# Assessing the Capability of a Regional-Scale Weather Model to Simulate Extreme Precipitation Patterns and Flooding in Central Texas

MARLA R. KNEBL LOWREY AND ZONG-LIANG YANG

*Department of Geological Sciences, University of Texas at Austin, Austin, Texas*

## ABSTRACT

A regional-scale weather model is used to determine the potential for flood forecasting based on model-predicted rainfall. Extreme precipitation and flooding events are a significant concern in central Texas, due to both the high occurrence and severity of flooding in the area. However, many current regional prediction models do not provide sufficient accuracy at the watershed scale necessary for flood mitigation efforts. The Weather Research and Forecasting (WRF) model, created with the purpose of improving upon the current fifth-generation Pennsylvania State University–National Center for Atmospheric Research (PSU–NCAR) Mesoscale Model (MM5), is specifically designed for regional grid spacings of 1–10 km. Previous research by the authors resulted in the development of a regional-scale prediction system over the San Antonio River basin, using a geographic information system (GIS) database, a hydrologic model, and a hydraulic model. Observed precipitation drives the prediction system; the authors hypothesize that the WRF model has the potential to predict flooding, at a lead time of several days, with a level of accuracy near that of observed precipitation. Causes of model error are also investigated, to determine the relative errors caused by model physics, initialization interval, buffer zone and domain size, and small-amplitude random errors. Results show that the Betts–Miller–Janjić cumulus and Lin et al. microphysics schemes, 48-h initialization interval, and two-domain configuration covering minimal ocean and having a parent-to-nest area ratio of greater than 10 best simulates a recent (July 2002) large storm event over the San Antonio River basin. This particular storm was selected because it produced extreme rainfall volumes and intensities, and also because its meteorological characteristics are typical of central Texas storm events. Location errors in rainfall are most significant because of their typically nonlinear patterns (increasing location error does not linearly modify streamflow output). Errors in intensity and timing show a more predictable (linear) watershed response that may be useful in the estimation of streamflow ranges for flood forecasting.

## 1. Introduction

Precipitation is the single most difficult and often erroneously modeled parameter in numerical weather models (Wang and Seaman 1997; Nielsen-Gammon et al. 2006). Determining the appropriate grid spacing that captures both small- and large-scale processes is vital to accurately representing storm events. For example, neglecting the subgrid-scale variability of precipitation has been shown to result in an underestimation of the total volume and runoff and, consequently, an overestimation of the evapotranspiration (Wang et al. 2005). Mitigation decisions based on local-scale engi-

neering models may not fully consider the hydrologic effects of large storm systems. Regional weather models have been successfully applied to the forecasting of hydrological phenomena and quantitative precipitation forecasts (Fovell 2006; Marchok et al. 2007; Westrick et al. 2002). A nested regional model provides a realistic method for modeling flood events at the watershed scale by capturing synoptic-scale triggers over the entire domain and downscaling to capture mesoscale factors over the region of interest. This approach has proved successful at various grid spacings (Liang et al. 2004a,b; Smedsmo et al. 2005).

Systematic errors still exist in regional models. Major causes of model error are investigated in this study. Since causes of error vary with the specific region and climate of interest, a regional model must be adapted for each application (Giorgi and Mearns 1999). Major errors arise from model physics, domain and buffer

*Corresponding author address:* Marla R. Knebl Lowrey, Dept. of Geological Sciences, 1 University Station, C-1100, University of Texas at Austin, Austin, TX 78712-0254.
E-mail: mlowrey@trcsolutions.com

zone treatments, and initial and boundary conditions. Even after accounting for major sources of error, small-amplitude random errors may remain in the model.

The implementation of various physics schemes causes a large variation in the forecast output (Zhang et al. 2006), especially the choice of cumulus scheme (Gallus and Segal 2001). The particular skill of a cumulus scheme in simulating rainfall is dependent upon the region and storm being modeled (Giorgi and Mearns 1999). The Grell–Devenyi (GD) scheme has displayed widespread skill in simulating precipitation (Giorgi and Shields 1999; Warner and Hsu 2000; Xu and Small 2002; Zhang et al. 2006). However, several studies have shown the Kain–Fritsch (KF) scheme to produce realistic precipitation over the North American monsoon region (Gochis et al. 2002; Leung et al. 2003; Liang et al. 2004b); other research has demonstrated the strength of the Betts–Miller–Janjić (BMJ) scheme at producing accurate forecasts (Baik et al. 1991; Vaidya and Singh 2000). Emanuel and Zivkovic-Rothman (1999) found that the BMJ scheme performs consistently with observations for relative humidity from the surface up to 500 mb. Jankov and Gallus (2004) found that neither the BMJ nor the KF scheme accurately reproduced mesoscale convective systems (MCS) in the Upper Midwest, with the BMJ scheme consistently overestimating and the KF scheme underestimating rainfall.

Cloud microphysics is also an important parameterization in the production of precipitation (Fritsch and Carbone 2004; Smedsmo et al. 2005). Cloud microphysics schemes vary widely. Some schemes only include water classes (e.g., Kessler 1969). Other schemes include some ice hydrometeors such as the Ferrier (1994) method with four total hydrometeor classes, the Hong and Lim (2006) six-class scheme with graupel, or the Thompson et al. graupel scheme (Thompson et al. 2004). Microphysics schemes for the present study were chosen from those available in version 2 of the Weather Research and Forecasting (WRF) model. The scheme most suited to the model application varies with location; for example, the Lin scheme (Lin et al. 1983) may be appropriate for hail-bearing storms in the Midwest (Smedsmo et al. 2005).

Emanuel and Zivkovic-Rothman (1999) assert that correct microphysics is integral to determining cumulus convection and hence precipitation. They postulate that reevaporation of condensed water, currently not included in most parameterizations, allows for correct estimates of convective moistening. Their cumulus parameterization controls entrainment and detrainment by allowing the level of natural buoyancy to vary depending on the moisture levels in cloud air. Both BMJ (Janjić 1994) and Emanuel's scheme are comparable to the observed relative humidity profiles; other schemes tested were less so.

Errors in the driving initial and boundary conditions can cause large variations in forecast output. Liang et al. (2004b) found large uncertainties in boundary conditions, mostly over oceans and other areas lacking complete data, contributed greatly to model error. The use of the National Centers for Environmental Prediction–National Center for Atmospheric Research (NCEP–NCAR) reanalysis caused less sensitivity to model domain when compared to the European Centre for Medium-Range Weather Forecasts (ECMWF) reanalysis (Liang et al. 2001). The initialization interval is important because of the spinup time needed for model adjustment, as well as the inability of most weather models to accurately forecast beyond several days. In this study, the initialization interval is taken to mean the difference in initialization times between each series of model runs over the storm event. A more detailed description of this methodology can be found in section 4b.

The size of the model domains and the interaction between nested domains greatly impact forecasts. If the domain edges are too close to the area of interest, edge effects may cause inaccuracies in the forecast. For example, Seth and Giorgi (1998) found that small domains caused unrealistic responses in the inner domain that were inconsistent with large-scale forcing. Locations of domain edges must be chosen so that important regional meteorological features can be resolved, and the edges must be located away from areas where reanalysis inaccuracies exist; parent and nest domains should be spaced far enough apart to avoid edge distortions (Liang et al. 2001). Increasing the width across which the dynamical relaxation of boundary conditions occurs, or nudging over the entire domain, may help preserve large-scale waves and hence reduce errors due to domain size (Juang and Hong 2001).

The major objectives of the present study surround the following questions: 1) How well does WRF model the synoptic setting that produced the July 2002 storm event? 2) What are major causes of model error in terms of physics, initialization, domain, and small-amplitude errors? 3) What types of precipitation error propagate most strongly into streamflow: magnitude, location, or timing? How do these each affect flood forecast outcomes? In the next section, an overview of the July 2002 storm meteorology is presented. Section 3 discusses the background information on the driving datasets and the WRF model structure, section 4 presents experimental methods and results, and section 5 discusses the conclusions and implications of this research.

## 2. Meteorology of the July 2002 storm event

The high frequency of flooding in Texas results from a unique juxtaposition of meteorological factors including moisture influx from the Gulf of Mexico, easterly waves moving across the area, and orographic uplift from the Balcones Escarpment (Hirschboeck 1987). In fact, Texas has recorded some of the greatest precipitation intensities in the world, for storms with durations of up to 24 h (Patton and Baker 1977). Storm events over central Texas are generally controlled by large-scale anomalies in atmospheric circulation patterns. The interaction of these anomalous behaviors with local and mesoscale processes, such as topographic influences and the location of mesohigh outflow boundaries, determines whether or not a flood event occurs in any given region. Extreme storms in Texas may be triggered by tropical cyclogenesis, cyclonic dissipation, frontal processes, orographic lifting, or a combination of several local- and synoptic-scale events. Most extreme storm events over Texas show large-scale meteorology that is significantly different from climatology at the 95% significance level (Nielsen-Gammon et al. 2006). The July 2002 storm was related to a deep upper-level trough that became stationary over south-central Texas. Much of the subsequent discussion is based upon the work of Nielsen-Gammon et al. (2006), who performed a compositing analysis to determine typical meteorological factors present during Texas storms and particularly the July 2002 event.

At the 850-mb level during typical storm evolution in central Texas, strong trade winds blow across the Yucatan toward Mexico and Texas, bringing moisture into the area. Decreasing wind speed from the Gulf of Mexico to the Great Plains implies low-level moisture convergence over the region. This is most easily seen as a deceleration of winds over eastern Texas. However, this deceleration was less prominent in the July 2002 storm compared to other large historical storm events over Texas. Observations on 30 June, several days before the onset of heavy rainfall, still show strong southeasterlies over the Gulf of Mexico. Observations for 5 and 6 July (after peak rainfall has occurred) show weaker winds and less deceleration over Texas.

Typical midlevel development (500 mb) is dominated by cyclonic circulation in the form of a north–south-elongated trough over south Texas. This cyclonic circulation causes a high potential vorticity (PV) anomaly that directs moisture away from Central America and into Texas. Precipitation efficiency is increased by southerly winds blowing into Texas from the Gulf of Mexico. A strong low-level jet (LLJ) feature, defined by Bonner (1968) as a low-level local maximum in the

vertical wind profile, is generally observed in Texas nontropical storms during September and October. In summer storm events, the strongest LLJ features are located more to the north in the central Great Plains; the southerly winds present during the July 2002 event do not show such localized maxima and are better referred to as a southerly wind event (SWE; Mitchell et al. 1995). However, similar to the Great Plains LLJ, the strong SWE transports vast amounts of moisture into the region and plays a large role in precipitation formation during the storm. In the upper levels of the atmosphere (200 mb), typical storm evolution in the region shows a high PV anomaly producing upper-level divergence over Texas. This PV anomaly causes an upper-level ridge to develop over the southeastern states, which prevents further movement of the trough.

Locally, mechanical lifting is thought to be a factor in central Texas storm evolution, demonstrated by the location of Balcones Fault Zone (BFZ) in relation to heavy rain concentration areas. Southeasterly to easterly winds at the 850-mb level over the region imply upward motion due to topography (Patton and Baker 1977). Many convective cells and high precipitation accumulations were observed in July 2002 along the BFZ. In addition to trigger factors for heavy rainfall, basin flooding is aggravated by the shallow soils, karstic terrain, bedrock channels, and increasing levels of urbanization over the San Antonio River watershed.

