adjust the model fields that influence precipitation. In a physical initialization approach (Krishnamurti et al. 1984, 1991, 1993, 1994; Donner 1988; Treadon 1996), the moisture, temperature, and sometimes even the vertical velocity profiles are adjusted until the precipitation rates from the convective scheme and observations are in agreement. Physical initialization both reduces the spinup problem substantially and improves the resulting precipitation forecasts. In a latent heat nudging approach (Ninomiya and Kurihara 1987; Wang and Warner 1988; Manobianco et al. 1994; Chang and Holt 1994; Jones and Macpherson 1997; Cartwright and Ray 1999; MacPherson 2001; Lin et al. 2001), model-predicted latent heating profiles are scaled to reflect the total heating implied by the observed precipitation. This modification of the heating profile acts as a source term in the model thermodynamic equation, inducing adjustments to the divergent circulation and vertical velocity field. This method has been found to be beneficial in midlatitudes and has a noticeable positive impact on rainfall forecasts. Rogers et al. (2000) use radar data to describe only where and when deep moist convection is occurring in the model domain by activating the model convective parameterization scheme in grid elements where convection is indicated. In this approach, the convective heating profiles and precipitation rates are compatible with the local environment instead of the observed precipitation rates. Results indicate that this approach also yields improved precipitation forecasts and better forecasts of the mesoscale environment. There also have been attempts to directly assimilate precipitation data by four-dimensional variational assimilation techniques (Zupanski and Mesinger 1995; Zou and Kuo 1996; Guo et al. 2000; Ishikawa 2002), showing positive impacts on the subsequent numerical forecasts of precipitation.

The recognition that precipitation forecasts are especially sensitive to initial condition uncertainty (Krishnamurti et al. 1994; Karyampudi et al. 1998; Ramamurthy and Jewett 1999) further suggests that precipitation forecasting may be profitably addressed in a probabilistic framework by using forecast ensembles as discussed by Fritsch and Carbone (2004). Ensembles most often incorporate uncertainty in the atmospheric initial state, although many ensembles also incorporate uncertainty in the model physical process schemes because the resulting forecasts are improved (Atger 1999; Buizza et al. 1999; Harrison et al. 1999; Evans et al. 2000; Stensrud et al. 2000; Ziehmann 2000; Fritsch et al. 2000; Hou et al. 2001; Wandishin et al. 2001; Alhamed et al. 2002). Stensrud et al. (2000) argue that diversity in the model physical process schemes is especially important for short-range ensembles, because the physics schemes play a large role in the evolution of the sensible weather events. One method of assimilating observations into an ensemble is the ensemble Kalman filter (Evensen 1994). In this approach, the statistical properties of the ensemble itself are used to determine the adjustment of the model fields by the observational information. This technique has been applied successfully in a number of different contexts in recent years (Evensen 1994, 1997; Houtekamer and Mitchell 1998; Miller et al. 1999; Anderson and Anderson 1999; Mitchell and Houtekamer 2000; Hamill and Snyder 2000; Anderson 2001; Mitchell et al. 2002; Whitaker and Hamill 2002; Snyder and Zhang 2003; Dowell et al. 2004; Zhang et al. 2004; Zhang 2005; Zhang et al. 2006). When the precipitation processes are explicitly resolved on the model grid, as is true for model grid spacings of ∼2 km or less, the ensemble Kalman filter is a very good assimilation method (Snyder and Zhang 2003; Dowell et al. 2004; Tong and Xue 2005). However, it is not clear that the ensemble Kalman filter is the best option to assimilate precipitation data when the model uses a convective parameterization scheme and precipitation is largely a subgrid-scale process. The usefulness of the covariances between precipitation and the model state variables in the presence of model error is questionable, leading to concerns about using an ensemble Kalman filter for assimilating subgrid-scale precipitation observations. Thus, for this initial study, a simple assimilation method for precipitation data is developed that operates alongside an ensemble Kalman filter as applied in a short-range mesoscale ensemble forecasting system. This method also is developed to operate externally to the numerical model integration to make it more amenable for use in ensembles with diversity in convective and microphysics schemes. Comparisons between forecasts with and without precipitation-data assimilation are compared to determine the benefits of this simple assimilation approach to improved numerical forecasts.

All of the experiments assimilate hourly surface observations over a 6-h assimilation period using the ensemble Kalman filter as in Fujita et al. (2007) in order to improve the predicted lower-tropospheric environment in which convection develops. Hourly precipitation estimates from a synthesis of radar and gauge observations are assimilated over the same 6-h period in parallel with the assimilation of surface data using a simple adjustment method based on the Rogers et al. (2000) approach. It is hoped that by combining the assimilation of surface observations and radar-derived
precipitation estimates the ensemble initial states can be specified more accurately, yielding improved probabilistic forecasts.

