Distributed Quantitative Precipitation Forecasting Using Information from Radar and Numerical Weather Prediction Models

AUROOP R. GANGULY AND RAFAEL L. BRAS

Ralph M. Parsons Laboratory, Civil and Environmental Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts

(Manuscript received 13 December 2002, in final form 25 May 2003)

ABSTRACT

The benefits of short-term (1–6 h), distributed quantitative precipitation forecasts (DQPFs) are well known. However, this area is acknowledged to be one of the most challenging in hydrometeorology. Previous studies suggest that the ''state of the art'' methods can be enhanced by exploiting relevant information from radar and numerical weather prediction (NWP) models, using process physics and data-dictated tools where each fits best. Tests indicate that improved results are obtained by decomposing the overall problem into component processes, and that each process may require alternative tools ranging from simple interpolation to statistical time series models and artificial neural networks (ANNs). A new hybrid modeling strategy is proposed for DQPF that utilizes measurements from radar [Weather Surveillance Radar-1998 Doppler (WSR-88D) network: 4 km, 1 h] and outputs from NWP models (48-km Eta Model: 48 km, 6 h). The proposed strategy improves distributed QPF over existing methods like radar extrapolation or NWP-based QPF by themselves, as well as combinations of radar extrapolation and NWP-based QPF.

1. Introduction

Distributed quantitative precipitation forecasts (DQPF) are acknowledged to be among the most challenging areas in hydrology and meteorology (Fritsch et al. 1998; Collier and Kryzysztofowicz 2000). Improvements to precipitation forecasting have been somewhat limited (Smith and Austin 2000) over the last few decades. Several factors contribute to making precipitation forecasting difficult. Precipitation processes can occur over a range of scales, ranging from large airmass movements to extremely localized convective events (Rogers and Yau 1989). Precipitation is further influenced by air motion and turbulent eddies, aerosol properties, and microphysical processes that dictate droplet growth and evaporation. The net effect is that precipitation is extremely variable in time and space.

The appropriate temporal and spatial resolutions for quantitative precipitation forecasts (QPF) vary with the application and the methodology used. At very short lead times (0–1 h), high forecast resolutions are needed. Forecasts based on extrapolation of observations like radar rainfall are often found to be most skillful in this context (Golding 2000; Collier 1991). At lead times greater than about 6 h, numerical weather prediction (NWP) models provide QPF at lower space–time resolutions. Antolik (2000) of the National Weather Service (NWS) mentions that future research needs to explore methods that combine the 48-km Eta NWP model outputs with ''stage III'' Weather Surveillance Radar-1998 Doppler (WSR-88D) data. Stage III is a radar–gauge analysis that has undergone manual quality control (Fulton et al. 1998). This study focuses on QPF at lead times of 1–6 h, for high-resolution hydrologic applications. A hybrid strategy that combines NWP model outputs and radar measurements leads to an improvement in distributed QPF.

The paper uses the WSR-88D network, which covers the continental United States providing hourly precipitation measurements at 4 km $\times$ 4 km resolutions, and the Eta Model, which generates forecasts for precipitation and other atmospheric variables at 48-km and 6-h resolution. Note that the Eta Model forecast resolution was 48 km during the study period. At the time of writing, the operational resolution is 12 km. Also, the Eta Model is now run 4 times daily (0000, 0600, 1200, and 1800 UTC). The proposed hybrid strategy is summarized in the following paragraphs.

The distributed QPF problem has been decomposed into four component processes. The motivation for the decomposition was to optimize the use of available information from radar and NWP, and to make the best use of the available process physics and data-dictated tools. The processes are modeled in succession, with each component superimposed on the results of the previous components. A schematic flowchart for the pro-

Corresponding author address: Professor Rafael L. Bras, Ralph M. Parsons Laboratory, Civil and Environmental Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139.
E-mail: rlbras@mit.edu

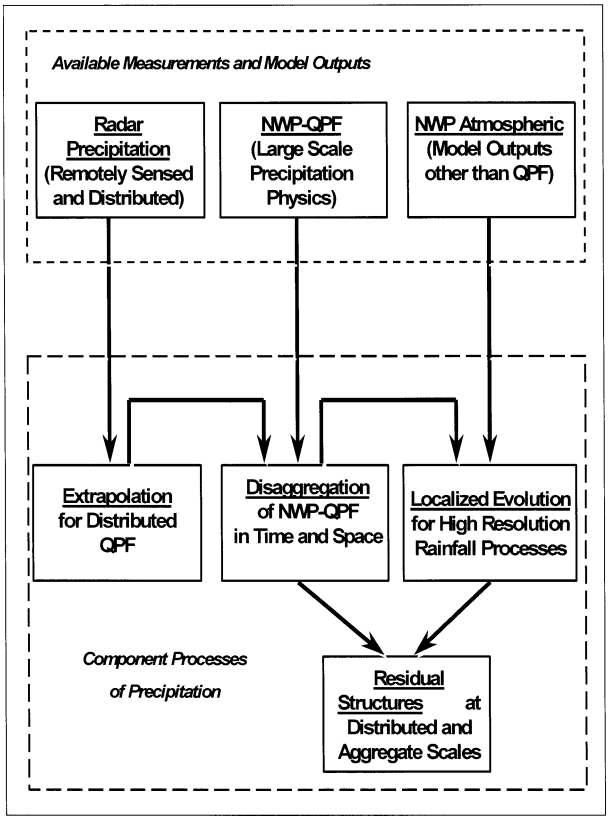

FIG. 1. A schematic flowchart for the proposed hybrid modeling strategy of the distributed QPF problem which has been decomposed into four components: radar extrapolation, large-scale physics, localized evolution, and residual structures.

posed hybrid modeling strategy is shown in Fig. 1. The output of the proposed approach is distributed QPF and probabilistic QPF (PQPF), that is, forecasts for rainfall in space and time, and a measure of the uncertainty associated with these forecasts.

The first component is radar extrapolation, handled through advection of the hourly radar maps. Velocity scales at each time in the past are estimated using two-dimensional spatial correlation by comparing the current map with the map at the previous time step. An exponential smoothing–based time series formulation is used for estimating the velocity scales in the future. This formulation applies relative weights to the past estimates of velocity scales with the weights decaying as one goes further into the past.

The second component is large-scale physics, as represented by NWP-based QPF. NWP-based QPF is grid averaged and cumulative (48 km and 6 h for NWP Eta). Disaggregation in time is performed by linear interpolation for QPF and other NWP model outputs. Temporal disaggregation is followed by an autoregressive (AR)-based error correction for the grid-averaged QPF, by comparing it with aggregated radar rainfall. The temporally disaggregated error-corrected, grid-averaged QPF is ''superimposed'' on the results of advection. The distributed structure obtained from advection is retained, but the values of the pixels with nonzero rainfall are scaled to reflect the large-scale mean from NWP-based QPF.

Localized evolution is modeled at radar pixel resolutions as a function of NWP atmospheric forecasts and the distributed precipitation ''state,'' which, in turn, is the result of radar extrapolation followed by NWP-based QPF disaggregation. The NWP-based QPF ''simple disaggregation'' models the grid average precipitation forecast based on advection, while the ''localized evolution'' models the perturbations from this ''mean'' structure. Evolution models localized precipitation processes that cannot be handled well by the lower-resolution NWP model physics (e.g., convective processes). The assumption is that evolution is a function of the rainfall at a pixel and its neighbors, as well as, the large-scale NWP atmospheric forecasts. Because this functional form is highly nonlinear, data-dictated artificial neural network (ANN)–based tools are used for the approximation.

The final component is residual structures or errors at both distributed and aggregate scales. Distributed errors are handled at higher resolutions by an ANN-based strategy that combines the results of the two component processes described earlier. Low-resolution errors at aggregate scales are handled by combining the spatial means of the rainfall intensities from the individual component processes.