## 3. Reanalysis dataset and meteorological model

### a. North American Regional Reanalysis

Initialization and boundary conditions were interpolated from the NCEP North American Regional Reanalysis (NARR; Mesinger et al. 2006). NARR is a comprehensive hydrometeorological dataset used to drive regional-scale models. Based on the Global Reanalysis (GR) Project that ran 30 yr of simulations using the Eta Model and assimilated observations, NARR assimilates observational data for analysis, boundary conditions, and execution of the Eta Model.

Data are output as 3-hourly files on a 32-km Eta grid with 29 pressure levels, distributed on a Lambert conformal grid. Data are available for free download from the NCAR Web site in gridded binary (GRIB) format. The regional reanalysis appears to correctly capture important elements of the hydrologic cycle; NARR streamflow output compares well with the measured discharge, and has been shown to perform better than the previous 40-yr ECMWF Re-Analysis (ERA-40; Déry et al. 2007). Although simulations with NARR have been unable to completely balance the moisture
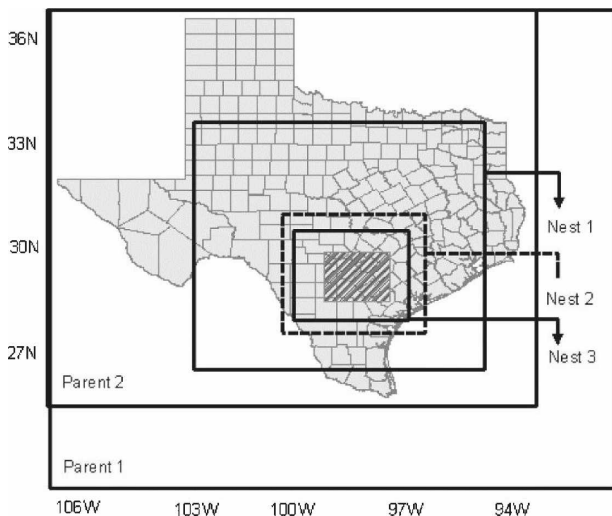
FIG. 1. Model domains. Control domain configuration includes parent 2 and nest 2 (other domains are referenced in section 4). The region of interest for forecasting (SAB) is represented by the hatched region.

budget, summer flux convergence over Texas and the northern Gulf of Mexico is much better represented with NARR than with previous global reanalyses (Nigam and Ruiz-Barradas 2006).

### b. WRF model

WRF version 2 (Michalakes et al. 2005), a nonhydrostatic, mesoscale weather model developed in a collaborative effort between NCAR, the National Oceanic and Atmospheric Administration (NOAA), and the Department of Defense (DOD), is used in this study. Initial conditions from NARR drive the model at the

first time step, and boundary conditions (also from NARR) are assimilated continuously at hourly time steps. The preliminary WRF configuration used in this study consists of a 12-km outer domain with a horizontal grid spacing of $105 \times 105$ grid points, and a 4-km inner domain with a horizontal grid spacing of $97 \times 94$ grid points. The area of interest for flood modeling, the San Antonio River basin (SAB), lies in the center of the inner domain; for reference, the basin area accounts for approximately 1/15 of the inner-domain area (Fig. 1). Both domains have 31 variably spaced, vertical grid cells in sigma coordinates, $\sigma = (P - P_t)/(P_s - P_t)$, with its model top at 100 mb. Ten levels reside at or below the planetary boundary layer (1.5 km); high vertical resolution in the boundary layer is thought to improve the simulation of the LLJ or SWE (Ting and Wang 2006), hence improving the model's representation of moisture transport and storm generation. The model physics are variable and described further in the following section.

## 4. Experimental design and results

### a. Physics parameterizations

The WRF model contains multiple options for the physics parameterizations. Adjustable physics components include the microphysics, cumulus parameterization, long- and shortwave radiation, boundary layer turbulence (PBL), surface layer, land surface parameterization, and subgrid-scale diffusion. Warm season precipitation output is especially sensitive to the choice of cumulus scheme and moderately sensitive to the microphysics, radiation, and boundary layer scheme (Wang and Seaman 1997; Liang et al. 2004a; Jankov et

TABLE 1. Various combinations of physics parameterizations used to run the WRF model in this study. Experiments are named based upon the different schemes used in each: MP, CU, NC, and SW refer to microphysics, cumulus, no convection, and shortwave radiation schemes, respectively. Best-fit combination of physics is shown in boldface.

| Expt | Microphysics | Cumulus | Radiation | Parameterization of convection in nest |
|---|---|---|---|---|
| MP1CU1 | Kessler | KF | Dudhia | Yes |
| MP1CU2 | Kessler | BMJ | Dudhia | Yes |
| MP1CU3 | Kessler | GD | Dudhia | Yes |
| MP2CU1 | Lin | KF | Dudhia | Yes |
| MP2CU1NC | Lin | KF | Dudhia | No |
| **MP2CU2** | **Lin** | **BMJ** | **Dudhia** | **Yes** |
| MP2CU2SW2 | Lin | BMJ | Goddard | Yes |
| MP2CU3 | Lin | GD | Dudhia | Yes |
| MP3CU1 | WSM | KF | Dudhia | Yes |
| MP3CU2 | WSM | BMJ | Dudhia | Yes |
| MP3CU3 | WSM | GD | Dudhia | Yes |
| MP4CU1 | Ferrier | KF | Dudhia | Yes |
| MP4CU2 | Ferrier | BMJ | Dudhia | Yes |
| MP4CU3 | Ferrier | GD | Dudhia | Yes |
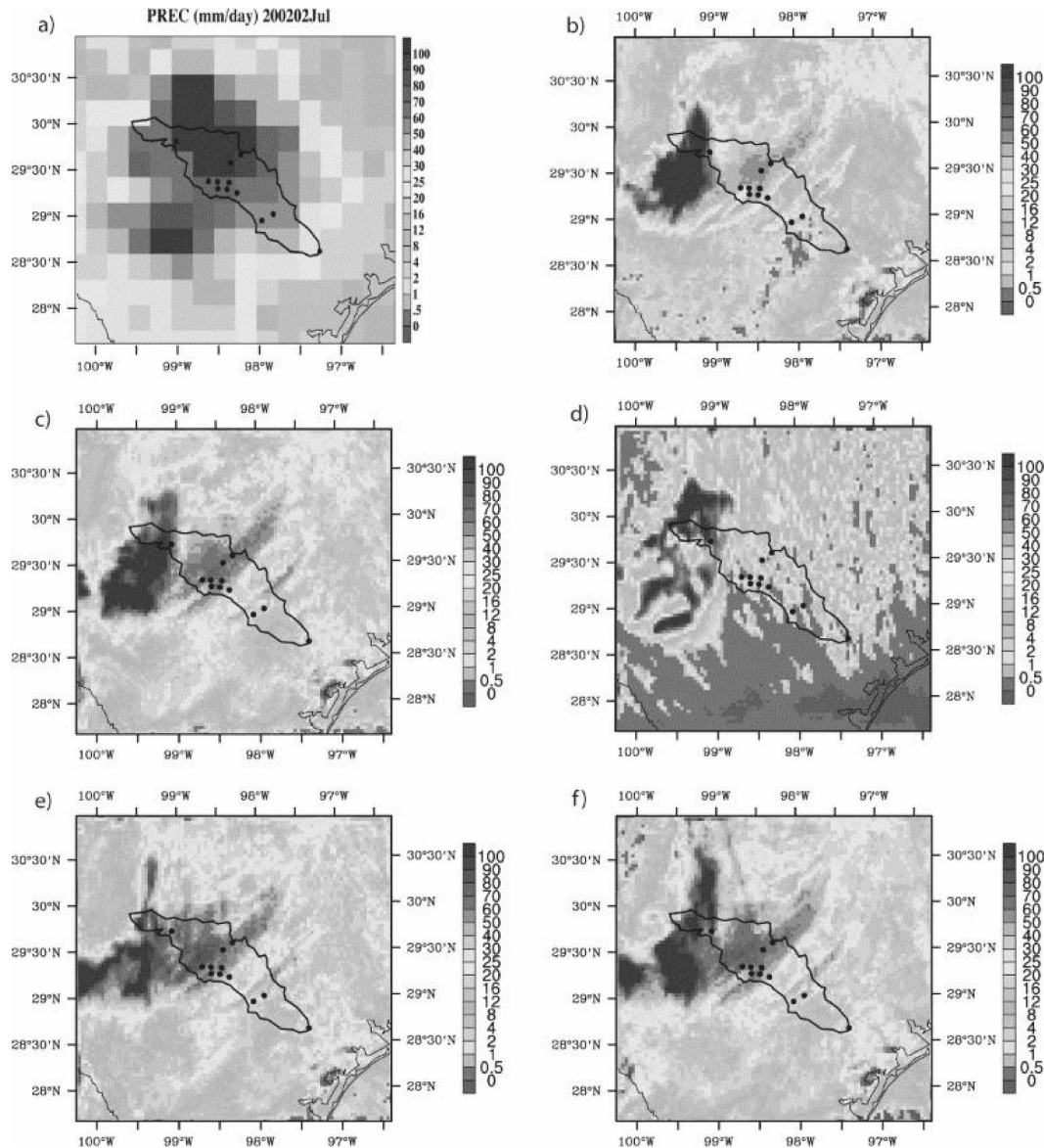
FIG. 2. Total daily precipitation (mm day$^{-1}$) for 2 Jul 2002 compared between (a) observations and WRF output using the KF cumulus parameterization, and various microphysics including (b) Kessler, (c) Lin, (d) Lin with no convection applied to the nest domain, (e) WSM, and (f) Ferrier. All model runs shown use RRTM shortwave radiation unless otherwise specified.