It is important to emphasize that the assimilation of the precipitation data is not accomplished by using the ensemble Kalman filter. Rather, each ensemble member is modified using a simple adjustment method that depends to some extent upon the convective parameterization scheme being used by each ensemble member. Similar to the scheme proposed by Rogers et al. (2000), the precipitation data are only used to inform the model when and where convection is occurring. This method is based on studies of mesoscale numerical models that indicate that the evolution of mesoscale convective systems and their environments are sensitive to the timing and location of the initial convection (e.g., Fritsch and Chappell 1981; Zhang and Fritsch 1986; Kain and Fritsch 1992; Stensrud and Fritsch 1994; Zhang and Harvey 1995; Rogers and Fritsch 1996).

The assimilation procedures are discussed in the following section. The experimental design is described in section 3. The results obtained from the experiment are discussed in section 4, followed by a summary in section 5.

2. Assimilation procedure

Surface and precipitation observations are assimilated only during the first 6 h of the experimental period: from 1200 to 1800 UTC (hereafter, the assimilation period). During this period, the assimilation of the observational data (hereafter, the analysis step) and the ensemble forecasts between two successive analysis steps (hereafter, the forecast step) are repeated one after the other. The analysis step for the surface observation takes place every hour, while that for the precipitation observation is performed every 15 min because hourly precipitation assimilation was found to be much less effective. When both of the observations are assimilated in a single analysis step, the assimilation of the precipitation observation is performed prior to that of the surface observation. By using this ordering, the surface observations constrain any modifications to the low-level model fields. Detailed descriptions of the analysis step for both of the observation types follow.

a. Surface observations

As in Fujita et al. (2007), surface observations of potential temperature, dewpoint temperature, and $u$ and $v$ wind components are assimilated every hour using the framework of the ensemble square root filter (Whitaker and Hamill 2002). The observations are distributed over the land area of the experimental domain at approximately 1500–1600 routine surface observation stations. The observational error standard deviations used in the ensemble square root filter are 2.0 K for potential temperature, 2.0 m s$^{-1}$ for the $u$ and $v$ components of the wind, and 2.0 K for dewpoint temperature (Zapotocny et al. 2000). The half-radius of the localization factor in the filter is taken to be five grid points in the horizontal direction (150 km) and 10 levels in the vertical direction (reaching to about 700 hPa from the surface). The motivation for using potential temperature and dewpoint temperature, the filter equations, and a discussion of the observational operator are all found in Fujita et al. (2007).

b. Precipitation observations

The precipitation data used are the hourly National Centers for Environmental Prediction (NCEP) stage IV accumulated precipitation data, a synthesis of radar and gauge observational data valid over the contiguous 48 states at 4-km resolution (Baldwin and Mitchell 1997). While the assimilation may benefit from more frequent observations of precipitation, they are not yet available operationally. The hourly data are resampled to the 30-km model grid by averaging all of the 4-km precipitation observations whose center point lies within each 30-km model grid box. Only the information on whether or not precipitation is observed (i.e., greater than 0) at each model grid element is used to influence the model fields and the precipitation data are assimilated every 15 min during the assimilation period. Because the precipitation data are hourly, the identical data are referenced four times during a 1-h period. The benefits of using a 15-min assimilation cycle are more fully discussed below.

In the assimilation procedure, the precipitation forecasts and resampled observations are compared at every grid element for each ensemble member, and an adjustment of the model fields in the corresponding vertical column (and in some cases also in the columns around it) is performed if they do not coincide. The adjustment is performed on the model vertical profiles of temperature ($T$), water vapor mixing ratio ($q$), and vertical velocity ($w$), and is the same for two cases studied as described in detail below. Some parts of the adjustment procedure depend on the convective parameterization scheme used for the ensemble member. Three different convective schemes—those of Grell (Grell et al. 1994), Kain–Fritsch (Kain and Fritsch 1993), and Betts–Miller (Betts 1986; Betts and Miller 1986; Betts and Miller 1993; Janjic 1994)—are used in the ensemble. There are two situations when the model fields need to be modified.
1) PRECIPITATION IS OBSERVED, BUT NOT PREDICTED

The simple adjustment technique begins by calculating the convective available potential energy (CAPE) of the ensemble member for the grid element at which precipitation is observed but not predicted. The source layer of the convection is taken to be the level that yields the highest equivalent potential temperature in the lowest 300 hPa above ground level (e.g., Baldwin et al. 2002). If CAPE is positive, the $T$ and $q$ fields at levels between the cloud base [the lifting condensation level (LCL)] and cloud top [the equilibrium level (EL)] are modified to be more representative of cloudy conditions. It is assumed that the cloud $T$ and $q$ profiles follow a moist adiabat defined by the source layer parcel. Thus, the resulting $T$ and $q$ fields after adjustment (indicated by the subscript “new”) are given by an equal mixture of the cloudy (subscript “cloud”) and the original environmental (subscript “env”) air, such that

$$T_{\text{new}} = (T_{\text{env}} + T_{\text{cloud}})/2 \quad \text{and} \quad q_{\text{new}} = (q_{\text{env}} + q_{\text{cloud}})/2.$$  

By referencing the precipitation data 4 times an hour, the model environmental vertical profile adjusts fairly quickly toward the moist adiabat defined by the source layer parcel using a time scale that is typical of those used in convective parameterizations. Results (not shown) indicate that this adjustment often leads to the activation of the convective scheme in less than 1 h, although it also can lead to the development of explicit grid-resolved precipitation if the model vertical motion field is supportive.