The proposed approach is compared to existing methods like persistence, advection, and a combination of advection and QPF from NWP models. Results with six precipitation events in the Arkansas–Red River Basin River Forecast Center (ABRFC) region demonstrate that this approach has the potential to improve forecast skills.

2. Recent advances in distributed QPF

a. Extrapolation and statistical methods

The commonly used methods for DQPF remain extrapolation techniques. Simple persistence assumes that rainfall intensities at forecast lead times are identical to the most recent measurements. Advection or Lagrangian persistence translates the measured rainfall fields using an estimate of the velocity scales, which in turn could be derived from spatial correlation (Grecu and Krajewski 2000) or wind vectors. Predictability from advection has been studied in detail by recent researchers (e.g., Germann and Zawadzki 2002).

Researchers have attempted to obtain mathematical or stochastic formulations for the description and forecasting of rainfall. An example is the cluster model of Rodriguez-Iturbe and Eagleson (1987), where statistical distributions are used to describe the evolution of rainfall cells. McLaughlin et al. (1990) demonstrated how the parameters of these models could be determined

from data through the use of a discrete Kalman filter formulation.

b. QPF from NWP models and combination with extrapolation

Currently achievable NWP resolutions are not adequate for generating QPF at hydrologic scales (Antolik 2000), although it is conceivable that forecast resolutions might approach these scales in the future (Droegemeir et al. 2000). There is a belief that improvements to NWP model physics and resolutions could be the best way to enhance distributed QPF. Given the complex nature of the distributed precipitation processes and the limited understanding of the precipitation physics at those scales, this assumption remains to be proven.

Recent developments include hybrid approaches that statistically combine QPF from NWP and extrapolation. Golding (2000) mentions that improving QPF at 1–6-h lead time is a key requirement, and this could be achieved by statistically combining radar and NWP information. The operational Nimrod system in the Met Office, designed for 1–6-h QPF, combines radar advection with NWP-based QPF using relative weights. Smith and Austin (2000) describe this as a great step forward in improving QPF.

There is some indication that the future of QPF might be directed toward the convergence of statistical and numerical modeling procedures (Wilson et al. 1998), that is, the development of statistical methodologies for rapid initialization or update of high-resolution physical numerical methods. For example, there are approaches that relocate ''storms'' and restart numerical models based on the relocated storms (Brewster 1998). Other approaches adjust the forcing in numerical models to follow the precipitation dynamics (Hou et al. 2001).

c. Utilizing NWP model outputs other than QPF

NWP model output statistics (MOS) studies, which linearly regress NWP outputs with point rainfall, have discovered that QPF from NWP is not the most important predictor of precipitation (Antolik 2000), as compared with other atmospheric variables produced by the numerical models, like temperature, pressure, and humidity (Mesinger 1996). These results suggest that while a simple disaggregation of NWP-based QPF might have limited predictability, forecasts of atmospheric state variables and instability indices from NWP could have additional information content, especially in convective situations. This possibility is also suggested from the results of Perica and Foufoula-Georgiou (1996), Nakakita et al. (1996), and Sugimoto et al. (2001). Some researchers have also suggested parameterized, physically based models, and estimated the parameters thereof from NWP model outputs (Lee and Georgakakos 1996). These are conceptually appealing, but the inherent physics or statistics are typically a crude parameterization, and the applicability is often limited to specific situations for which these models were designed (Smith and Austin 2000).

d. Complex statistical approaches like artificial neural networks

Estimation and forecasting based on ANN have shown promise for geophysical applications, including quantitative precipitation estimation (QPE) and QPF (Kuligowski and Barros 2001; Elshorbagy et al. 2002a,b; Hsu et al. 1999). ANNs are complex data-dictated tools that have been shown to act as ''universal function approximators,'' and converge faster than other traditional approximators. However, there is a need for caution and the applicability of ANN for a specific situation needs to be determined from statistical data analysis and domain knowledge. This is demonstrated in the QPF context by contrasting the applications of Toth et al. (2000) where ANN improved forecasts, and Grecu and Krajewski (2000) where no significant improvements were observed.

3. Data

a. Eta NWP model outputs

Model outputs from the 48-km Eta NWP are archived for the continental United States by the NWS. Black (1994) provides a detailed description of the model. Of the NWP model outputs, QPF is usually the one with the highest variability in space and time, and is thought to be least accurate (Mesinger 1996). This could be due to the inherent variability of precipitation, presence of thresholds and intermittence, and the fact that QPF is a derived quantity (Antolik 2000). Other NWP model outputs, corresponding to forecasts of atmospheric variables [e.g., lifted index (LI), temperature (T)], are more uniform in space and time, and more accurate. The Eta NWP model was run twice daily during the study period, at 0000 and 0012 UTC. Archived model outputs are available at selected stations in the United States as Forecast Outputs for the United States (FOUS). This paper uses 6-h precipitation amounts and model outputs at a spatial resolution of 48 km. At this time, though, some versions of the Eta are updated at 6-h intervals, and 3-h precipitation accumulations are available in grid format at a 12-km spatial resolution.

b. WSR-88D rainfall

We have used radar rainfall from the WSR-88D network, which consists of S-band WSR-88D, radars. The geographical coverage includes the continental United States. There is continuous spatial and temporal coverage. Detailed analysis is performed on the Hydrologic Rainfall Analysis Project (HRAP) grid. The grid size in the HRAP region is approximately 4 km $\times$ 4 km, the

TABLE 1. The precipitation events used for calibration and verification. The details of the six cases or precipitation events (hours used for calibration and validation) used for model building and verification are listed. Sixteen hours of stage-III data were used for each storm. The first 10 h were used for calibration and model building, while the last 6 h were used for verification and forecast generation.

Primary date	Calibration—Start hour	Verification—End hour
27 Apr 1988	04269823UTC	04279814UTC
4 May 1999	05049915UTC	05059906UTC
16 Jun 1999	06169908UTC	06169923UTC
20 Feb 1997	02209712UTC	02219703UTC
5 Oct 1998	10049822UTC	10059813UTC
17 Oct 1998	10179811UTC	10189801UTC

temporal resolution is 1 h. In the United States, the highest quality gauge-corrected radar precipitation estimates are archived in the ABRFC, especially for the regions in and around Oklahoma. Stage-III data are calibrated using ground-based rain gauges. Young et al. (2000) evaluate WSR-88D data for operational hydrologic forecasting. To ensure that the calibration and validation was limited to the ABRFC area of responsibility, a central portion of the domain was considered during the analysis.

c. Precipitation events

Calibrated stage-III WSR-88D radar data were used for six storms in the ABRFC. The three ''winter'' events were for 20 February 1997, 5 October 1998, and 17 October 1998. The three ''summer'' events were for 27 April 1998, 4 May 1999, and 16 June 1999. A total of 16 hourly maps were used for each event, 10 for calibration, the 10th for initialization, and the rest for 1–6-h forecast generation and verification. Table 1 lists the details of these cases (hours used in calibration and validation).

4. Hybrid model

a. Overview

The distributed QPF problem has been decomposed into four component processes. The motivation for the decomposition was to optimize the use of available information from radar and NWP, and to make the best use of the available process physics and data-dictated tools. These were guided by a detailed literature review, physical insights, and preliminary data analyses (not shown here). A schematic flowchart for the proposed hybrid modeling strategy is shown in Fig. 1, while Fig. 2 depicts the individual component processes.

The component processes are modeled in succession, with each component superimposed on the results of the previous components. The first component is radar extrapolation, which is handled through advection of the hourly radar maps. The second component is large-scale physics, which disaggregates NWP-based QPF in space and time using the results from extrapolation. The third component is localized evolution, which models the change in rainfall intensity at individual pixels using an ANN-based approach. This component handles the evolution in intensity that occurs over and above large-scale physics and advection processes. The final component is residual structures, which combines the forecasts from the individual components using error statistics. The output of the proposed approach is a distributed QPF and a PQPF, that is, forecasts for rainfall intensities in space and time and a measure of the uncertainty associated with these forecasts. Details of each component follow.