al. 2005; Nielsen-Gammon et al. 2006). Warm season convective precipitation is the most difficult variable to model accurately (Olson et al. 1995; Nielsen-Gammon et al. 2006), possibly due to the importance of the initial precipitation peak locations in providing feedback mechanisms that dictate the subsequent event evolution. For this study, sensitivity is tested using various combinations of schemes (approximately 40 different combinations); only those combinations that produce significantly different simulations of precipitation are described in this paper. The physics options are com-

bined as shown in Table 1 and included the Grell–Devenyi, Kain–Fritsch, and Betts–Miller–Janjić cumulus schemes; the Kessler, Lin, Eta, and WRF Single-Moment 3-Class (WSM) cloud microphysics; the Rapid Radiative Transfer Model (RRTM) and Eta longwave radiation; and Dudhia, Goddard, and Eta shortwave radiation. Several studies have suggested that for domains using a fine grid spacing (<5 km), a regional model may simulate the convective processes equally or better than using a parameterization of the convection (Bélair and Mailhot 2001; Smedsmo et al. 2005); hence,
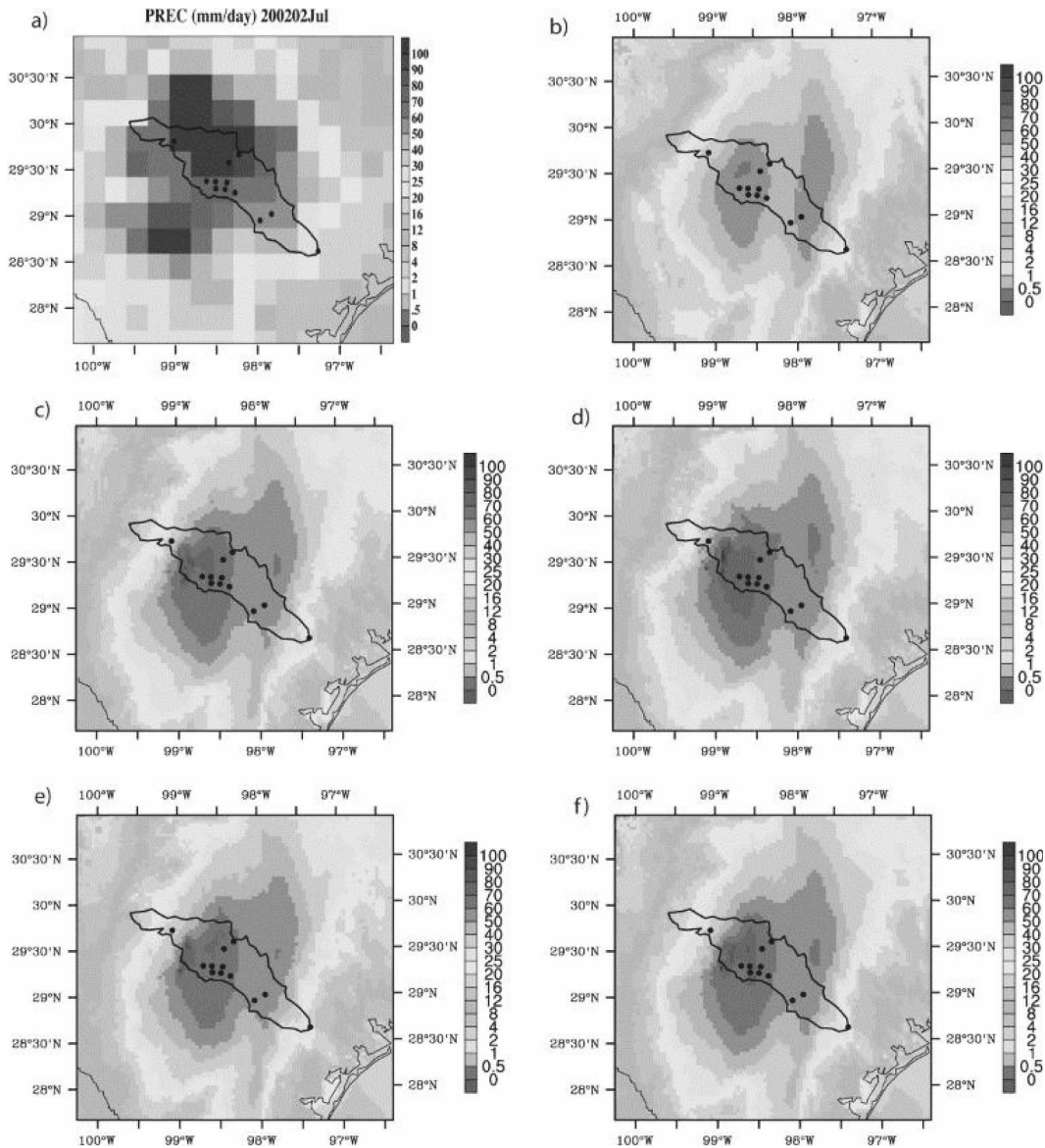
FIG. 3. Total daily precipitation (mm day$^{-1}$) for 2 Jul 2002 compared between (a) observations and WRF output using the BMJ cumulus parameterization, and various microphysics including (b) Kessler, (c) Lin, (d) Lin with Goddard shortwave radiation, (e) WSM, and (f) Ferrier.

the effect of removing the convection parameterization from the nested domain is also studied. Each model is run for the entire 11-day storm event using only a single initialization at day 1 of the storm.

Comparisons between observed and modeled rainfall are performed by upscaling from the fine-resolution WRF output to the coarser-resolution Higgins observed dataset (Higgins et al. 2000). The Higgins dataset is derived from rain gauge records from approximately 2500 stations on a 0.25° latitude × 0.25° longitude grid (about 32 km$^2$). Upscaling is performed by averaging the rain rate (mm day$^{-1}$) across the WRF

grid points; the total water volume is conserved in the upscaling process. Regional averages of WRF precipitation output are computed for 32 km$^2$ regions aligning with the Higgins rainfall grid that intersect the study basin, for a total of 25 grid cells. These grid regions are the comparison points for all subsequent analyses.

Precipitation is found to be most sensitive to the cumulus and to a lesser extent the microphysics schemes used. In Figs. 2–4, the San Antonio River watershed and its gauges are superimposed on the WRF rainfall map for the purpose of identifying areas of large errors across the basin. Although comparisons between the
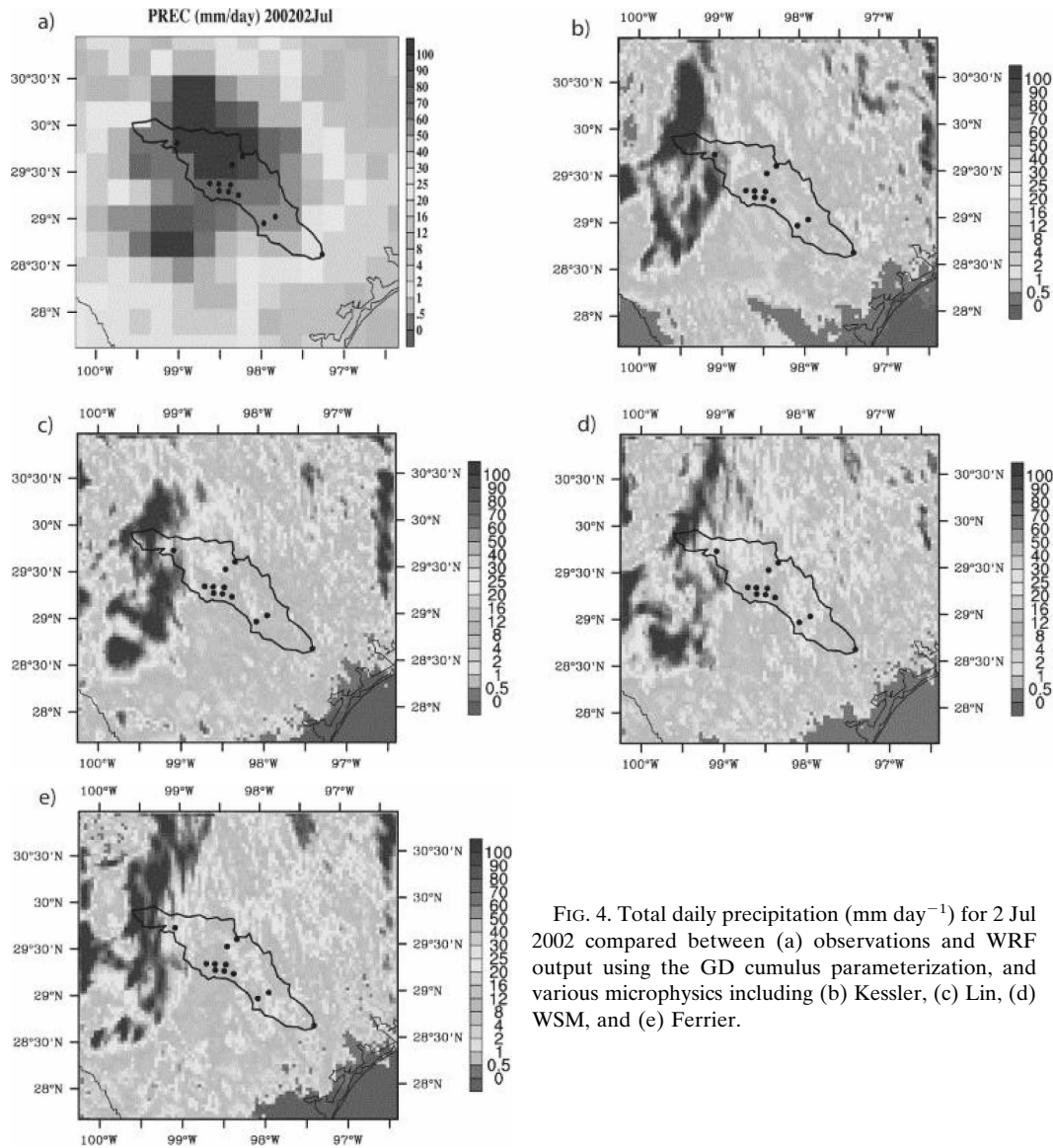
FIG. 4. Total daily precipitation (mm day$^{-1}$) for 2 Jul 2002 compared between (a) observations and WRF output using the GD cumulus parameterization, and various microphysics including (b) Kessler, (c) Lin, (d) WSM, and (e) Ferrier.

model and the observations were made over the entire 10-day storm, the following discussion pertains specifically to the peak days of rainfall (2–3 July). The Kain–Fritsch scheme (Fig. 2) produces spurious locations of high precipitation surrounding the basin on 2 July, especially in combination with the RRTM shortwave radiation scheme, while underestimating precipitation inside the watershed. The Grell–Devenyi cumulus scheme (Fig. 4) largely underestimates precipitation on 2 July, completely missing the large peak over San Antonio. Removing convection from the nested domain, regardless of the cumulus scheme used, results in a gross underestimation of the rainfall over the basin (Fig. 2d). This result is consistent with Kain and Fritsch (1998), who found the parameterization of convection

to be necessary at grid lengths as fine as 5–10 km. Variations in both the longwave and shortwave radiation schemes appear to have insignificant effects on the precipitation forecasts over the basin (Fig. 3d for 2 July; others not shown). Overall, the Betts–Miller–Janjić cumulus scheme (Fig. 3) most accurately simulates the precipitation and subsequent river flooding over the SAB; however, this scheme produces a wider swath of peak precipitation than observed on 3 July, accounting for the large errors in the lower portions of the basin on that day (Table 2). Results from experiment MP2CU2 most closely match the observed mean and standard deviation of rainfall for the basin. Despite the emphasis on the correct microphysical properties in modeling deep convection (Fritsch and Carbone 2004; Smedsmo

TABLE 2. Mean and standard deviation of precipitation over the SAB (as highlighted in Fig. 1) during the 48-h period of peak rainfall, and correlation with observations for each of the physics configurations used in this study. MAE is shown for both peak days over the upstream, midstream, and downstream portions of the watershed. Best-fit combination of physics is shown in boldface.