If the ensemble member convective scheme is either the Kain–Fritsch or Grell scheme, then the trigger function that determines whether or not convection activates is very sensitive to the sign of the model vertical motion field. Thus, the condition $w_{\text{new}} \geq 0.005$ m s$^{-1}$ is also imposed on the vertical velocity from the surface to the EL to make certain that this grid point is evaluated by the convective scheme. This alteration to vertical motion is not long lived and does not appear to produce significant gravity waves that might negatively impact the forecasts.

This simple adjustment technique both reduces the convective inhibition and warms and moistens the atmosphere between the level of free convection (LFC) and the EL. By successively moving the environment toward a moist adiabat after each 15-min analysis step, the convective schemes become more and more likely to activate through their own internal trigger functions. In addition, with the environment moving toward saturation, grid-resolved precipitation also is more likely to develop.

If CAPE is found to be zero or negative, an identical procedure is applied regardless of what convective scheme is used. Because CAPE $\geq 0$ is needed for convection to form, it is assumed that both CAPE is underforecast and that this underforecast is due to model mixing ratio errors in the lower troposphere. Thus, the values of $q$ in the lower model layers are increased to create positive CAPE. The resulting $q$ values ($q_{\text{new}}$) between the surface and one model level above the LCL are determined using the saturation mixing ratio at the environmental temperature ($q_{\text{sat}}$), such that

$$q_{\text{new}} = (q_{\text{env}} + q_{\text{sat}})/2.$$  

The vertical velocity also is set to no less than 0.005 m s$^{-1}$ over the same layer. Because hourly surface observations are being assimilated with the ensemble Kalman filter, yielding reasonably accurate boundary layer predictions, this CAPE modification is not used very often. In regions lacking surface observations, but with precipitation observations, this portion of the algorithm becomes more important.

This simple adjustment approach is similar in philosophy, but different in application, from the one discussed in Rogers et al. (2000). In their approach, the Kain–Fritsch convective scheme is activated by forcing a parcel from the source layer to its LFC and allowing the convective scheme to produce the resulting heating and moistening profiles. Thus, the trigger function within the convective scheme is modified to allow convection to develop at the right time and place. The present approach is designed to obtain the same end result—the development of convection at the right time and place with the resulting heating and moistening profiles determined in response to the local environment—but the adjustment approach operates external to the model integration so that it can be used more easily in ensembles that contain multiple convective schemes. It also may help produce grid-scale saturation and the development of precipitation from resolvable-scale processes.

2) NO PRECIPITATION IS OBSERVED, WHILE MODEL PREDICTS PRECIPITATION

In Rogers et al. (2000), the convective tendencies are set to zero when the model predicts convection yet it is not observed. Unfortunately, simply setting the convective tendencies to zero in the two cases examined in this study appears to only delay convective activation slightly, suggesting that a different approach is needed.
If the grid-scale precipitation of the ensemble member is nonzero, then the conditions $w_{new} < 0.0$, rainwater = 0.0, and relative humidity <90% are imposed at all model levels. If the convective precipitation is nonzero, then the convective precipitation and tendencies over the previous 15-min period at the grid element are set to zero. The vertical profiles of $T$ and $q$ are replaced by their values prior to when the convection began, that is, their values at the time of the previous analysis step (indicated in the following by subscript “−15min”). These values of $T$ and $q$ are then modified depending on the convective scheme used in the ensemble member to reduce the likelihood of convection activating in the next hour. In the Kain–Fritsch and Grell convective schemes, convective inhibition is an important factor in convective scheme activation. Thus, convective inhibition is increased by altering the $T$ and $q$ values in the levels between the source layer of the convection (the level with the highest equivalent potential temperature in the lowest 300 hPa) and the LFC, such that

$$T_{new} = T_{-15\text{min}} + 2.0 \quad \text{and} \quad q_{new} = q_{-15\text{min}} \times 0.8$$

(4)

(5)

throughout these layers. For the Betts–Miller–Janjic (BMJ) scheme, cloud-layer moisture determines whether or not the scheme initiates convection. Thus, the mixing ratio between the cloud base (LCL) and the cloud top (EL) is reduced in order to deactivate the convection, so that $q_{new} = q_{-15\text{min}} \times 0.8$ between the LCL and the EL.

Initial tests suggest that altering a single grid element to deactivate convection is not very effective, as convection often activates at neighboring grid elements over the next hour. Thus, the same modification to either convective inhibition or cloud-layer moisture also is applied to nearby grid elements where both the observation and the model indicate no precipitation, or where the model indicates no precipitation and the observation is missing. However, the strength of this modification is reduced as the distance from the original grid element increases, with the radius of influence set to five grid points (150 km) using the function of Gaspari and Cohn (1999). If a grid point has more than one neighboring point where the model indicates convection and the observation does not, then the strength of the adjustment is based upon the distance to the closest grid point where convection is deactivated. This adjustment of the neighboring grid points is performed to provide smoother transitions between adjusted and unadjusted regions. Admittedly, other values for this radius of influence may yield better results.