1) RADAR EXTRAPOLATION COMPONENT

Figure 2a depicts the radar extrapolation component. A radar map at a given time t is compared to the lagged map at time $(t - 1)$. A two-dimensional spatial correlation is calculated, by ''moving'' the current map over the lagged map. The maximum correlation obtained in this fashion provides an estimate for the velocity scale at any time t. The metric (for 2D covariance) to be maximized takes the form

$$S = \sum \{[X(i - i_g, j - j_g, t) - \langle X(t)\rangle] \cdot [X(i, j, t - 1) - \langle X(t - 1)\rangle]\}. \quad (1)$$

In (1), $X(t)$ represents the rainfall intensity at pixel (i, j) in time t, and $\langle X(t)\rangle$ represents the spatial average of $X(t)$, over all radar pixels. The two-dimensional correlation in space is calculated in terms of the metric S. The values of (i_g, j_g) that maximize the metric S are denoted by (i_{gM}, j_{gM}). These provide estimates of the velocity scales for time t.

Velocity could similarly be estimated for all times t in the past, in the context of the precipitation event. The time series of estimated velocities are used for forecasting into the future, using a single exponential smoothing formulation:

$$i_{\mathrm{gMF}} = \alpha i_{\mathrm{gM}}(t) + \alpha(1 - \alpha) i_{\mathrm{gM}}(t - 1) + \cdots + \alpha(1 - \alpha)^n i_{\mathrm{gM}}(t - n), \quad (2)$$

$$j_{\mathrm{gMF}} = \alpha j_{\mathrm{gM}}(t) + \alpha(1 - \alpha) j_{\mathrm{gM}}(t - 1) + \cdots + \alpha(1 - \alpha)^n j_{\mathrm{gM}}(t - n). \quad (3)$$

The parameter α is the exponential smoothing parameter. It determines the decay applied to the weights for estimates further in the past. The parameter is less than unity, so that the most recent velocity estimate gets the highest relative weight. Note that in the single exponential smoothing strategy (Mills 1990), the velocity scale remains identical for all lead times. This parameter was estimated using the calibration portion of the data.

The radar forecasts are initiated with the current map, which is translated in the future for each radar pixel, using the forecasts for the velocity scales:

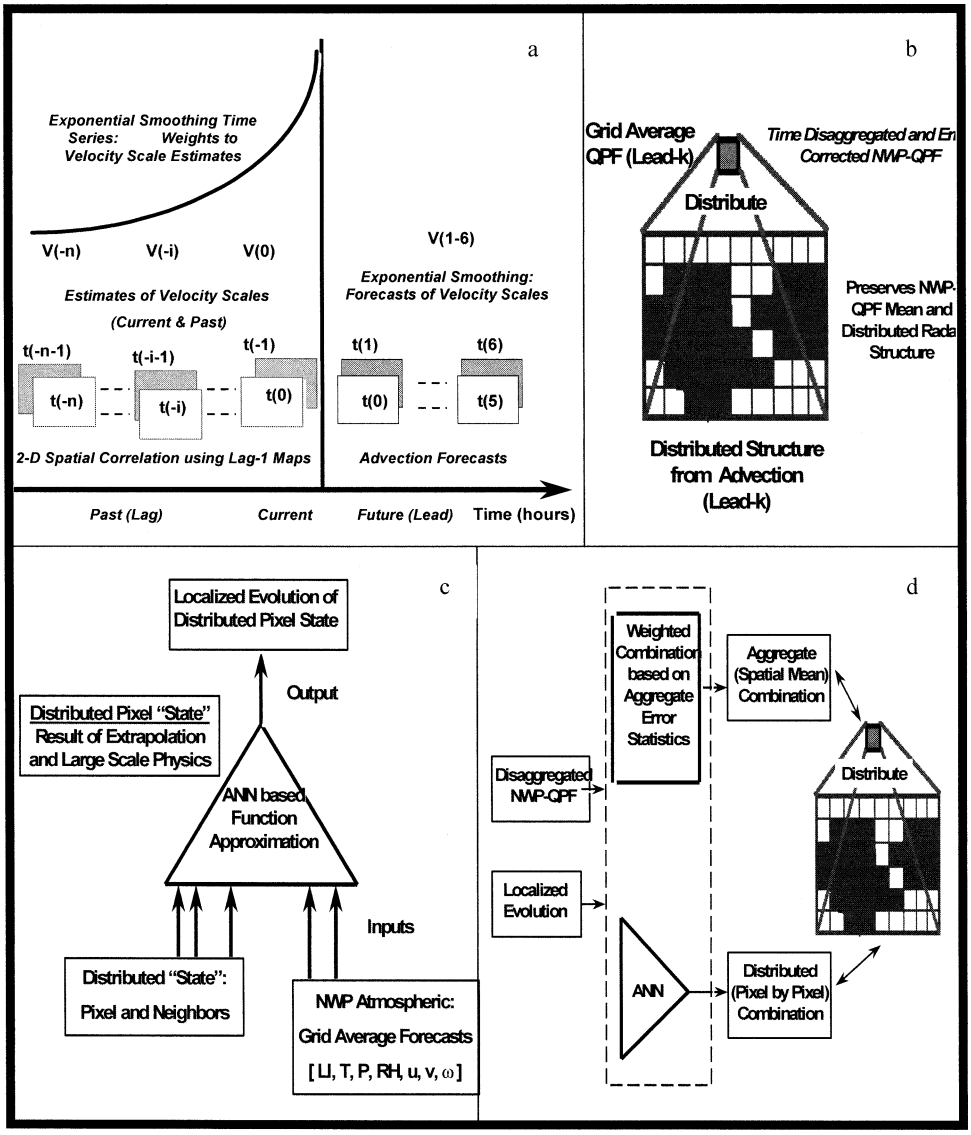

FIG. 2. Details of the four component processes of the proposed hybrid strategy for: (a) radar extrapolation, (b) Spatial disaggregation aspect of the large-scale physics component, (c) the artificial neural network strategy used for localized evolution, and (d) overview of the residual structures component.

$$X^g(i, j, t + t_L) = X(i + i_{\mathrm{gMF}}, j + j_{\mathrm{gMF}}, t) \quad \text{for all } (i, j). \tag{4}$$

The extrapolation strategy does not use NWP wind forecasts. Preliminary studies indicated that the wind available from FOUS at a single level and at 6-h resolutions have no significant correlation with the calculated advection velocities and do not improve velocity forecasts.

2) LARGE-SCALE PHYSICS COMPONENT

(i) Temporal disaggregation

The QPF from NWP represents a cumulative forecast for 6 h, so this needs to be converted to hourly rainfall intensities. The first step is to temporally disaggregate the NWP-based QPF from 48-km and 6-h resolution to 48-km and 1-h resolution in the following manner:

$$r(t + k \times \Delta t) = \{[(N_T + \Delta N_T) - (t + k \times \Delta t)][R(N_T + \Delta N_T)/6] + [(t + k \times \Delta t) - N_T][R(N_T)/6]\}/\Delta T. \tag{5}$$

In (5), r represents disaggregated 1-h precipitation prediction from the NWP over the NWP grid, 48 km $\times$ 48 km pixels. In (5), t is the current time, Δt the radar time step (1 h), and k an integer. The left-hand side of the equation represents the aggregate forecasts for time steps at lead k. On the right-hand side, N_T is the current (or last available) NWP time, $(N_T + \Delta N_T)$

is the next available NWP time, and R is the average (6 h) NWP-based QPF. This assumes that the forecast lead times lie within the intervals for which NWP outputs are available. The hourly QPF obtained from (5) are compared with spatially aggregate radar rainfall. The errors are modeled and forecast as an AR time series. The time series of the spatially aggregated errors is $Z(\tau) = [\langle X(t - i\Delta t)\rangle - r(t - i\Delta t)]$, where i is the index for times in the past, $\tau = (t - i\Delta t)$, and $\langle X\rangle$ is the spatial mean of the observed radar rainfall. The AR model for error correction (at lead k) is $Z_{t+k} \cong \alpha_1 Z_{t+k-1} + \alpha_2 Z_{t+k-2} + \cdots + \alpha_p Z_{t+k-p}$.