| Expt | Combined 2 and 3 July | | | MAE (%), 2 July | | | MAE (%), 3 July | | |
|---|---|---|---|---|---|---|---|---|---|
| | Mean (mm) | SD | R | Upper | Middle | Lower | Upper | Middle | Lower |
| Observed | 64 | 32 | — | — | — | — | — | — | — |
| MP1CU1 | 22 | 17 | −0.18 | 69 | 65 | 52 | 68 | 54 | 56 |
| MP1CU2 | 36 | 16 | −0.41 | 69 | 36 | 1 | 82 | 23 | 252 |
| MP1CU3 | 36 | 40 | −0.36 | 44 | 85 | 67 | 53 | 33 | 98 |
| MP2CU1 | 25 | 46 | −0.39 | 61 | 55 | 42 | 73 | 53 | 25 |
| MP2CU1NC | 13 | 10 | 0.27 | 69 | 87 | 83 | 86 | 83 | 72 |
| **MP2CU2** | **70** | **32** | **0.17** | **69** | **36** | **1** | **4** | **285** | **364** |
| MP2CU2SW2 | 49 | 28 | 0.15 | 65 | 29 | 10 | 56 | 86 | 374 |
| MP2CU3 | 24 | 11 | 0.19 | 59 | 77 | 62 | 71 | 38 | 9 |
| MP3CU1 | 27 | 6 | 0.60 | 62 | 53 | 41 | 68 | 40 | 9 |
| MP3CU2 | 45 | 17 | −0.32 | 69 | 33 | 6 | 61 | 77 | 313 |
| MP3CU3 | 28 | 9 | 0.47 | 55 | 71 | 50 | 67 | 20 | 4 |
| MP4CU1 | 30 | 9 | 0.52 | 52 | 52 | 41 | 67 | 34 | 3 |
| MP4CU2 | 45 | 16 | −0.24 | 68 | 33 | 7 | 57 | 74 | 278 |
| MP4CU3 | 29 | 20 | 0.28 | 46 | 74 | 53 | 56 | 44 | 31 |

et al. 2005), model runs using varying microphysics schemes produce less striking differences in precipitation than do model runs with varying cumulus schemes. This observation is most prominent when the BMJ convective scheme is used (Figs. 3b–f), likely because the BMJ dries the atmosphere and reduces the grid-resolved precipitation, minimizing the role of the microphysics in generating precipitation (Jankov et al. 2005). In Fig. 2b, the Kessler scheme underestimates rainfall over much of the basin on 2 July, especially the lower portions of the area. In the WSM scheme (Fig. 2e), the peak rainfall signal is much weaker than is observed. The Lin (Fig. 2c) and Ferrier (Fig. 2f) schemes produce more accurate rainfall over all areas of the basin, despite both having a peak that is located southwest of the observed rainfall peak. The finding that rainfall rate is most sensitive to changes in convective scheme and less sensitive to changes in microphysics and PBL schemes agrees with the recent results of Jankov et al. (2005), who also found the microphysics along with the convective treatment to be most influential on the total rain volume.
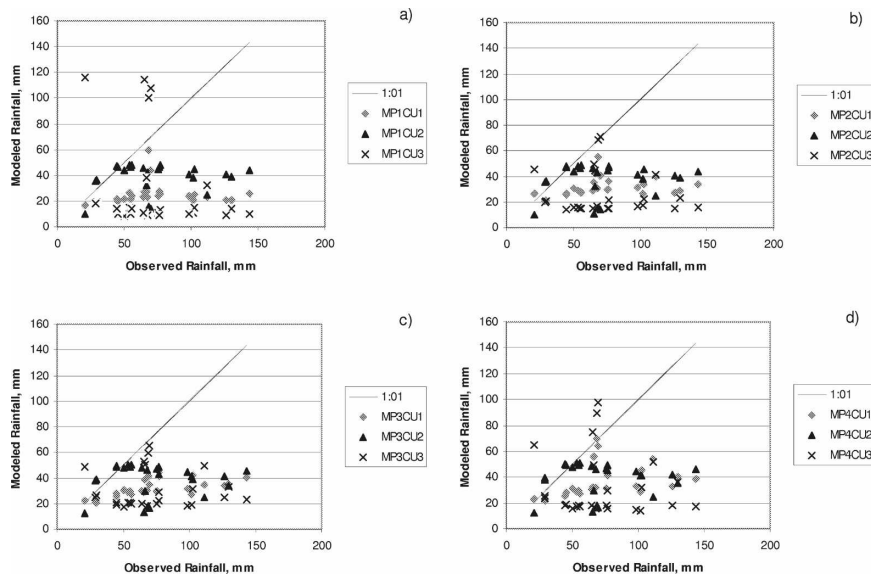


FIG. 5. Comparison of daily precipitation for 2 July, over 25 grid cells for the various experiments listed. A 1:1 line (perfect match with observations) is drawn for reference.
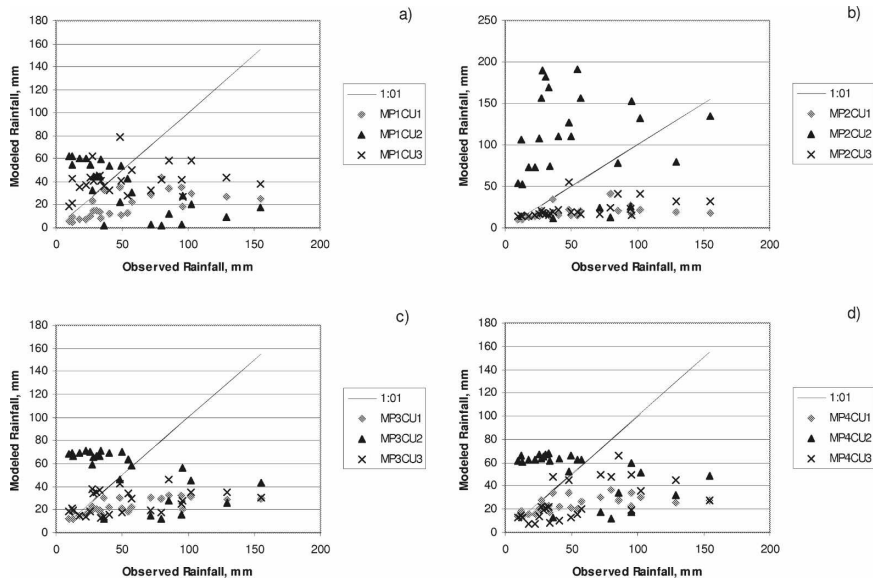
FIG. 6. Same as Fig. 5, but for 3 July.

In Figs. 5 and 6, it is clear that most physics combinations underestimate precipitation during the peak days of 2 and 3 July; on average, the combination runs underestimated precipitation by 47%. Experiments utilizing the KF cumulus scheme consistently underestimated rainfall for both days over nearly the entire basin. Experiments with the BMJ scheme tended to underestimate rainfall for 2 July and overestimate rainfall for 3 July, which may be attributed to error in the timing of the convection. However, the BMJ scheme produces some of the most accurate values over the central basin region. Experiments using the GD cumulus scheme show sporadic results, with the highest accuracy over the southern portions of the basin on 3 July. Daily totals averaged over the upper, middle, and lower basin areas (not shown) further support these characteristics. The physics configuration for the succeeding experiments are as follows: Lin microphysics, BMJ cumulus parameterization, Janjić Eta surface layer scheme, Rapid Update Cycle (RUC) land surface model, Mellor–Yamada–Janjić Eta boundary layer physics, and RRTM (longwave) and Dudhia (shortwave) radiation.

### b. Initialization interval

Decision making in flood management demands a high accuracy of weather model output, and many weather forecasts are considered impractical for use beyond 36 h. The potential performance of WRF in modeling future storm events is investigated in this study; in the following experiments, the effect of forecast length on convective rainfall output is addressed through variation of the initialization interval in WRF.

The WRF model is initialized at intervals of 24 h (1 day), 48 h (2 days), 72 h (3 days), 120 h (5 days), and 264 h (11 days, covering the entire storm length). A series of model runs are completed with each interval as necessary to cover the 11-day storm (Table 3). For example, the 24-h initialization experiment consists of 11 runs, the 48-h initialization consisted of 6 runs, etc. Results from the model runs are combined for each interval length and plotted for the entire storm length. Precipitation output grids are then run through the hydrologic model and the streamflow output compared between experiments.

Comparisons between precipitation output for different initialization intervals are shown in Figs. 7–9. Experiment results demonstrate that a 48-h initialization best reproduces the observed rainfall (Fig. 9). The 24-h run misses or underestimates several precipitation peaks, while the longer forecasts severely underestimate precipitation, and in some cases produce spurious areas of precipitation. Although this finding may be due to numerous factors, one possible reason that an intermediate initialization interval is more accurate in

TABLE 3. Initialization intervals used for WRF experimental runs.

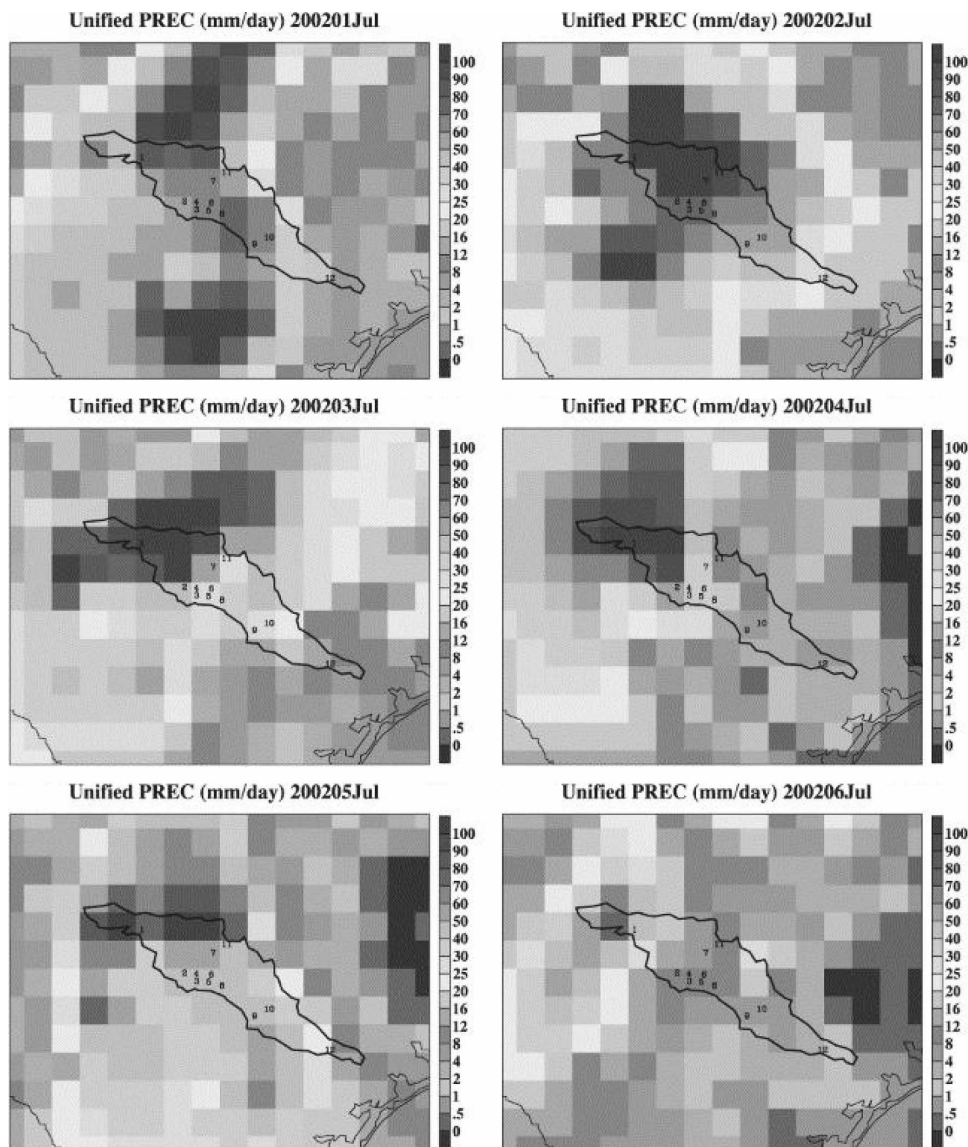| Expt | Initialization interval (h) | No. of model runs |
|------|------------------------------|--------------------|
| 1D | 24 | 11 |
| 2D | 48 | 6 |
| 3D | 36 | 4 |
| 5D | 120 | 3 |
| 11D | 264 | 1 |