While it is possible to argue with a number of the particular choices used in developing this simple approach to activating and deactivating the convective parameterizations, it is important to recognize that the ability to assimilate precipitation observations in an ensemble that includes variability in the convective parameterization scheme is greatly simplified by using an approach that operates external to the model integration. This necessitates making changes to the environment that are consistent with the behavior of the various schemes, yet can be formulated in a simple algorithm. While the approach developed here is far from perfect, it offers a starting point for exploring precipitation assimilation methods that work in parallel with an ensemble Kalman filter.

3. Experiment design

The two cases selected for further study are the same as those in Fujita et al. (2007): 1200 UTC 8 May–0000 UTC 9 May 2003 (hereafter, the May case) and 1200 UTC 1 July–0000 UTC 2 July 2003 (hereafter, the July case). Both cases are associated with severe weather, with the May case dominated by deep convection that develops along a frontal boundary and dryline within southwesterly flow and the July case dominated by a decaying tropical depression downstream of a large-scale ridge. The environments from these two cases are very different, with most of the convection occurring within environments of moderate CAPE in the May case and in environments of low CAPE in the July case. Further information on each of the cases is found in Fujita et al. (2007). In each case, the assimilation of the surface and precipitation observations is performed during the first 6 h, followed by a 6-h ensemble forecast.

The forecast model used in the experiment is the nonhydrostatic fifth-generation Pennsylvania State University–National Center for Atmospheric Research (PSU–NCAR) Mesoscale Model (MM5; Dudhia 1993; Grell et al. 1994). The simulations are performed on a Lambert conformal projection grid domain of 180 × 120 grid points, centered at 37.0°N, 97.0°W, with grid spacing of 30 km (hereafter, the experimental domain). The simulations have 24 full-sigma levels in the vertical direction, with the constant pressure at the model top taken to be 100 hPa.

The forecast model used in the experiment is the nonhydrostatic fifth-generation Pennsylvania State University–National Center for Atmospheric Research (PSU–NCAR) Mesoscale Model (MM5; Dudhia 1993; Grell et al. 1994). The simulations are performed on a Lambert conformal projection grid domain of 180 × 120 grid points, centered at 37.0°N, 97.0°W, with grid spacing of 30 km (hereafter, the experimental domain). The simulations have 24 full-sigma levels in the vertical direction, with the constant pressure at the model top taken to be 100 hPa.

The forecast model used in the experiment is the nonhydrostatic fifth-generation Pennsylvania State University–National Center for Atmospheric Research (PSU–NCAR) Mesoscale Model (MM5; Dudhia 1993; Grell et al. 1994). The simulations are performed on a Lambert conformal projection grid domain of 180 × 120 grid points, centered at 37.0°N, 97.0°W, with grid spacing of 30 km (hereafter, the experimental domain). The simulations have 24 full-sigma levels in the vertical direction, with the constant pressure at the model top taken to be 100 hPa.

The forecast model used in the experiment is the nonhydrostatic fifth-generation Pennsylvania State University–National Center for Atmospheric Research (PSU–NCAR) Mesoscale Model (MM5; Dudhia 1993; Grell et al. 1994). The simulations are performed on a Lambert conformal projection grid domain of 180 × 120 grid points, centered at 37.0°N, 97.0°W, with grid spacing of 30 km (hereafter, the experimental domain). The simulations have 24 full-sigma levels in the vertical direction, with the constant pressure at the model top taken to be 100 hPa.
boundary conditions are incorporated using the breeding of growing modes (BGM) technique (Toth and Kalnay 1993, 1997) as developed at NCEP. The breeding cycles are performed over an 84-h period prior to the ensemble start time on a 90-km Lambert conformal projection grid domain of 180 x 130 grid points, centered at 37.0°N, 97.0°W (hereafter, the BGM domain). Thus, the BGM domain is approximately three times as large as the experimental domain and extends well out into both the Pacific and Atlantic Oceans. The initial perturbations for the breeding cycles are generated using a Monte Carlo method (Errico and Baumhefner 1987) applied to \( T, u, v \) wind components, geopotential height, relative humidity, and sea level pressure fields. During the 84-h breeding period, the perturbations of \( T \), dewpoint temperature, \( u \), \( v \), and perturbation pressure are scaled every 12 h. Following the approach developed at NCEP, the scaling factor is determined for all fields and vertical levels by calculating the factor needed for the BGM domain average \( T \) perturbation at model level \( z = 0.825 \) to be equal to 0.7 K (J. Du 2004, personnel communication). Twelve breeding cycles are performed to generate 12 perturbations, which are added to and subtracted from the NCEP global analysis valid at the initial time of the experimental period, resulting in 24 perturbed fields. Including the control field, which is the NCEP global analysis field itself without perturbations, an ensemble of 25 initial conditions is created. After the breeding cycles, forecast integrations are performed over the 12-h experimental period on the BGM domain from the 25 initial conditions, in order to obtain the boundary conditions for the experimental period. An identical model physical process scheme, which is listed in the first row of Table 1, is used for all of the ensemble members to create the bred modes.