Note that $\Delta t = 1$ h, and forecasts generated at each lead time could be used in the AR formulation for forecasts at successive lead times. The forecasts for the spatially averaged hourly forecast following error correction, F, becomes

$$F(t + k) = r(t + k) + Z_{t+k} = r(t + k) + \sum \alpha_i Z_{t+k-i}. \quad (6)$$

The value of the autoregressive window size and the corresponding parameters need to be estimated from the calibration portion of the data. In this study, the window size (p) was restricted to a value of two from preliminary data analyses, while the parameters are estimated dynamically. Note that error correction is used only if the forecasts are improved at analysis time, which is checked by forecasting the errors at cross-validation time steps.

The correction described in (6) is an additive correction and is used to adjust raining pixels (i.e., pixels with nonzero rainfall intensities based on advection) only, as described next.

(ii) Spatial disaggregation

The temporally aggregated and error-corrected QPF obtained from (6) is used to scale the distributed map obtained from radar advection (Fig. 2b). The pixels with advected radar rainfall below a minimum threshold (the value used in this study is 0.001 mm) are left unaltered, while those with nonzero rainfall are changed to reflect the grid average QPF. Note that only pixels with nonzero values are considered when computing the grid average QPF value. To preserve the nature of grid-averaged NWP-based QPF, the nonzero pixels are modified equally using the average deficit or excess grid average rainfall, as indicated by (7),

$$\text{If} \quad X^g(i, j, t + k \times \Delta t) > \varepsilon$$
$$X^D(i, j, t + k \times \Delta t) = X^g(i, j, t + k \times \Delta t) + \frac{N_{\text{pixel}}[F(t + k \times \Delta t) - \langle X^g(t + k \times \Delta t)\rangle]}{N_{\text{wet-pixel}}}. \quad (7)$$

In (7), N_{pixel} is the total number of radar pixels, $N_{\text{wet-pixel}}$ is the total number of radar pixels with nonzero rainfall, and X^D is the distributed rainfall forecast from this disaggregation strategy. To prevent the occurrence of unreasonably large or small changes, the corrected results were restricted to fixed ratios (0.5 to 2.0 for the cases in this study) of the original.

3) LOCALIZED EVOLUTION COMPONENT

We assume that there is a functional form that can model the evolution of the distributed rainfall state (after advection and scaling), as shown in Fig. 2c,

$$d\mathbf{X}/dt = \Phi^{(\Omega)}(\mathbf{X}, \mathbf{X}_N, \boldsymbol{\Theta}^{(\Omega)}, \boldsymbol{\Omega}) + \boldsymbol{\eta}. \quad (8)$$

In (8), $\mathbf{X}$ represents the vector of the distributed state in a Lagrangian reference frame conditioned on and scaled by NWP-based QPF. This evolution of $\mathbf{X}$ is assumed to be a function of the state $\mathbf{X}$, and the state at neighboring pixel states, $\mathbf{X}_N$. We assume that the functional form is static for each NWP grid for each precipitation event, and is dependent on the large-scale atmospheric state $\boldsymbol{\Omega}$. Finally, we assume a set of parameter vectors $\boldsymbol{\Theta}^{(\Omega)}$, also a function of $\boldsymbol{\Omega}$, derived from NWP model outputs. The term $\boldsymbol{\eta}$ denotes model errors, which account for the processes that occur at resolutions greater than radar observations.

To model the evolution in (8), we use data-dictated ANN techniques that do not make a priori assumptions about the functional form or the parameters relating the input vectors to the outputs. The ANN model is expressed as (Fig. 2c)

$${}^E X - {}^D X = f_{\text{ANN}}({}^D X, {}^D X_N; \text{LI}, P, T, \text{RH}, u, v, \omega) + \boldsymbol{\eta}. \quad (9)$$

In (9), ${}^E X$ denotes the effects of evolution and ${}^D X$ denotes the pixel state in a Lagrangian frame of reference conditioned on large-scale NWP-based QPF [from Eq. (7)]. The term ${}^D X_N$ denotes the state of the neighboring pixels of the scaled advection map.

NWP model outputs other than QPF were also disaggregated in time. Spatial disaggregation was not attempted for these outputs because there were no high-resolution forecasts or measurements to serve as bases for disaggregation, and the atmospheric outputs are more uniform in space than QPF (Antolik 2000; Mesinger 1996). The NWP model outputs used are the LI, sea level pressure (P), average air temperature (T), average relative humidity (RH), and wind vector components (u, v, and ω). We include all of these as ANN inputs because we cannot reject a priori the hypothesis that these do not have information pertinent for distributed rainfall. The implicit assumption is that the manner in which these variables influence rainfall can be deduced in a data-dictated fashion by the ANN.

The ANN used was a modified form of the Nonlinear Autoregressive moving average (NARMA) model (Connor et al. 1994). NARMA models handle the autoregressive and moving average components of a tradi-

tional Autoregressive Moving Average (ARMA) time series model (Mills 1990), using nonlinear ANN functions instead of linear formulations. We extended the NARMA implementation such that the nonlinear autoregressive (NAR) and nonlinear moving average (NMA) components were separated out and modeled using individual multilayer perceptron (MLP; see Bishop 1996). This implies a linear separation of the autoregressive and the moving average components, which results in some loss of generality but still includes models like the Autoregressive-Integrated Moving Average (ARIMA) and the nonlinear Threshold Autoregressive (TAR) as special cases. The NMA is used only if forecasts are improved at analysis times on a cross-validation data.

The ''Bayesian ANN'' methods of MacKay (1994) and the implementation strategy suggested therein are used and slightly extended to generate ''most probable'' rainfall forecasts (which is not necessarily the mean of the forecast ensemble) and corresponding error bars. This method first trains an ANN on calibration (stage-III rainfall) data and calculates the error statistics on cross-validation data, and then uses these statistics to generate an ensemble of identically distributed random-error simulations. Realizations from the random simulations are then summed with the observed output variables to generate ''pseudo outputs.'' An ensemble of ANN models is obtained by retraining the originally trained ANN with each of the pseudo outputs. This ensemble of models provide ensemble forecasts at each verification time step, which could then be used to calculate the expected error statistics and, hence, the confidence bounds. Training individual members of the ensemble is not too computationally expensive, because these members are initialized with previously calibrated ANN. However, to reduce forecast generation time, only the most likely forecasts (obtained from the original ANN), and the upper and lower confidence bounds, were used for forecast generation at successive time steps. The number of training time steps, the neighboring window size (i.e., the number of neighboring pixel states used as inputs), and the optimal MLP architecture were chosen from trial and error using subsets of the data. The selected training times were lower for winter storms compared to summer storms, reflecting the higher degree of nonlinearity in the latter. The functional form to be modeled by the ANN was assumed to remain invariant within the context of the localized evolution event for the training and forecast generation times in the domain of interest. The pixels in space at each training time step were divided randomly into ''training'' and ''cross validation'' data, with the ratio of roughly 2:1.

To summarize, the localized evolution model calculates the rainfall intensities at individual pixels through an ANN-based nonlinear function approximation. This function approximation is conditioned on the advected radar map scaled with spatially averaged NWP-based QPF. The inputs to the ANN model are the NWP forecasts of atmospheric variables and the scaled and advected rainfall intensities.

4) RESIDUAL STRUCTURES COMPONENT

Figure 2d depicts the residual structures component. This component attempts to combine the complementary skills of the large-scale physics and localized evolution components.