FIG. 7. Observed total daily precipitation (mm day$^{-1}$) from the Higgins dataset. Days 2–7 of the storm-event area shown: (top left) 1 July, (top right) 2 July, (middle left) 3 July, (middle right) 4 July, (bottom left) 5 July, and (bottom right) 6 July.

representing rainfall may be the adjustment time in the model, as the model integrates the NARR initial boundary conditions (IBCs) and begins producing grid variables on its own (i.e., the time needed for the model to reach equilibrium). In this particular case, the adjustment time may also include an adjustment between the differences in the Eta land surface model used in the NARR data and the RUC land surface model used in the WRF model runs for this experiment.

The correlation with the observations is significantly higher for smaller initialization intervals, with an r value of 0.47 for a 1-day initialization and 0.63 for a

2-day initialization (Table 4). The streamflow output shows patterns that are similar to those for the precipitation (Fig. 10), with all but the 48-h run underestimating the streamflow. The 48-h run produces a slight time lag compared to the observed streamflow, but captures the overall peak intensity and timing of the basin streamflow. Correlation values are high for both experiments 1D and 2D. At urban gauges, the correlation is very low (Table 5); this low correlation was also observed when driving the streamflow model with observed radar data (Knebl et al. 2005) and, therefore, is likely due to error inherent in the hydrological model
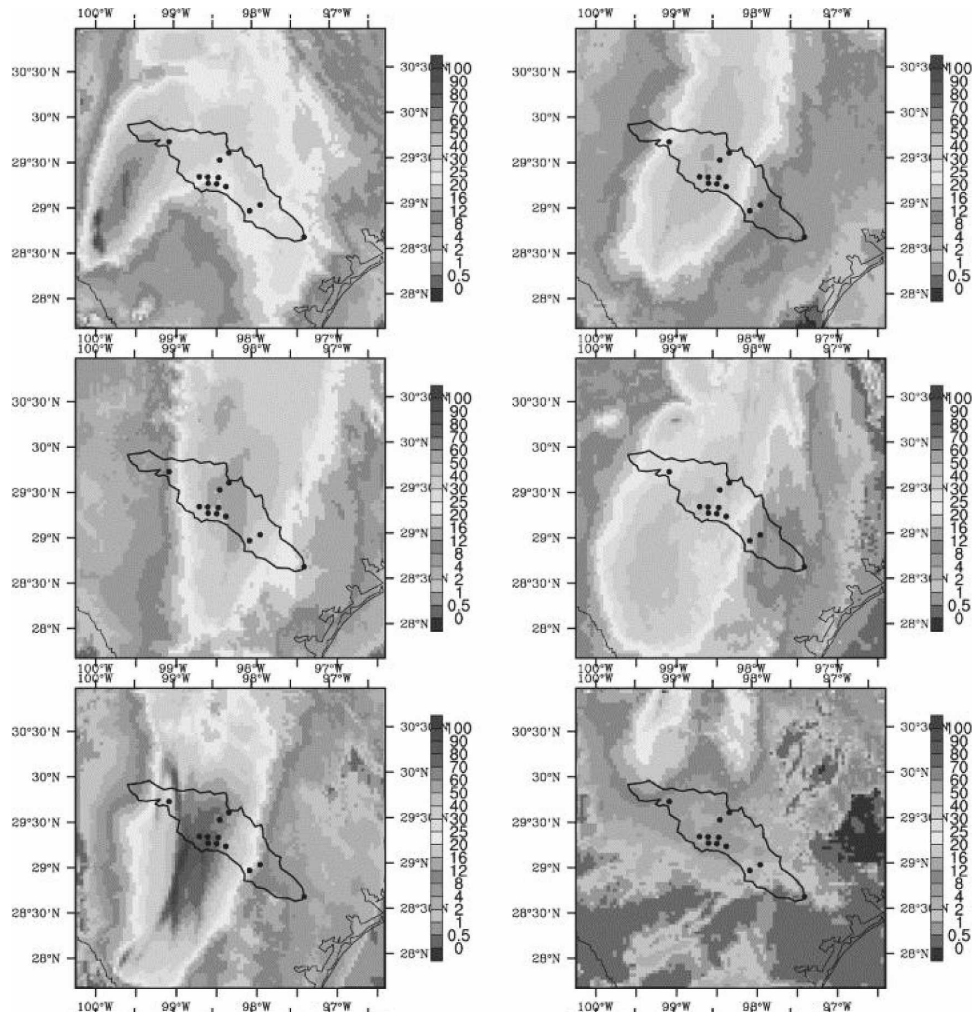
FIG. 8. Same as Fig. 7, but for total daily precipitation (mm day$^{-1}$) from WRF output using 11-day initialization.

and not error in the WRF output. A comparison of the mean biases for each experiment (Table 5) as well as an analysis of the bias distribution (Fig. 11) clearly demonstrates that the 2-day initialization best represents the entire storm. These results demonstrate that WRF performs reasonably well at producing precipitation on its own (without assimilation of observations) for 48-h intervals.

### c. Domain and boundary size

To study the sensitivity of precipitation forecasts to the size and location of the model domain, various domain configurations are designed as shown in Fig. 1. Six experiments are completed based on the six configurations (Table 6). All experiments utilize two domains, and the outer domain is driven by NARR forcing data. Physics and initialization details are dictated by best-fit

results from the previous sections. Experiments are named using capital letters for larger domains and lowercase letters for smaller domains, according to the following convention: P and p for parent domains 1 and 2; N, n, and nn for nest domains 1, 2, and 3, and BZ for increased boundary zone width. Experiment pn is designated as the control run, with the outer domain covering the entire state of Texas, and the inner domain covering the area of interest (SAB), including a wide radius around the basin to allow for the adjustment of parameters between the two domains. The outer domain in experiment Pn is enlarged to cover a significant portion of the Gulf of Mexico at its southeast corner (20 points added to the south and east sides). The objective of this configuration is to allow more adjustment time between ocean and land variables. To lessen the boundary effects that may occur when the forecast area is too
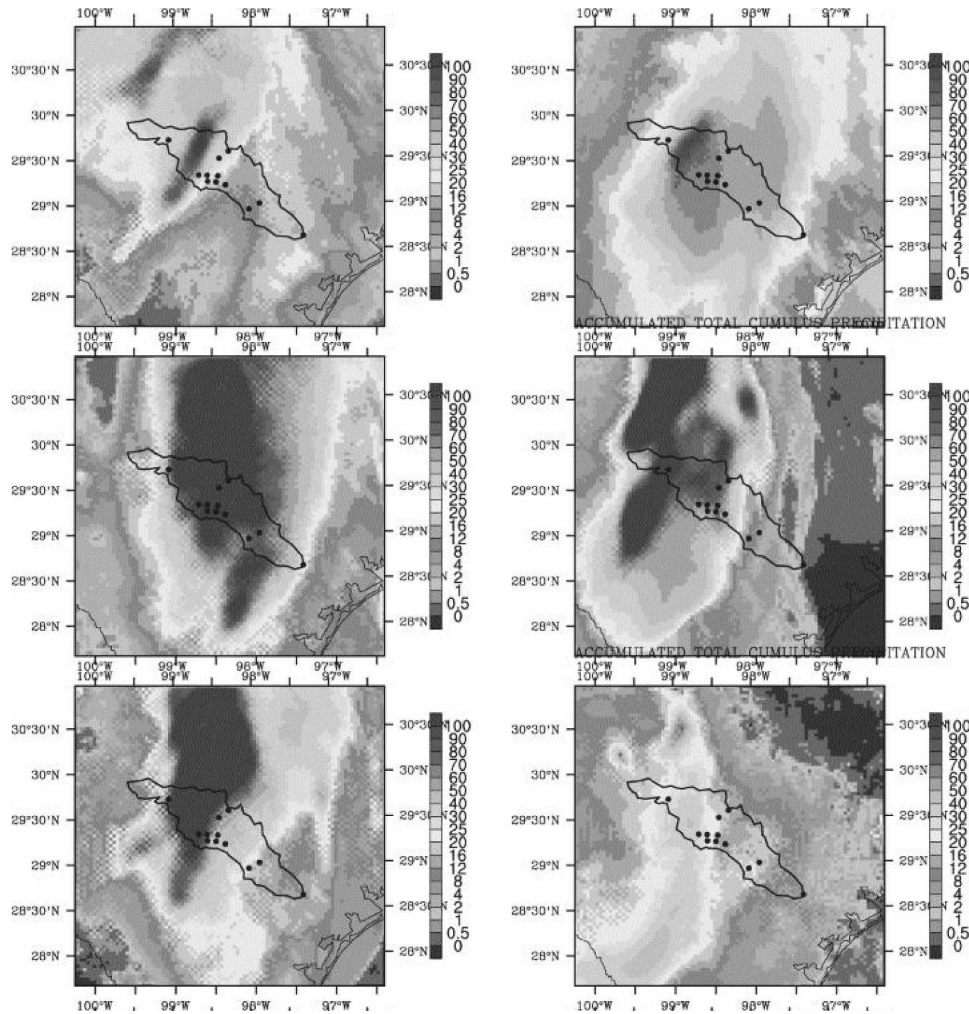
FIG. 9. Same as in Fig. 8, but for a 2-day initialization.

close to the parent–nest boundary, experiment pN enlarges only the nested domain, approximately doubling its grid points in both directions. Both the parent and nest domains are enlarged in experiment PN, to observe the combined effects of the previous two configurations. Experiment pnn shrinks the nested domain eight points in the north–south direction and four points in the east–west direction to singularly model the area of flood interest. Last, experiment pnBZ doubles the default buffer zone width from 5 to 10 grid cells in an effort to reduce edge effects from nudging between the parent and nest domains.

A qualitative comparison of the precipitation grids appears to show the control experiment pn to be the most reasonable configuration for modeling the 2002 storm event over the San Antonio River basin (Figs. 12–13; others not shown). Statistical analysis, however, demonstrates that each configuration has strengths and

weaknesses (Tables 7 and 8). Enlarging the parent domain results in location errors of rainfall over the San Antonio River basin during times of peak precipitation (experiment Pn; results not shown). These errors in the rainfall simulations may be, in part, due to the placement of the southern boundary in the Gulf of Mexico

TABLE 4. Precipitation mean, standard deviation, and correlation values for WRF experiments in which initialization is varied. Values are calculated over the SAB as highlighted in Fig. 1.