The additional uncertainty created by imperfect model physics is taken into account by using different model physical process schemes for each of the 25 ensemble members during the 12-h experimental period only. The physical process schemes varied are the convection, planetary boundary layer (PBL), radiation, and land surface schemes as listed in Table 1 [see Fujita et al. (2007) for further discussion]. Together, the 25 different physics configurations and the 25 different initial and boundary conditions yield the ICPH ensemble.

Runs with two different ensembles are performed and compared. Both use the identical ICPH ensemble configuration described above, but in the first one both the surface and precipitation (P) observations are assimilated (hereafter, the ICPH + P ensemble), while in the second one, only the surface observations are assimilated (hereafter, the ICPH ensemble).

<table>
<thead>
<tr>
<th>Member</th>
<th>Convective</th>
<th>PBL</th>
<th>Radiation</th>
<th>Surface</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Grell</td>
<td>MRF</td>
<td>Cloud</td>
<td>Noah</td>
</tr>
<tr>
<td>2</td>
<td>Grell</td>
<td>Blackadar</td>
<td>Cloud</td>
<td>5 layer</td>
</tr>
<tr>
<td>3</td>
<td>Grell</td>
<td>Burk–Thompson</td>
<td>Cloud</td>
<td>Force–restore</td>
</tr>
<tr>
<td>4</td>
<td>Grell</td>
<td>Eta</td>
<td>Cloud</td>
<td>Noah</td>
</tr>
<tr>
<td>5</td>
<td>Grell</td>
<td>Blackadar</td>
<td>CCM2</td>
<td>Five layer</td>
</tr>
<tr>
<td>6</td>
<td>Grell</td>
<td>Burk–Thompson</td>
<td>CCM2</td>
<td>Force–restore</td>
</tr>
<tr>
<td>7</td>
<td>Grell</td>
<td>Eta</td>
<td>CCM2</td>
<td>Noah</td>
</tr>
<tr>
<td>8</td>
<td>Grell</td>
<td>MRF</td>
<td>CCM2</td>
<td>Noah</td>
</tr>
<tr>
<td>9</td>
<td>Kain–Fritsch</td>
<td>Blackadar</td>
<td>CCM2</td>
<td>Five layer</td>
</tr>
<tr>
<td>10</td>
<td>Kain–Fritsch</td>
<td>Burk–Thompson</td>
<td>CCM2</td>
<td>Force–restore</td>
</tr>
<tr>
<td>11</td>
<td>Kain–Fritsch</td>
<td>Eta</td>
<td>CCM2</td>
<td>Noah</td>
</tr>
<tr>
<td>12</td>
<td>Kain–Fritsch</td>
<td>MRF</td>
<td>CCM2</td>
<td>Noah</td>
</tr>
<tr>
<td>13</td>
<td>Betts–Miller</td>
<td>Blackadar</td>
<td>CCM2</td>
<td>Five layer</td>
</tr>
<tr>
<td>14</td>
<td>Betts–Miller</td>
<td>Burk–Thompson</td>
<td>CCM2</td>
<td>Force–restore</td>
</tr>
<tr>
<td>15</td>
<td>Betts–Miller</td>
<td>Eta</td>
<td>CCM2</td>
<td>Noah</td>
</tr>
<tr>
<td>16</td>
<td>Betts–Miller</td>
<td>MRF</td>
<td>CCM2</td>
<td>Noah</td>
</tr>
<tr>
<td>17</td>
<td>Kain–Fritsch</td>
<td>Blackadar</td>
<td>RRTM</td>
<td>Five layer</td>
</tr>
<tr>
<td>18</td>
<td>Kain–Fritsch</td>
<td>Burk–Thompson</td>
<td>RRTM</td>
<td>Force–restore</td>
</tr>
<tr>
<td>19</td>
<td>Kain–Fritsch</td>
<td>Eta</td>
<td>RRTM</td>
<td>Noah</td>
</tr>
<tr>
<td>20</td>
<td>Kain–Fritsch</td>
<td>MRF</td>
<td>RRTM</td>
<td>Force–restore</td>
</tr>
<tr>
<td>21</td>
<td>Betts–Miller</td>
<td>Blackadar</td>
<td>RRTM</td>
<td>Five layer</td>
</tr>
<tr>
<td>22</td>
<td>Betts–Miller</td>
<td>Burk–Thompson</td>
<td>RRTM</td>
<td>Force–restore</td>
</tr>
<tr>
<td>23</td>
<td>Betts–Miller</td>
<td>Eta</td>
<td>RRTM</td>
<td>Noah</td>
</tr>
<tr>
<td>24</td>
<td>Betts–Miller</td>
<td>MRF</td>
<td>RRTM</td>
<td>Noah</td>
</tr>
<tr>
<td>25</td>
<td>Kain–Fritsch</td>
<td>MRF</td>
<td>Cloud</td>
<td>Noah</td>
</tr>
</tbody>
</table>

TABLE 1. Model configurations used in the experiment. Member number 1 is the model configuration used in the breeding cycles.
ensemble mean and the surface observations of $\theta$ at 2 m above the ground averaged over all of the observation locations for the May case. The rms difference is calculated before and after the analysis step during the assimilation period (0–6-h simulation time). Results are from the ICPH (solid line) and ICPH + P (dashed line) ensembles.