Scaling the advection maps with low-resolution NWP-based QPF yields a set of distributed forecasts for precipitation. Superimposing the localized evolution component provides another. The first procedure is referred to as a ''disaggregation model,'' while the second is an ''evolution model.'' This ANN-based model handles localized evolution conditioned on the disaggregation model. Both approaches consider advection as a baseline. The first strategy preserves the distributed radar structure, as well as the aggregate information from NWP, while the second tries to capture the localized evolution at smaller scales in a Lagrangian frame of reference. If the localized evolution component were to consistently improve forecasts, then the evolution model would always outperform the disaggregation model. As indicated in Fig. 3, however (ignoring the results of the ''proposed'' strategy for the purposes of this discussion), the evolution component results in significant improvement in certain forecast scenarios, but results in loss of information in others. This behavior is not unexpected, and is elaborated on later.

Figure 3 shows the skills from the individual component processes as a function of lead time. These are discussed in detail later; however, the complementary skills from the large-scale physics and the localized physics components are useful for the purposes of this discussion. We observe from our preliminary data analyses (not shown) and from Fig. 3 that the strategy that combines extrapolation and NWP-based QPF does relatively better in the winter in terms of aggregate statistics, and for longer lead times. The approach that models high-resolution evolution using NWP atmospheric variables and ANN exhibits relatively better skills in the summer (not shown), at distributed scales, and for shorter lead times ($\sim$1–3 h). This indicates that there are dominant residual structures, both at distributed and at aggregate scales, as a result of each strategy. We hypothesize that an approach that captures the distributed errors and preserves the aggregate information, by combining the results of both these individual strategies, could be an improvement over either.

To achieve a distributed combination, we use a data-dictated ANN approach that regresses the observed rainfall on the results of both the disaggregation model and the evolution model. The combination might need to be nonlinear, so the use of ANN precludes the necessity to assign a functional form a priori.

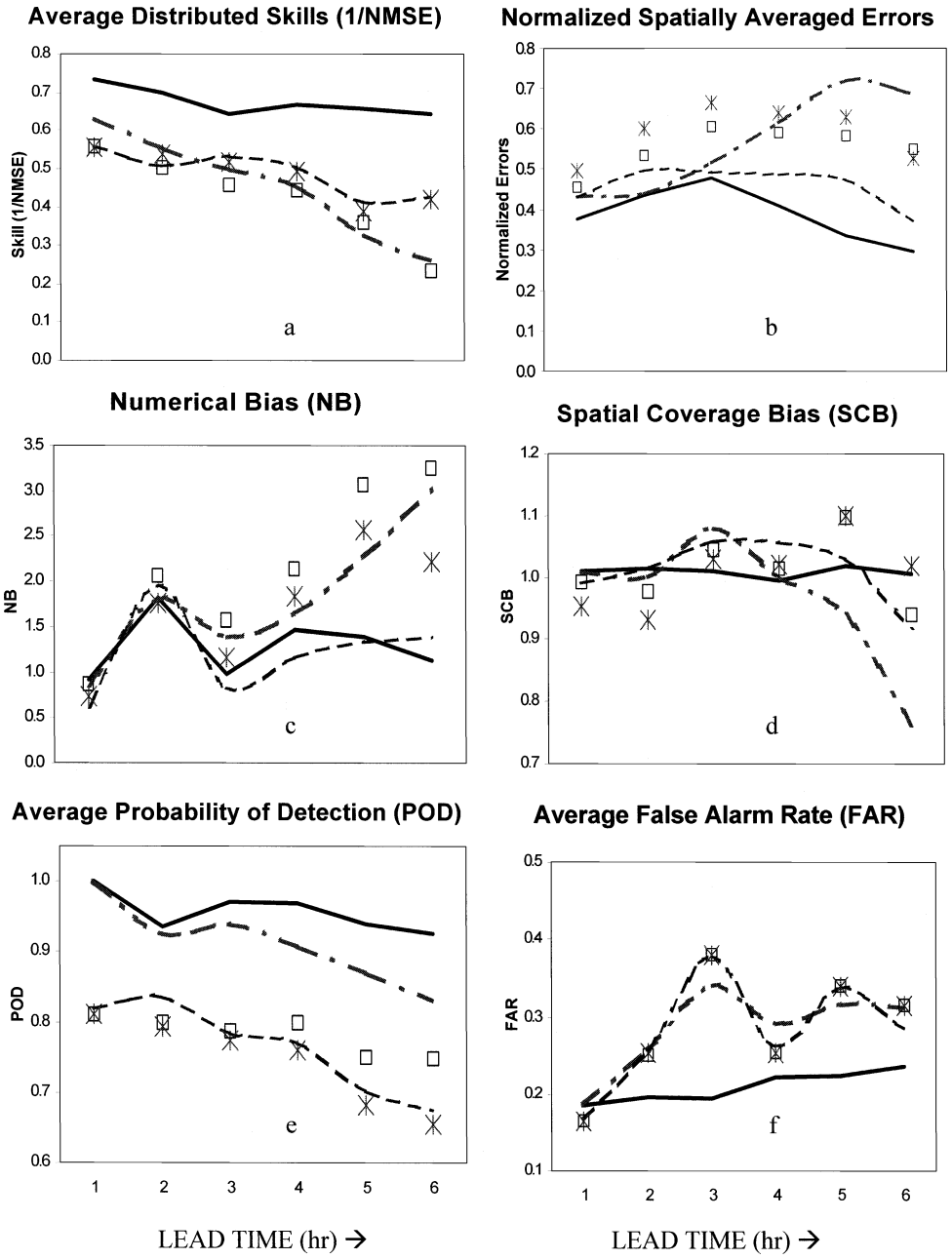

FIG. 3. The QPF skills as a function of lead time are shown. The proposed QPF strategy has been compared with existing methods and component processes. QPF from persistence (asterisks), advection (squares), large-scale physics (dashed lines), localized evolution (dotted-dashed lines), and the proposed strategy (bold lines). The performance in terms of a selected skill score is shown for (a) 1/NMSE, (b) normalized spatially averaged errors, (c) numerical bias, (d) spatial coverage bias, (e) POD, and (f) FAR.

$$
\begin{aligned}
{}^{DE}X(i, j, t) - {}^{A}X(i, j, t) \\
= g_{\mathrm{ANN}}\{[{}^{D}X(i, j, t) - {}^{A}X(i, j, t)], \\
[{}^{E}X(i, j, t) - {}^{A}X(i, j, t)]\} + \boldsymbol{\eta}. \qquad (10)
\end{aligned}
$$

In (10), ${}^{A}X$ is the result of advection alone, ${}^{D}X$ the scaled advection, and ${}^{E}X$ the localized evolution. The result of pure advection is removed prior to the ANN-based function approximation, and added back to the outputs later. This preprocesses the data to make the ANN-based function approximation task easier, and improves performance. The details of the ANN formulation were similar to the one used for localized evolution. Note that ${}^{DE}X$ denotes the ''distributed combination'' of the disaggregation and the evolution strategies.

To retain the optimal spatial aggregate information,

we combine the spatial mean from the disaggregation and evolution models by assigning relative weights. These weights are calculated from the aggregate error statistics of the individual components:

$$\langle {}^{O}X\rangle = \frac{[(1/\sigma_D) \times \langle {}^{D}X\rangle + (1/\sigma_{\mathrm{DE}}) \times \langle {}^{\mathrm{DE}}X\rangle]}{[(1/\sigma_D) + (1/\sigma_{\mathrm{DE}})]}. \quad (11)$$

In (11), $\langle {}^{O}X\rangle$ is the overall spatial aggregate, ${}^{D}X$ represents scaled advection, ${}^{\mathrm{DE}}X$ represents the combined map that accounts for distributed errors, σ_D is the aggregate error (rmse) for the scaled advection, σ_{DE} is the aggregate error of the distributed combination obtained from the Bayesian ANN strategy, described earlier.

The distributed QPF map obtained by combining the results of the disaggregation and evolution models in (10) is assumed to best account for the errors at distributed scales. The weighted aggregate forecast in (11) is assumed to best preserve the information at spatially aggregated scales. We combine these by scaling the distributed map with the estimate for the aggregate spatial mean. The details of the scaling algorithm are similar to the spatial disaggregation strategy for NWP-based QPF described earlier in (7).