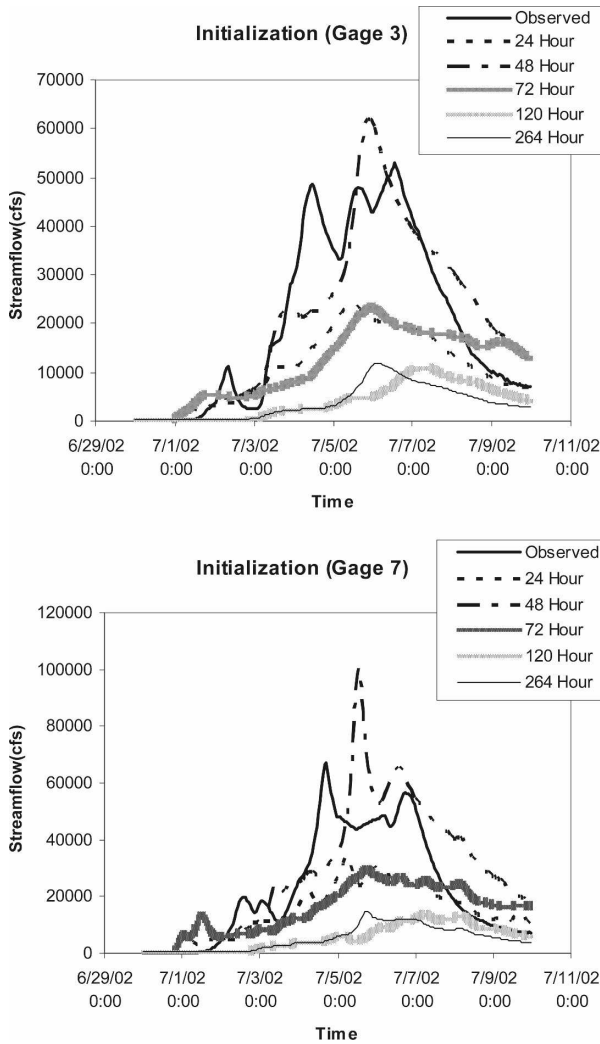| Expt | Storm mean (mm day$^{-1}$) | Daily deviation | Correlation with obs (%) |
|------|------|------|------|
| Observed | 31.6 | 13.6 | — |
| 1D | 38.8 | 10.8 | 0.47 |
| 2D | 66.8 | 23.1 | 0.63 |
| 3D | 33.4 | 7.7 | 0.27 |
| 5D | 20.6 | 4.46 | 0.04 |
| 11D | 19.9 | 4.4 | 0.17 |

FIG. 10. Select streamflow output ($m^3\,s^{-1}$) from the hydrological model at U.S. Geological Survey (USGS) gauge location (a) 3 and (b) 7 that compare the effects of different driving data (rainfall) from WRF.

TABLE 5. Correlation of modeled streamflow with observations for hydrological model runs driven with gridded precipitation output from WRF model runs of varying initialization intervals. Correlation coefficients at each gauge area are included to demonstrate the wide spatial variability of the error.

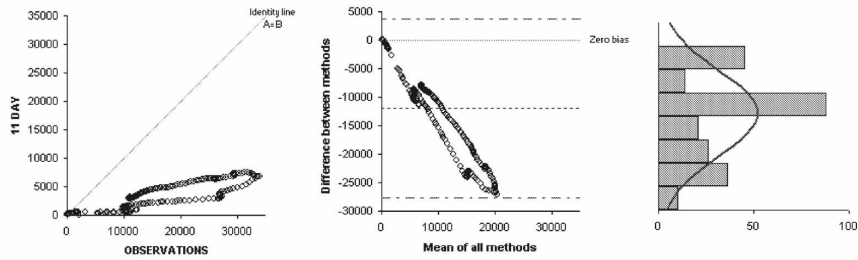| Expt | 11D | 5D | 3D | 2D | 1D |
|---|---|---|---|---|---|
| Correlation: upstream | | | | | |
| Gauge 1 | 0.89 | 0.5 | 0.69 | 0.7 | 0.66 |
| Gauge 2 | 0.75 | 0.5 | 0.69 | 0.84 | 0.9 |
| Gauge 3 | −0.04 | −0.18 | −0.06 | 0.01 | 0.05 |
| Gauge 4 | 0.26 | −0.04 | 0.26 | 0.34 | 0.47 |
| Correlation: midstream | | | | | |
| Gauge 5 | 0.78 | 0.55 | 0.75 | 0.88 | 0.93 |
| Gauge 6 | 0.32 | 0.11 | 0.23 | 0.18 | 0.34 |
| Gauge 7 | 0.73 | 0.49 | 0.74 | 0.78 | 0.88 |
| Gauge 8 | 0.02 | −0.01 | −0.01 | 0.04 | 0.04 |
| Gauge 9 | 0.84 | 0.6 | 0.81 | 0.84 | 0.92 |
| Correlation: downstream | | | | | |
| Gauge 10 | 0.94 | 0.91 | 0.96 | 0.96 | 0.95 |
| Gauge 11 | 0.47 | 0.16 | 0.56 | 0.63 | 0.78 |
| Gauge 12 | 0.88 | 0.88 | 0.94 | 0.96 | 0.98 |
| Mean (basin) correlation | 0.98 | 0.96 | 0.94 | 0.88 | 0.88 |
| Mean bias ($ft^3\,s^{-1} \times 10^3$) | −11.9 | −11.7 | −6.0 | 1.8 | −5.9 |

ing the area of the nest in experiment pnn (result not shown) has a similar effect to increasing the buffer zone width; precipitation patterns remain similar but decreased in intensity, and boundary edge effects are clearly observable.

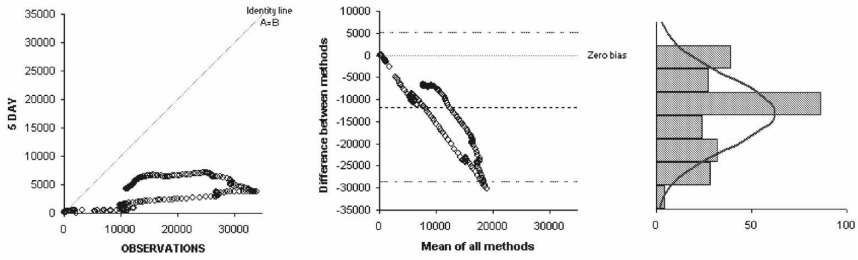### d. Performance of best-fit WRF model

The best-fit WRF model, using a 48-h initialization, domain pn, and physics as described previously, is investigated in more detail in order to assess its ability to simulate the July 2002 storm event. Figure 14 compares the horizontal winds and divergence between the NARR reanalysis data and two WRF configurations: the best-fit configuration and a poor-performing configuration (i.e., one that produces significant precipitation error, domain PN).

Figure 14 displays the observed winds (from the NARR reanalysis) and the modeled wind vectors for the upper, middle, and lower levels of the atmosphere. Strong trade winds are apparent at the 850-mb level over the Gulf of Mexico; these winds bring moisture into the central Texas region and weaken as the storm progresses. As these southerly winds move across land, they begin to weaken slightly, triggering low-level convergence. Weakening winds are also apparent as air flows over the BFZ in central Texas. Both the best-fit and poor-performing WRF models appear to simulate this near-surface wind pattern quite well (Figs. 14b and 14c).
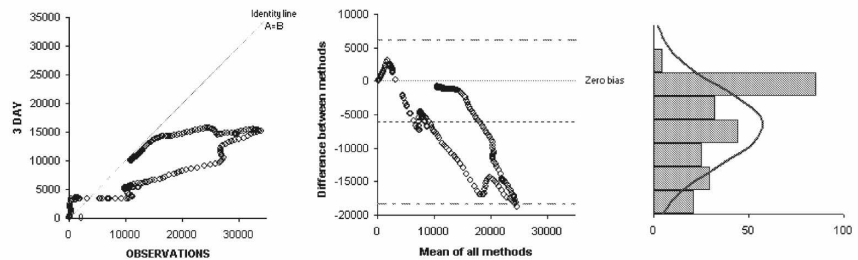
where large forcing errors exist (Liang et al. 2001); they may also be, in part, introduced via competition for convection at the land–ocean boundary (Nigam and Ruiz-Barradas 2006). In experiment pN (Fig. 13), rainfall is underestimated, suggesting that enlarging the nest domain may reduce the effects of the synoptic forcing from the parent domain. In experiment PN (result not shown), the effects of enlarging both domains are compounded, with large errors in both the intensity (underestimation) and location of the rainfall. Doubling the buffer zone width in experiment pnBZ (result not shown) results in an underestimation of the precipitation, which is presumably caused by the diminished importance of small-scale mechanisms in producing convective precipitation in the nested domain. Decreas-
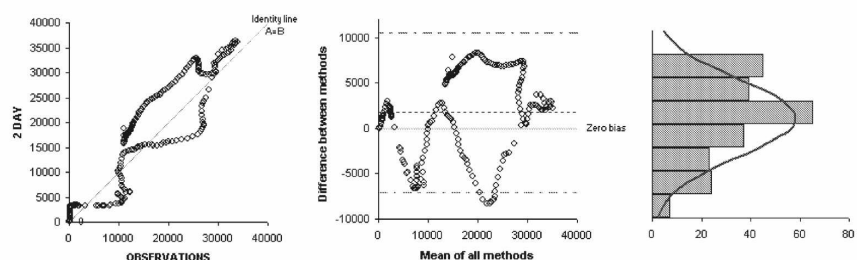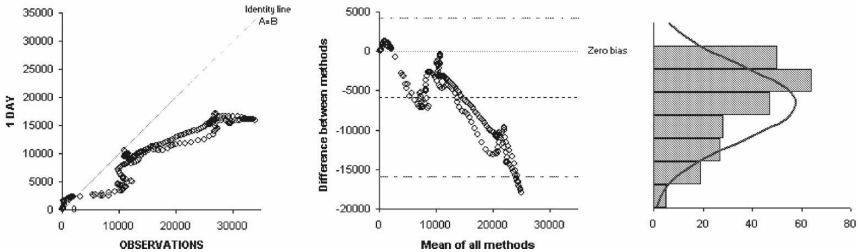
FIG. 11. Bias plots for streamflow, using WRF output from initialization experiments to drive the hydrological model, averaged over the watershed. (left) Comparison of observed data to modeled streamflow, (middle) residuals, and (right) distribution of residuals in order to show the relative amount of underestimation and overestimation of streamflow: (top to bottom) 11 to 1 day simulation.