4. Results

a. 8–9 May 2003 case

A surface low pressure system is located in western Kansas at 1800 UTC 8 May 2003 with a pressure trough and dryline stretching southward into Oklahoma. A warm front extends eastward from the low across northern Kansas and then southeastward across Missouri. Rainfall is observed primarily to the north of the frontal boundary between 1200 and 2100 UTC. Deep convection finally develops in the warm sector across Oklahoma and central Kansas in association with a short-wave trough and a dryline between 2100 and 2200 UTC. Values of CAPE in the warm sector at the time of convective initiation are above 3000 J kg$^{-1}$.

The time sequence of the potential temperature root-mean-square (rms) difference between the ensemble mean and the surface observations (averaged over the observation points) shows that the ICPH + P ensemble results are similar to those from the ICPH ensemble (Fig. 1). The same also can be said for dewpoint temperature, and the $u$ and $v$ wind components (not shown). This indicates that the assimilation of precipitation data does not disturb the description of the surface fields, suggesting that the simple assimilation method proposed here can consistently incorporate simultaneous precipitation information in parallel with surface data assimilation using the ensemble Kalman filter.

The ensemble mean 1-h accumulated rainfall at 1800 UTC 8 May from the ICPH + P ensemble, at the end of the assimilation period (Fig. 2a), shows that the assimilation procedure largely moves the precipitation distribution closer to the observations (Fig. 2c). The ICPH ensemble (Fig. 2b) shows relatively large regions of rainfall over 1 mm stretching along the axes of higher CAPE from eastern Kansas to southwestern Oklahoma and from central Tennessee to central Alabama. Neither of these rainfall regions are seen in the observations or the ICPH + P ensemble. Thus, the extent of the rainfall region clearly is reduced by the assimilation of the precipitation data. In the ICPH + P ensemble, the precipitation associated with a frontal boundary stretching from southeastern South Dakota southeastward to southern Illinois, near the western border of Tennessee, and across the borders of Virginia and Maryland is more clearly concentrated than in the ICPH ensemble. Also note that the precipitation in the ICPH ensemble over Utah, Nevada, and northern California is suppressed in the ICPH + P ensemble. These features of the rainfall distribution in the ICPH + P ensemble are consistent with the NCEP stage IV rainfall observations (Fig. 2c), indicating that the assimilation procedure is working properly to locate the model precipitation distribution according to the observational information. This also is seen in calculations of ensemble spread for hourly rainfall, in which the ICPH + P ensemble has spread values 23%–53% smaller than those from the ICPH ensemble during the 6-h assimilation period (not shown).

The effect of the precipitation assimilation on the rainfall distribution can be seen in the time sequence of the bias score (Wilks 1995) of 1-h accumulated ensemble mean rainfall for different thresholds (Fig. 3). Bias scores close to 1 indicate agreement between the forecast and observation for the areal coverage of rainfall exceeding the given threshold. The bias scores of the ICPH + P ensemble for lower thresholds of 0.1 and 0.5 mm h$^{-1}$ (Figs. 3a and 3b) are much closer to 1.0 than those of the ICPH ensemble during the assimilation period. These results indicate that the information on the location of the precipitation from the observations is used to restrain the region of precipitation, suppressing the overforecasting of the rainfall region. Bias scores closer to 1.0 also are seen for the threshold of 1.0 mm h$^{-1}$ (Fig. 3c), suggesting that many regions with relatively stronger precipitation are captured well by the ICPH + P ensemble. However, rainfall distributions from the ensemble mean are generally broad and less peaked compared to the observations, as reflected in the bias scores for the higher threshold of 2.5 mm h$^{-1}$ (Fig. 3d). Both the ICPH + P and the ICPH ensembles produce bias values of less than 1. The improvement of the bias score in the threshold amounts less than 1 mm h$^{-1}$ persists in the forecast period after the assimilation,
Fig. 2. Ensemble-average accumulated rainfall (mm) from 1700 to 1800 UTC 8 May 2003 from the (a) ICPH + P and (b) ICPH ensembles. (c) Observed accumulated rainfall (mm) from 1700 to 1800 UTC 8 May 2003 (NCEP stage IV at 4-km grid spacing), resampled to the experimental 30-km domain. Thick and thin solid lines indicate the isolines of 1.0- and 0.1-mm accumulated rainfall, respectively. The dashed line in (c) outlines the region where the precipitation data are available over the United States; grid points outside this region are not influenced by the precipitation-data assimilation. The low center, warm front, and dryline location are shown.
particularly in the first 2 h, with the ICPH + P ensemble having a bias score closer to 1 than the ICPH ensemble over the length of the experimental period. The tendency for convective parameterization schemes to produce too much light precipitation (Kain 2004) partially explains the smaller improvements in the ensemble mean bias after the first 2 h of the forecast period because these schemes are no longer constrained by the observations.