The scaling takes the form

$$\text{If } {}^{\mathrm{DE}}X(i, j, t + k \times \Delta t) > 0 \quad \text{(or, a small threshold } \varepsilon\text{)},$$

$$X^{C}(i, j, t + k \times \Delta t) = {}^{\mathrm{DE}}X(i, j, t + k \times \Delta t) + \frac{N_{\mathrm{pixel}}[\langle {}^{O}X(t + k \times \Delta t)\rangle - \langle {}^{\mathrm{DE}}X(t + k \times \Delta t)\rangle]}{N_{\mathrm{wet\text{-}pixel}}}. \quad (12)$$

In (12), X^{C} represents the final combined map.

5. Distributed QPF skill metrics

We considered several skill measures that might be of interest for hydrologists or meteorologists during forecast evaluation. Besides statistical measures of skill, it might be of interest to visually inspect the improvement in distributed QPF from the precipitation maps.

The following statistical measures of skill were considered and averaged across the precipitation events:

1) The inverse of the normalized root-mean-square errors, or 1/NMSE, for pixels with nonzero rainfall. A perfect forecast gives 1/NMSE = ∞. For a stationary process, 1/NMSE = 1 implies the forecast is no better than the mean of the verification data. However, hourly precipitation processes usually exhibit significant nonstationarities. The root-mean-square errors were normalized with the standard deviation of the verification (forecast) datasets. Only pixels with nonzero-observed rainfall were considered in the calculations. This statistic provides a measure for the distributed skill in an averaged sense.
2) The normalized spatially averaged errors as a function of lead time. Higher errors indicate lower skills. This provides an estimate of the spatially averaged forecast bias. This is not to be confused with the numerical bias skill score, which is usually calculated as a ratio, and is considered next.
3) The numerical bias of the forecasts defined as the mean forecast value divided by the mean observed value based on values within all pixels. Values closer to unity indicate lower bias, while higher or lower values indicate greater bias and worse forecasts.
4) Spatial coverage bias defined as the number of forecast pixels with rainfall greater than, or equal to, 0.2 mm divided by the number of observed pixels with rainfall greater than, or equal to, 0.2 mm. Values closer to unity indicate lower coverage bias, while higher or lower values indicate greater coverage bias and worse forecasts.
5) Statistics on the probability of detection (POD) with respect to 0.2-mm precipitation. The POD is defined as the ratio of the number of ''hits'' (forecast of rain given rain occurred) to the total number of hits and ''misses'' (forecast of no rain given that rain occurred). The POD varies from 0 to 1, with 0 representing no skill in detection (no hits) and 1 representing perfect skill (no misses).
6) Statistics on the false alarm ratio (FAR) with respect to 0.2-mm precipitation. The FAR is defined as the ratio of the number of ''false alarms'' (forecast of rain given that no rain occurred) to the total number of false alarms and hits (forecast of rain given that rain occurred). The FAR varies from 1 to 0, with 1 representing no skill (high rate of false alarms) and 0 representing perfect skill (no false alarms).

6. Results and discussion

The proposed strategy for distributed QPF comprises four components, which are applied in sequence. This section summarizes the performance of the individual component processes in terms of the distributed and aggregate error statistics, and/or measures of skills. For comparison, results from persistence are also presented. Each error or skill measure (shown in Figs. 3a–f) quantifies a slightly different aspect of the forecast performance. In general, extrapolation appears to perform marginally better than persistence. Large-scale physics and localized evolution appear to exhibit complementary skills (other than for the POD and FAR). Within the 1–6-h lead times, localized evolution appears to perform better than large-scale physics at lower (1–2 or 3 h) leads, while the latter performs better for higher lead times. The proposed hybrid approach improves over the existing DQPF strategies and component processes compared in this study. For the cases considered in this study, the proposed strategy demonstrates the ability to successfully blend the complementary skills from large-

scale physics and localized evolution, which in turn results in improvements over advection and persistence.

Occasional outliers were removed during the calculations of errors or skill measures. For a given lead time, if the skill or error metrics deviated sharply from the overall trend, and if that deviation was caused by just one storm, the specific value was treated as an outlier. Typical adjustments entailed substituting the outliers with the average values from the adjacent time steps. For example, in the winter, the distributed skills from persistence at 2-h leads were unexpectedly high, and this was caused by one storm. This aberration in the skills from persistence caused the other strategies to exhibit a similar aberrant behavior. The distributed skills at 2-h leads in the winter, from all the forecasting strategies, were adjusted by replacing with the average of the skills at 1- and 3-h leads. A similar adjustment was done for the distributed skills in the summer at 3-h leads. The results from one storm (16 June 1999) were ignored from the 1/NMSE calculations at leads greater than 1 h. This storm decayed suddenly in intensity at certain locations after about 1 h, which caused the distributed skills to exhibit anomalous behavior.

a. 1/NMSE

As shown in Fig. 3a, the average distributed skill (in terms of 1/NMSE) for all seasons combined exhibits a low but nearly steady rate of decay with lead time in the 1–6-h range. The existing methods show some skills, especially at short lead times. In terms of this skill metric, 1-h advection does not improve over persistence, primarily due to occasionally missed cells and patterns not being preserved for hourly rainfall. The large-scale physics method that combines advection and NWP-based QPF usually improves over pure advection or persistence in terms of aggregate skill (Fig. 3b), but the difference is marginal in terms of distributed skill (Fig. 3a). For 1-h resolutions and 1–6-h leads, the skills from 1-h advection and NWP-based QPF have been found to be comparable (Golding 2000). Further, the process of distributing the low-resolution NWP-based QPF in space and time introduces significant errors. The localized evolution strategy appears to exhibit greater skill than large-scale physics until about 2–3 h, and decays sharply thereafter. The relative skills from the large-scale physics and localized evolution components vary by forecast leads, resolution, and season. The proposed method significantly improves over existing techniques, like persistence, advection, and advection combined with NWP-based QPF in terms of the distributed skill measured by 1/NMSE.

b. Spatially averaged errors

Figure 3b shows the spatially aggregated normalized errors as a function of lead time averaged for all seasons. Advection marginally improves over persistence, while large-scale physics improves over advection. On the average, localized evolution marginally improves over large-scale physics for all seasons only until about 2–3-h lead time, but performs worse (relative to large-scale physics) thereafter. The proposed strategy beats all other approaches at 1–6-h leads.

c. Numerical bias

Figure 3c shows the spatially aggregated numerical bias as a function of lead time. For about 1–3-h lead, all of the DQPF strategies perform equally well relative to each other. However, after around 3-h lead, the skill from persistence and advection quickly grow worse, and localized evolution follows this trend. The skill from large-scale physics is better at 4–6-h leads, relative to persistence, advection, and localized evolution. The skill from the proposed strategy closely follows that of large-scale physics.

d. Spatial coverage bias

The bias in the spatial coverage is shown in Fig. 3d. For 1–2-h leads, both large-scale physics and localized evolution appear to show improvements over advection, which, in turn, is better than persistence. In terms of this measure, forecast skill from localized evolution is significantly worse at 5- and 6-h leads. The proposed method appears to perform better than all existing DQPF strategies considered in this study, as well as the individual component processes.

e. POD

Figure 3e shows the POD as a function of lead time. Large-scale physics does not improve over persistence or advection. At 4–6-h leads, advection appears to perform marginally better than persistence. Both the localized evolution and the proposed hybrid strategy perform significantly better than the other strategies, which demonstrate the ability of these methods to predict rain versus no-rain areas. The proposed strategy performs the best overall.

f. FAR

Figure 3f shows the FAR as a function of lead time. The performance from persistence, advection, large-scale physics, or localized evolution does not exhibit a significant or consistent difference relative to each other in terms of this measure. The proposed method exhibits improvements over all existing strategies in terms of this measure. However, the behavior at 1-h lead (where both localized evolution and the proposed strategy perform slightly worse compared to the other three methods) appear to be a slight anomaly.