TABLE 6. Domain configuration and grid dimensions used for domain experiments.

| Expt | Parent/nest | Parent grid ($i \times j$) | Nest grid ($i \times j$) |
|------|-------------|----------------------------|--------------------------|
| pn | 2/2 | $105 \times 105$ | $97 \times 94$ |
| Pn | 1/2 | $125 \times 125$ | $97 \times 94$ |
| pN | 2/1 | $105 \times 105$ | $193 \times 196$ |
| PN | 1/1 | $125 \times 125$ | $193 \times 196$ |
| pnn | 2/3 | $105 \times 105$ | $79 \times 70$ |
| pnBZ | 2/2/large boundary zone | $105 \times 105$ | $96 \times 93$ |

At midlevels of the atmosphere, a strong cyclonic circulation, consistent with the observations, develops in the best-fit model (Fig. 14e). The poor-performing WRF places this circulation too far north, over the Texas–Oklahoma boundary, which may contribute to its underestimation of the storm's extent (Fig. 14f).

Both the upper-level divergence and a related anticyclone east of the storm are clearly displayed in the WRF best-fit model (Fig. 14h). The poor-performing model run (Fig. 14i) demonstrates much weaker divergence and only minimal anticyclonic winds. This lack of divergence in the upper atmosphere may inhibit convection and contribute to precipitation underestimation.

Additionally, water vapor mixing ratios display patterns that are consistent with the observations (not shown), with increasing values as the storm progresses and significantly decreasing values several days after the storm has passed. The largest variation in moisture occurs at midlevels of the atmosphere, where cyclonic circulation dominates, and over the middle and upper regions of the basin, where convective precipitation is strongest.

### e. Small-amplitude random errors

Even after adjusting the model physics, initialization, and domain configurations, the resulting best-fit model
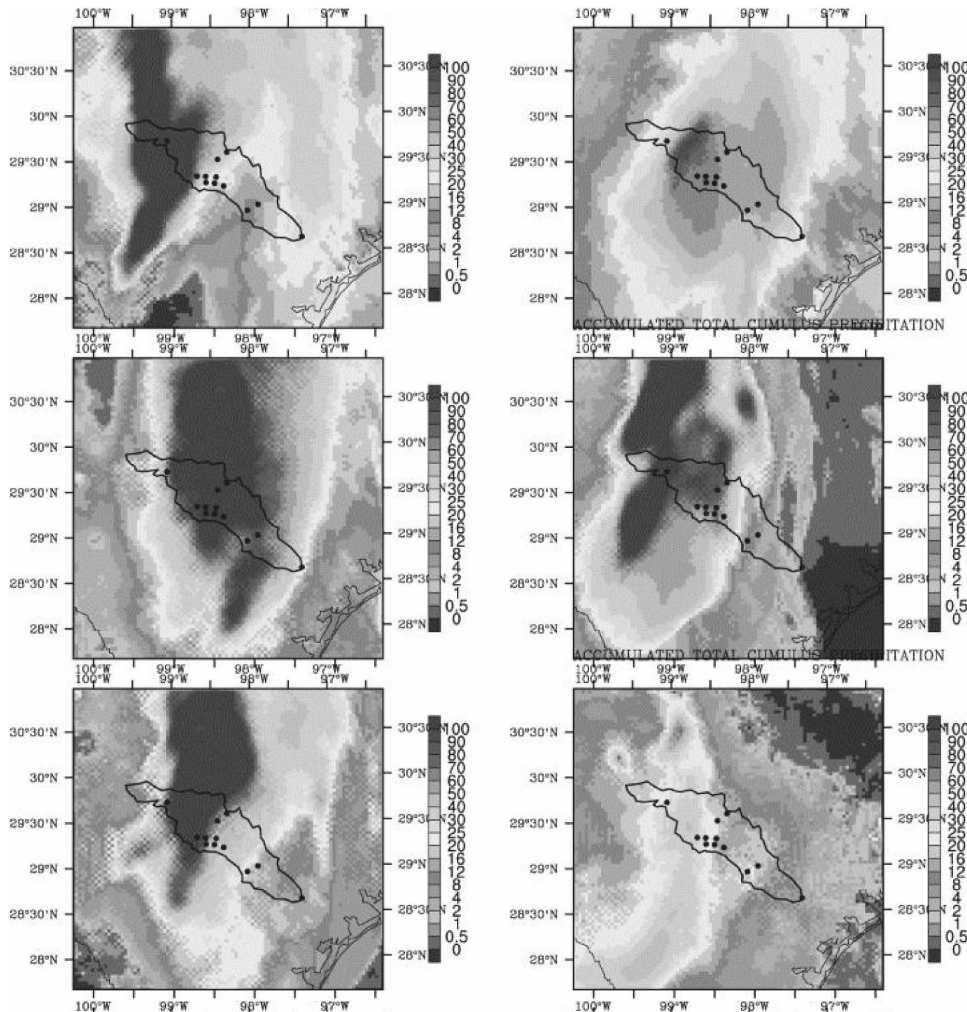


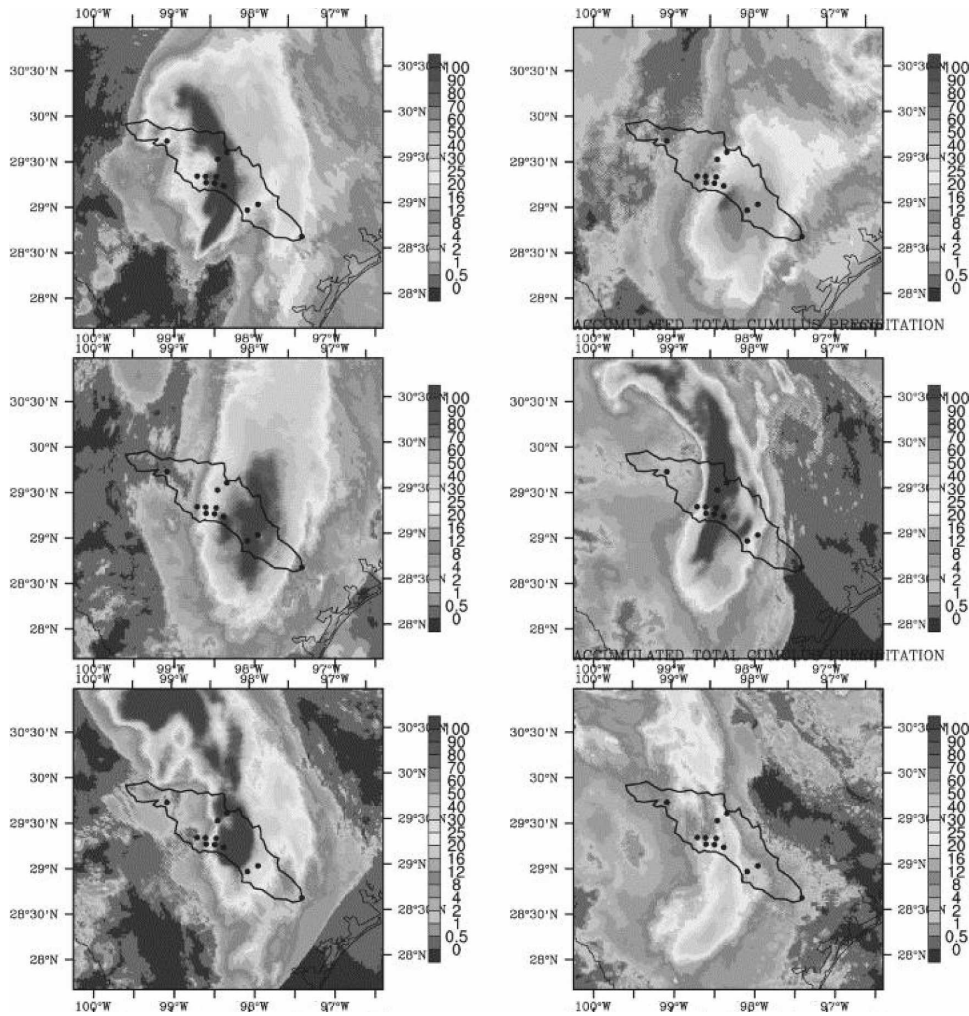FIG. 12. Same as Fig. 8, but for WRF output using domain configuration pn.

FIG. 13. Same as Fig. 12, but for domain configuration pN.

still shows significant errors compared to the observations. The remaining errors may be partially attributed to random error but also partially attributed to errors in the best-fit model physics (it should be noted that this study only reduces errors attributable to physics by finding the best-fit physics set; the set of physics schemes still contains substantial errors). Random errors can develop from small inaccuracies in the initial conditions that grow exponentially with time. Introducing small perturbations into the model can aid in understanding how these initial errors grow. Zhang et al. (2006) perturbed their model domain using thermal bubbles and found that nonlinear error growth mechanisms are of secondary importance to practical errors such as erroneous physical parameterization. Their results implied that error growth is greatest for high-resolution domains and in regions of high convective activity. Major error types encountered in regional

models include intensity, location, and timing errors, with erroneous location shifts most common.

Since precipitation errors are nonlinear, their effects on various hydrologic fluxes may be manifest in either an amplification or a dampening of the variable. To test the effects of common precipitation errors in the WRF

TABLE 7. Same as in Table 4, but for domain configuration experiments.

| Expt | Storm mean (mm day$^{-1}$) | Daily deviation | Correlation with obs (%) |
|---|---|---|---|
| Observed | 31.6 | 13.6 | — |
| pn | 42.34 | 14.0 | 0.72 |
| Pn | 42.43 | 8.1 | 0.73 |
| pN | 25.30 | 6.4 | 0.68 |
| PN | 22.77 | 5.4 | 0.71 |
| pnn | 36.06 | 6.5 | 0.62 |
| pnBZ | 27.26 | 7.3 | 0.48 |

TABLE 8. Same as in Table 5, but for domain configuration experiments.

| Expt | pn | Pn | pN | PN | pnn | pnBZ |
|---|---|---|---|---|---|---|
| Correlation: upstream | | | | | | |
| Gauge 1 | 0.29 | 0.94 | 0.64 | −0.09 | 0.28 | 0.84 |
| Gauge 2 | 0.8 | 0.43 | −0.34 | −0.33 | 0.8 | 0.56 |
| Gauge 3 | −0.02 | −0.09 | 0.0 | 0.0 | −0.03 | −0.06 |
| Gauge 4 | 0.27 | 0.16 | −0.11 | −0.06 | 0.25 | 0.1 |
| Correlation: midstream | | | | | | |
| Gauge 5 | 0.85 | 0.5 | −0.62 | −0.49 | 0.84 | 0.63 |
| Gauge 6 | 0.18 | 0.14 | 0.01 | 0.2 | 0.019 | 0.16 |
| Gauge 7 | 0.75 | 0.27 | −0.48 | −0.42 | 0.74 | 0.53 |
| Gauge 8 | 0.03 | 0.0 | 0.27 | 0.27 | 0.06 | 0.01 |
| Gauge 9 | 0.79 | 0.32 | −0.48 | −0.46 | 0.78 | 0.64 |
| Correlation: downstream | | | | | | |
| Gauge 10 | 0.97 | 0.8 | −0.92 | −0.91 | 0.97 | 0.94 |
| Gauge 11 | 0.53 | 0.38 | −0.36 | −0.36 | 0.5 | 0.33 |
| Gauge 12 | 0.9 | 0.84 | −0.96 | −0.96 | 0.93 | 0.86 |
| Mean (basin) correlation | 0.87 | 0.62 | −0.29 | −0.52 | 0.86 | 0.75 |
| Mean bias (cfs $\times 10^3$) | 2.8 | −6.9 | −13.7 | −13.7 | −1.2 | −8.4 |

model on streamflow output from a surface hydrologic model, WRF output grids were perturbed arbitrarily (in both the positive and negative directions) in intensity, location, and timing. Only those grid cells residing in the area of flood interest (Fig. 1) were perturbed. Intensity error was introduced by adding the following rainfall to each grid cell (in mm h$^{-1}$): ±0.01, 0.05, 0.1, 0.5, 1, 2, 3, 4, and 5. The absolute value, instead of a percentage, was used to obtain more meaningful results over the basin. Location error was introduced by shifting rainfall amounts 1, 3, 5, and 10 grid cells in each of four directions: NE, NW, SE, and SW. Finally, timing errors were introduced in two ways. First, hours were added or subtracted from each rainfall grid cell as follows: ± 3, 6, 12, and 18 h. This simple shift is not particularly realistic for typical rainfall errors, so additional experiments were run changing the duration of the rainfall. Rainfall grid cells during peak times were perturbed to (a) widen the peak, (b) narrow the peak, and (c) extensively narrow the peak (to mimic flash flooding).