The ICPH + P ensemble yields a smaller rms difference of hourly precipitation between the ensemble mean and observations (averaged over grid elements where the observations are available) than the ICPH ensemble over the length of the experimental period (Fig. 4). Although information on rain rates from the observations is not used in the assimilation, the suppression of model rainfall over the regions with no observed rainfall is believed to reduce the rms difference.

The probability of accumulated rainfall exceeding 1 mm h$^{-1}$ at 2000 UTC 8 May, 2 h after the end of the assimilation period, clearly shows the influence of the assimilation on the precipitation forecast (Fig. 5). Probabilities are calculated by simply assuming that each ensemble forecast is equally likely. While the ICPH ensemble produces higher probabilities from north-central Kansas southeastward into southern Missouri along a warm front, the probabilities over this region are reduced in the ICPH + P ensemble. Instead, the ICPH + P ensemble gives higher probabilities over

![Fig. 3. Time sequence of the bias score of the ensemble mean 1-h accumulated rainfall for thresholds of (a) 0.1, (b) 0.5, (c) 1.0, and (d) 2.5 mm. Results are from the ICPH (solid line) and ICPH + P (dashed line) ensembles.](image1)

![Fig. 4. Time sequence of the rms difference of the 1-h accumulated rainfall between the ensemble mean and the observations averaged over grid elements where the observations are available. Results are from the ICPH (solid line) and ICPH + P (dashed line) ensembles.](image2)
Fig. 5. Probability (%) of accumulated rainfall exceeding 1 mm between 1900 and 2000 UTC 8 May 2003 from the (a) ICPH + P and (b) ICPH ensembles. (c) Observed accumulated rainfall (mm) from 1900 to 2000 UTC 8 May 2003 (NCEP stage IV at 4-km grid spacing), resampled to the experimental 30-km domain. Thick and thin solid lines in (c) indicate the isolines of the 1.0- and 0.1-mm accumulated rainfall, respectively. Dashed line in (c) is as in Fig. 2c. The low center, warm front, and dryline location are shown.
Iowa, which is more consistent with the observations than is the ICPH ensemble. Other notable differences in the rainfall probabilities are seen in western Kentucky and Tennessee, and over eastern Virginia.

An attribute diagram (Wilks 1995) for accumulated rainfall of 1 mm h\(^{-1}\) or greater shows the reliability of all of the ensemble probability forecasts where observations are available for every hour over the whole experimental period (1200 UTC 8 May–0000 UTC 9 May 2003). Results are from the ICPH (solid line) and ICPH + P (dashed line) ensembles. Error bars show the 95% confidence interval.

b. 1–2 July case

A tropical depression is slowly crossing the southeastern states on the morning of 1 July 2003. The depression is located over Mississippi and Alabama at 1200 UTC and moves into Georgia as the day progresses. Most of the rainfall across the United States is associated with the tropical depression, with environmental lapse rates near moist adiabatic and values of CAPE between 0 and 500 J kg\(^{-1}\) across much of the southeastern United States. Times series of surface observations from the ICPH + P and ICPH ensembles again indicate that the precipitation assimilation is not negatively influencing the surface data assimilation from the ensemble Kalman filter (not shown).

The ensemble mean 1-h accumulated rainfall at 2000 UTC 1 July, 2 h after the end of the assimilation period, highlights that most of the rainfall in this case is associated with the tropical depression over the southeastern United States (Fig. 7). The observed region of the most intense rainfall, which is seen over Georgia, almost coincides with the higher predicted mean rainfall totals from both the ICPH and ICPH + P ensembles. Improvements in the mean rainfall distribution owing to the precipitation data assimilation are seen in the rainband stretching southwestward from southwestern Georgia, scattered spots of relatively intense rainfall distributed from western Virginia to northeastern Texas, and over the Florida peninsula. These mesoscale precipitation features are more clearly produced in the ICPH + P ensemble and qualitatively agree better with the observations than those found in the ICPH ensemble. The ensemble spread for hourly rainfall again shows that the ICPH + P ensemble has smaller spread values than the ICPH ensemble during the 6-h assimilation period, although for this case the spread is only decreased by 5% on average.