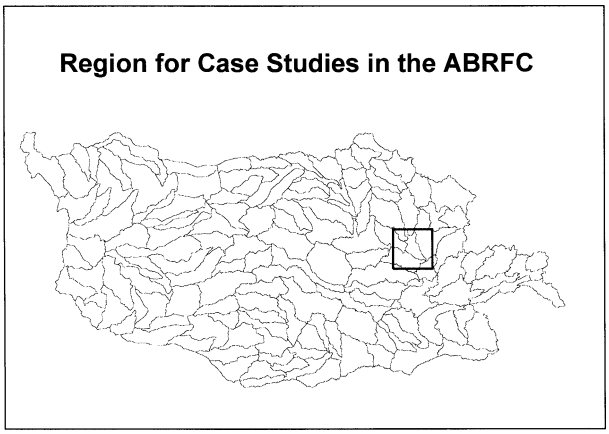

FIG. 4. The region within the ABRFC used for the case studies; the spatial extent of the study region lies within the ABRFC area of responsibility.

g. Qualitative performance

We have used hourly rainfall maps for model development and calibration, and hourly maps for verification and forecast generation. Previous research and data analysis indicate that the spatially distributed structure is not very well preserved for hourly rainfall. The spatial extent of the study region within the ABRFC region is indicated in Fig. 4. Figure 5 illustrates that the proposed strategy still misses much of the detailed spatial structure, but improves over persistence and advection. It should be noted, though, that contouring can be misleading.

7. Conclusions

Distributed QPF has diverse and important applications. In the area of flood forecasting, this has the potential of saving human lives and property. Improving distributed QPF also happens to be among the most challenging problems in hydrology and meteorology. Availability of better NWP models and high-quality radar observations, as well as emerging data-dictated tools, offer a window of opportunity for improving distributed QPF. The presented hybrid modeling strategy improves distributed QPF by making the best use of information from radar measurements and NWP model outputs, as well as the available precipitation physics and data-dictated tools. The overall strategy consists of four elements that need to be applied in sequence. These are described as (i) radar extrapolation, (ii) large-scale physics, (iii) localized evolution, and (iv) residual structures.

High-quality WSR-88D data were used in conjunction with 48-km Eta NWP model outputs to demonstrate improvements in distributed QPF. The proposed hybrid model improves distributed QPF over techniques like radar extrapolation alone, NWP-based QPF alone, and hybrid models that combine radar extrapolation with NWP-based QPF. This is indicated through skill scores and error measures. Case studies (not shown) also indicated that the hybrid model would perform better than pure data-dictated tools or parameterized physically based models.

Although not discussed, it is important to note that this research confirmed the seasonal dependence of the ability to quantitatively predict rainfall. Predictability is higher in the winter than in the summer. In the winter, QPF from NWP has more information content than other NWP model outputs at aggregate scales, and linear modeling strategies seem to perform relatively well. The predictability of the space–time-distributed QPF, and the spatially aggregated QPF at short leads, could be improved in the summer using nonlinear strategies that exploit the information content in the NWP forecasts of atmospheric variables, like in the suggested hybrid strategy.

The hybrid strategy proposed in this paper postprocesses the information already available from NWP and radar, and uses statistical approaches to model distributed physical processes that cannot be easily handled by the current generation of NWP models. It has been argued that no method has succeeded in consistently improving distributed QPF (at scales useful for hydrologic applications) over and above Lagrangian persistence at very short (e.g., 1–2 h) lead times, although QPF from NWP has been found to be useful at slightly higher leads (e.g., 3–6 h). Based on the results presented here, the new hybrid strategy appears to have the potential to improve over these approaches, even though further validation is necessary.

Researchers have speculated that the future of QPF lies in numerical modeling at increasing resolution. There is a need to investigate the value of these approaches to generate DQPF at hydrologic scales and to compare them to methods like the proposed hybrid strategy that attempt to statistically combine observations and numerical model outputs. Statistical methods for initialization, nudging, assimilation, or validation of numerical weather models are somewhat complementary to statistical postprocessing approaches, and it is possible that significant improvements in DQPF can be realized by strategies that blend these approaches.

Acknowledgments. This work was partially supported through the National Weather Service (NOAA, U.S. Department of Commerce) Award NA07WH0437 as part of the Cooperative Agreement on the Comparison of Distributed and Lumped Hydrologic Models, and the Consiglio Nazionale delle Ricerche of Italy as part of the Cooperative Agreement in Climate Change and Hydrogeological Disasters. We would like to thank Profs. Dennis McLaughlin, Amar Gupta, Elfatih Eltahir, and Dara Entekhabi of the Massachusetts Institute of Technology, Prof. Shafiqul Islam of the University of Cincinnati; as well as Prof. Ana Barros of Harvard University and her collaborator Dr. Robert Kuligowski, for