As in the May case, the bias scores for the precipitation thresholds of 0.1 and 0.5 mm h\(^{-1}\) (Fig. 8) show that the ICPH + P ensemble yields values closer to 1 than the ICPH ensemble over the entire experimental period, particularly during the first 2 h after the assimilation period. However, for the highest threshold amount of 2.5 mm h\(^{-1}\), the ICPH + P ensemble does not improve upon the bias score in comparison to the ICPH ensemble until 2 h into the forecast period (hour 8), although this improvement then lasts for the remainder of the forecast period. In this case, the region of intense rainfall is focused over Georgia, with fine-structured patterns scattered around the tropical depression. Thus, the improvement of the overall pattern of the rainfall distribution is not necessarily reflected in the grid-element-wise comparison to the observation during the assimilation period. The 1-h accumulated rainfall rms difference between the ensemble mean and the observations (Fig. 9) shows that the ICPH + P ensemble often yields slightly larger difference values than the ICPH ensemble during the assimilation period. Even the ICPH ensemble without the assimilation of precipitation produces well the location of the region of intense rainfall in most ensemble members. It appears that improvements to the locations of convection within the relatively weaker rainfall zones around the tropical depression do not have a significant influence on the rms difference. The rms differences are nearly identical after 2000 UTC (8 h).
Fig. 7. As in Fig. 2 but for accumulated rainfall (mm) from 1900 to 2000 UTC 1 Jul.
An attribute diagram of accumulated rainfall for the threshold of 1.0 mm h\(^{-1}\) (Fig. 10) from the two ensembles again highlights the value of the precipitation-data assimilation. The probabilities from the ICPH + P ensemble are skillful over a larger range of probabilities (0.32–1.0) in comparison to the ICPH ensemble (0.45–1.0), and for model forecast probabilities <0.32 the ICPH + P ensemble shows improvements to the relationship between the model forecast probability and the observed relative frequency. Thus, the ensemble probabilities from the ICPH + P ensemble in general are an improvement over those from the ICPH ensemble. However, the probabilities from the ICPH + P ensemble are slightly less reliable than those from the ICPH ensemble for probabilities between 0.7 and 0.8.

The grid elements contributing these higher probabil-
ties are mainly distributed over Georgia, where intense rainfall by the tropical depression is located. Because the location of the center of the rainfall is well specified even without the information from the precipitation observation, the forecast precipitation probability from the ICPH ensemble yields a high correlation with the observed relative frequency over this region.

5. Summary

A simple adjustment procedure is developed to assimilate precipitation information from radar and gauge-derived observations alongside an ensemble Kalman filter system that assimilates hourly surface observations. The precipitation assimilation only uses the information on the location and timing of the rainfall as in Rogers et al. (2000). Values of temperature, mixing ratio, rainwater, and vertical motion from each ensemble member at each grid element are modified to increase the likelihood of convective development during the assimilation period in accordance with the information from the corresponding observation in the grid element. This procedure is applied in parallel with an ensemble Kalman filter system that assimilates hourly surface observations over a 6-h period into a 25-member ensemble with both different initial and boundary conditions and different physical process schemes. A synthesis of radar and gauge observational data is assimilated every 15 min from 1200 to 1800 UTC for two very different cases, and the results are compared with runs without precipitation assimilation over the following 6-h forecast period.

Results show that the simple precipitation assimilation is able to develop precipitation where it is observed, resulting in the precipitation area being reproduced more accurately than in the run without precipitation-data assimilation. In addition, the positive results from the surface data assimilation are maintained. The overforecasting of light precipitation is greatly reduced by the precipitation assimilation, showing the largest decrease in ensemble mean bias scores for the lowest threshold amounts. The precipitation distribution and ensemble spread measures suggest that the procedure reduces the displacement between the model precipitation and observations, yielding an overall areal precipitation distribution that is closer to the observations. However, at larger threshold values the ensemble mean rainfall amounts may not be improved. The reliability of the precipitation probabilities for a 1 mm h⁻¹ threshold from the model are improved significantly by using the precipitation data for the May severe weather case, and are shown to be generally more reliable over a larger range of probabilities for the July tropical depression case. In general, the improvement of the rainfall forecasts can be seen for a few hours beyond the assimilation period, suggesting that the assimilation of the precipitation helps to provide an improved initial condition for ensemble forecasting in which convection can develop in accordance with the observations. Potential improvements to this procedure in the future are the inclusion of information on the rain rates as in Manobianco et al. (1994) and flow dependency.

It is encouraging that a simple method for assimilating precipitation data, which operates outside of the convective parameterization schemes, can be run in tandem with an ensemble Kalman filter assimilation scheme and yields improvements in the probabilistic quantitative precipitation forecasts (PQPFs) predicted by the model without degrading other aspects of the assimilation and forecasts. While it is clear that precipitation forecasts remain a very challenging facet of numerical weather prediction, it is hoped that increased use of physics and initial condition variability in ensembles will lead to forecast improvements, especially when combined with precipitation assimilation.

Acknowledgments. This research was conducted while the lead author (TF) was a visiting scientist at the National Severe Storms Laboratory and the Sasaki Institute, and funding support from the Office of Vice President for Research, College of Atmospheric and Geographic Sciences, and the School of Meteorology of the University of Oklahoma is gratefully acknowledged. The third author (DD) was supported by the National Science Foundation under Grant 0333872. Discussions with Drs. Y. K. Sasaki and K. L. Elmore were very helpful and are sincerely appreciated. The helpful comments from three anonymous reviewers led to improvements to the presentation and are gratefully acknowledged.

REFERENCES

——, J. S. Kain, and M. P. Kay, 2002: Properties of the convection


Manobianco, J., S. Koch, V. M. Karyampudi, and A. J. Negri, 1994: The impact of assimilating satellite-derived precipita-