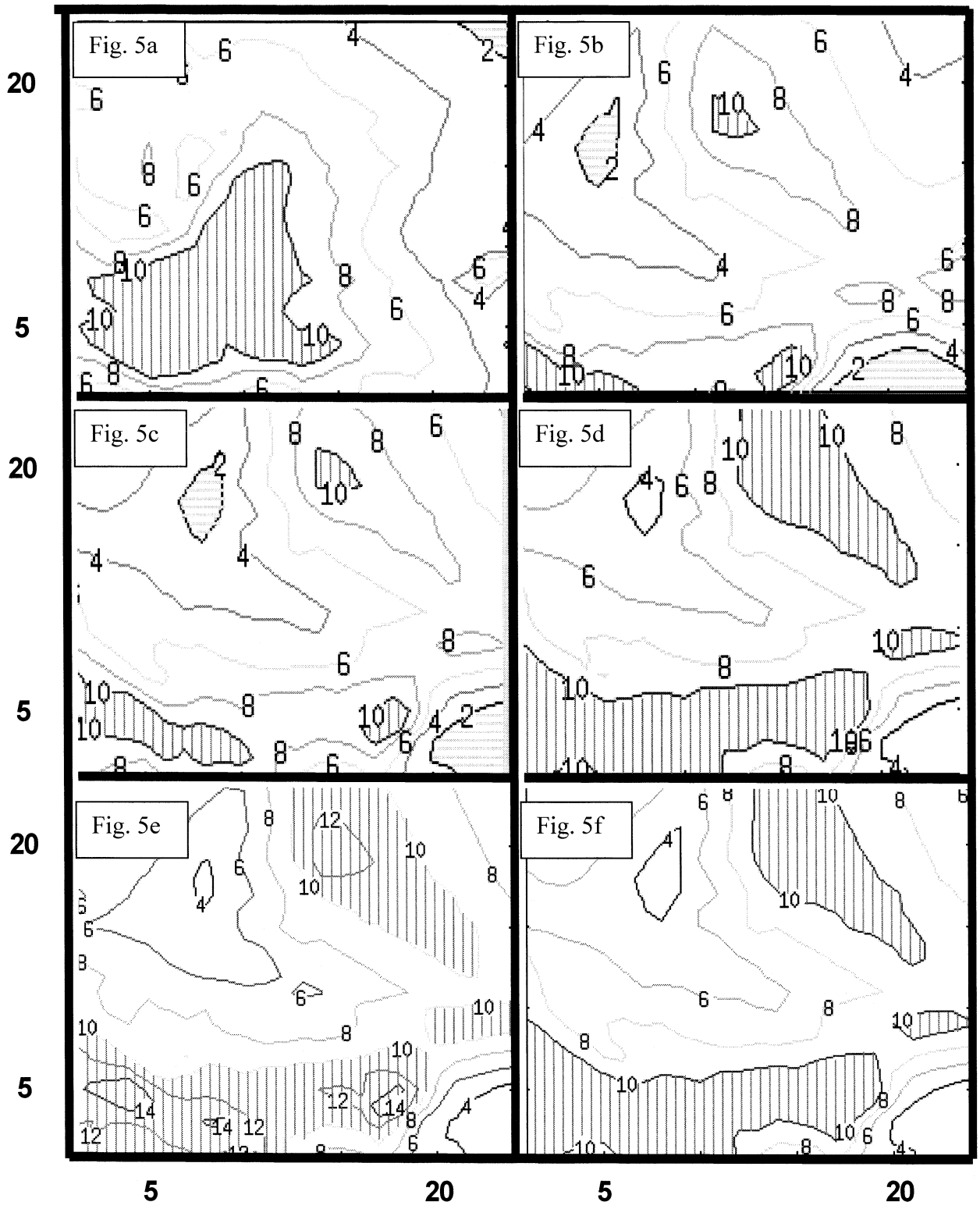

FIG. 5. Contour plots showing observed radar rainfall and QPF at 1-h lead for the storm of 27 Apr 1998. The *x* and *y* axes are in pixel units while rainfall intensity is in millimeters per hour. Contours higher than 10 mm are indicated with vertical stripes, while contours lower than 2 mm are indicated with horizontal stripes. The contours represent either observed or forecasted rainfall: (a) observed stage-III radar rainfall, (b) QPF from persistence, (c) QPF from advection, (d) QPF from the proposed strategy, (e) forecasted upper bounds, and (f) forecasted lower bounds.

their suggestions and help. We gratefully acknowledge the comments from the three reviewers, which significantly improved the quality of the paper. The first author would like to thank his family and friends, as well as his colleagues and management at Oracle Corporation where he was employed for the last 3 yr of his doctoral research. Both the authors would like to acknowledge the help from graduate research assistants at MIT in hydrology, who are currently collaborating with Prof. Rafael Bras, for producing Fig. 4.

REFERENCES

Antolik, M. S., 2000: An overview of the National Weather Service's centralized statistical quantitative precipitation forecasts. *J. Hydrol.,* **239,** 306–337.

Bishop, C. M., 1996: *Neural Networks for Pattern Recognition.* Oxford University Press, 256 pp.

Black, T. L., 1994: The use of NMC mesoscale eta model: Description and forecast examples. *Wea. Forecasting,* **9,** 265–278.

Brewster, K., 1998: Phase correcting assimilation of radar data for thunderstorm forecasting. Preprints, *12th Conf. on Numerical Weather Prediction,* Phoenix, AZ, Amer. Meteor. Soc., 181–184.

Collier, C. G., 1991: The combined use of weather radar and mesoscale numerical model data for short-period rainfall forecasting. *Hydrological Application of Weather Radar,* I. D. Cluckie and C. G. Collier, Eds., E. Horwood, 644 pp.

——, and R. Kzyzysztofowicz, 2000: Preface: Quantitative precipitation forecasting. *J. Hydrol.,* **239,** 1–2.

Connor, J. T., R. D. Martin, and L. E. Atlas, 1994: Recurrent neural networks and robust time series prediction. *IEEE Trans. Neural Networks,* **5,** 240–254.

Droegemeier, K., and Coauthors, 2000: Hydrological aspects of weather prediction and flood warnings. *Bull. Amer. Meteor. Soc.,* **81,** 2665–2680.

Elshorbagy, A., S. P. Simonovic, and U. S. Panu, 2002a: Estimation of missing streamflow data using principles of chaos theory. *J. Hydrol.,* **255,** 123–133.

——, ——, ——, 2002b: Noise reduction in chaotic hydrologic time series: Facts and doubts. *J. Hydrol.,* **256,** 147–165.

Fritsch, J. M., and Coauthors, 1998: Quantitative precipitation forecasting: Report of the eighth prospectus development team, U.S. weather research program. *Bull. Amer. Meteor. Soc.,* **79,** 285–299.

Fulton, R. A., J. P. Breidenbach, D.-J. Seo, D. A. Miller, and T. O'Bannon, 1998: The WSR-88D rainfall algorithm. *Wea. Forecasting,* **13,** 377–395.

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.

Golding, B. W., 2000: Quantitative precipitation forecasting in the UK. *J. Hydrol.,* **239,** 286–305.

Grecu, M., and W. F. Krajewski, 2000: A large-sample investigation of statistical procedures for short-term quantitative precipitation forecasting. *J. Hydrol.,* **239,** 69–84.

Hou, A. Y., S. Q. Zhang, A. M. da Silva, W. S. Olson, C. D. Kummerow, and J. Simpson, 2001: Improving global analysis and short-range forecast using rainfall and moisture observations derived from TRMM and SSM/I passive microwave sensors. *Bull. Amer. Meteor. Soc.,* **82,** 659–679.

Hsu, K.-L., H. V. Gupta, X. Gao, and S. Sorooshian, 1999: Estimation of physical variables from multichannel remotely sensed imagery using a neural network: Application to rainfall estimation. *Water Resour. Res.,* **35,** 1605–1618.

Kuligowski, R. J., and A. P. Barros, 2001: Combining IR-microwave satellite retrieval of temperature and dewpoint profiles using artificial neural networks. *J. Appl. Meteor.,* **40,** 2051–2067.

Lee, H. T., and K. P. Georgakakos, 1996: Operational rainfall prediction on meso-γ scales for hydrologic applications. *Water Resour. Res.,* **32,** 987–1003.

MacKay, D. J. C., 1994: Bayesian non-linear modelling for the prediction competition. *ASHRAE Trans.,* **100,** 1053–1062.

McLaughlin, D., S. Islam, D. Entekhabi, M. French, and R. Bras, 1990: A distributed filtering approach to real-time rainfall forecasting. Preprints, *Eighth Conf. on Hydrometeorology,* Kananaskis Park, Alberta, Canada, Amer. Meteor. Soc., 150–155.

Mesinger, F., 1996: Improvements in quantitative precipitation forecasts with the eta regional model at the National Centers for Environmental Prediction: The 48-km upgrade. *Bull. Amer. Meteor. Soc.,* **77,** 2637–2649.

Mills, T. C., 1990: *Time Series Techniques for Economists.* Cambridge University Press, 377 pp.

Nakakita, E., S. Ikebuchi, T. Kanamuri, M. Okuda, A. Yamaji, and T. Takasao, 1996: Short-term rainfall prediction model using a volume scanning radar and grid point value data from numerical weather prediction. *J. Geophys. Res.,* **101,** 26 181–26 197.

Perica, S., and E. Foufoula-Georgiou, 1996: A model for multiscale disaggregation of spatial rainfall based on coupling meteorological and scaling descriptions. *J. Geophys. Res.,* **101** (D21), 26 347–26 361.

Rodriguez-Iturbe, I., and P. S. Eagleson, 1987: Mathematical models of rainstorm events in space and time. *Water Resour. Res.,* **23,** 181–190.

Rogers, R. R., and M. K. Yau, 1989: *A Short Course in Cloud Physics.* Pergamon, 293 pp.

Smith, K. T., and G. L. Austin, 2000: Nowcasting precipitation—A proposal for a way forward. *J. Hydrol.,* **239,** 34–45.

Sugimoto, S., E. Nakakita, and S. Ikebuchi, 2001: A stochastic approach to short-term rainfall prediction using a physically based conceptual rainfall model. *J. Hydrol.,* **242,** 137–155.

Toth, E., A. Brath, and A. Montanari, 2000: Comparison of short-term rainfall prediction models for real-time flood forecasting. *J. Hydrol.,* **239,** 132–147.

Wilson, J. W., N. A. Crook, C. K. Mueller, J. Sun, and M. Dixon, 1998: Nowcasting thunderstorms: A status report. *Bull. Amer. Meteor. Soc.,* **79,** 2079–2100.

Young, C. B., A. A. Bradley, W. F. Krajewski, and A. Kruger, 2000: Evaluating NEXRAD multisensor precipitation estimates for operational hydrologic forecasting. *J. Hydrometeor.,* **1,** 241–254.