Table 9 uses a moderate rainfall error of +2 mm h$^{-1}$ to demonstrate the variation in streamflow error with location in the basin, over the 72-h peak rainfall period. Correlation values over the storm period are high for all locations. The value of the best-fit to modeled ratio CTL/EXP2 allows study of the relative overestimation and underestimation of precipitation in each subbasin region. As expected, increasing rainfall error in a positive direction results in an overestimation of the streamflow in all subbasins, with the exception of the lowermost subbasin. This large subbasin consists of mostly agricultural soils, and may be better equipped to dampen the response to increased rainfall and stream-

flow upstream. In addition, this subbasin has a long response time; much of the impact from upstream rainfall may be delayed beyond the time scope observed in this study. In subbasins with a shorter response time, such as those with bedrock-dominated streams (gauge 2) or those in urban areas (gauge 7), increases in precipitation are quickly translated to runoff, and hence in these basins streamflow is greatly overestimated.

Statistics for various error types are displayed in Table 10 and Figs. 15–17. Positive (negative) perturbations to the rainfall magnitude show a comparable increase (decrease) in streamflow. Interestingly, positive rainfall errors are manifested in streamflow error throughout the entire storm, while negative errors result in lower overall streamflow error and are concentrated at the storm peak (Fig. 15). Intensity errors in precipitation show a linear translation pattern for streamflow, especially for positive rainfall perturbations; intensity errors also tend to retain the shape of the hydrograph. Location shifts in precipitation are a significant error type, because the streamflow response is erratic and difficult to predict. Rainfall location errors to the east, especially the southeast, result in the highest streamflow errors (Fig. 16). Location errors to the west (NW and SW) appear to retain the correct hydrograph shape and have fewer errors in discharge magnitude. This directional variability of the streamflow error is highly dependent on the actual location and size of the peak rainfall; glancing back at Fig. 12 (the best-fit rainfall), it is clear that rainfall changes most rapidly in an ESE direction away from the peak. Lagging or advancing the precipitation peak timing results in a linear lag or advance of the streamflow peak (Figs. 17a and 17b); however, the error is significantly
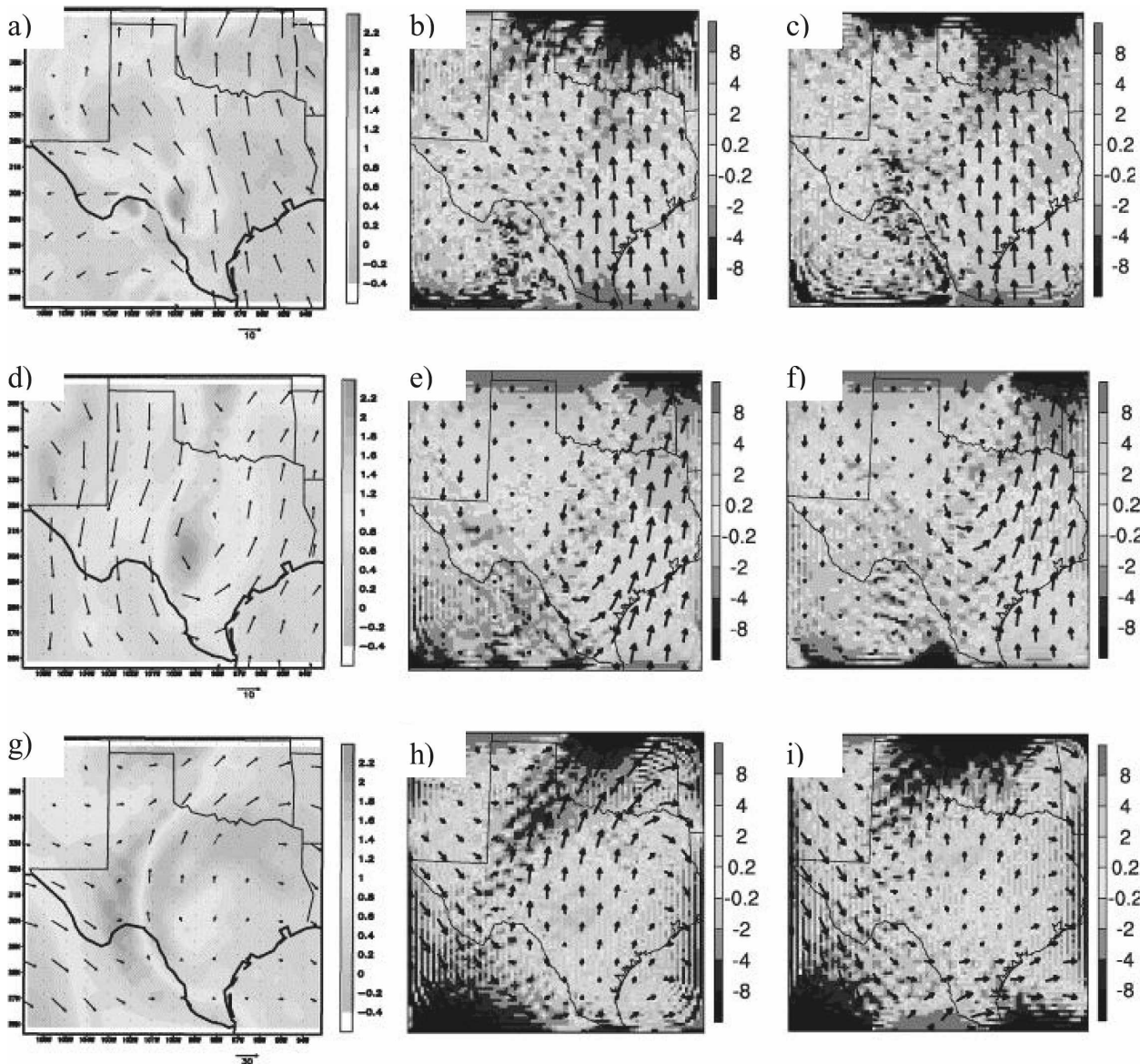
FIG. 14. Vector winds (m s$^{-1}$) and divergence (shading; $\times 10^{-6}$ s$^{-1}$) at three levels in the atmosphere, compared between observed winds (from NARR), best-fit WRF output, and poor-fit WRF (domain PN) output: at 850mb for (a) NARR, (b) best fit, and (c) poor fit; at 500 mb for (d) NARR, (e) best fit, and (f) poor fit; and at 200 mb for (g) NARR, (h) best fit, and (i) poor fit.

larger for time-lagged experiments. For example, perturbing the precipitation $-12$ h results in a mean absolute error (MAE) of 213.3%, while perturbing the precipitation $+12$ h results in an MAE of only 58.4%. In the case of a negative lag, the peak also increases with increased lag time, due to increased time of concentration, which allows the basin more time to accumulate and then translate runoff to the river network. The effect of the peak shape perturbations is most pronounced for an erroneously wide peak (Fig. 17c); in this case the overall volume of the streamflow is increased

to the extent that the hydrograph peak is not well defined compared to the best-fit result.

## 5. Conclusions and implications

This paper studies the sensitivity of WRF model precipitation forecasts to different physics, initialization, and domain configurations, and examines the ability of WRF to accurately reproduce a summer convective rainfall event. The propagation of error from rainfall to streamflow is investigated by applying a variety of per-

TABLE 9. Spatial variability of streamflow error. A precipitation error in WRF of 2 mm h$^{-1}$ over the basin area is used as an example. The ratio of best-fit to modeled streamflow (CTL/EXP) shows the overestimation (CTL/EXP < 1) and underestimation (CTL/EXP > 1) at various gauges across the basin. Time increments are averages during the peak 72 h of the storm.

| Time increment (h) | 1–12 | 13–24 | 25–36 | 37–48 | 49–60 | 61–72 | Mean storm $R$ |
|---|---|---|---|---|---|---|---|
| CTL/EXP: upstream | | | | | | | |
| Gauge 1 | 0.56 | 0.41 | 0.30 | 0.30 | 0.36 | 0.31 | 0.94 |
| Gauge 2 | 0.21 | 0.09 | 0.07 | 0.07 | 0.11 | 0.10 | 0.94 |
| Gauge 3 | 0.99 | 0.68 | 0.72 | 0.78 | 0.99 | 0.76 | 0.96 |
| Gauge 4 | 0.44 | 0.25 | 0.20 | 0.23 | 0.34 | 0.29 | 0.97 |
| CTL/EXP: midstream | | | | | | | |
| Gauge 5 | 0.25 | 0.09 | 0.06 | 0.06 | 0.10 | 0.08 | 0.95 |
| Gauge 6 | 0.31 | 0.59 | 0.52 | 0.45 | 0.62 | 0.50 | 0.98 |
| Gauge 7 | 0.18 | 0.08 | 0.05 | 0.05 | 0.07 | 0.06 | 0.97 |
| Gauge 8 | 1.06 | 0.71 | 0.70 | 0.74 | 0.96 | 0.75 | 0.96 |
| Gauge 9 | 0.18 | 0.07 | 0.01 | 0.01 | 0.06 | 0.05 | 0.97 |
| CTL/EXP: downstream | | | | | | | |
| Gauge 10 | 0.22 | 0.10 | 0.06 | 0.04 | 0.06 | 0.05 | 0.99 |
| Gauge 11 | 0.79 | 0.40 | 0.22 | 0.15 | 0.20 | 0.20 | 0.98 |
| Gauge 12 | 25.71 | 13.43 | 6.23 | 1.72 | 0.82 | 0.34 | 0.99 |

turbations to the precipitation and finding the relative significance of each in terms of streamflow.

WRF's ability to accurately simulate a July 2002 central Texas storm event is most sensitive to cumulus parameterization, slightly sensitive to microphysics parameterization, and nearly unaffected by the choice of shortwave and longwave radiation schemes. The Betts–Miller–Janjić cumulus scheme combined with Lin mi-

TABLE 10. Intensity, location, and timing error characteristics for streamflow: mean absolute error, bias, and standard deviation of streamflow resultingm each of 45 perturbations to the WRF output rainfall grid.

| Intensity error (mm h$^{-1}$) | Basin MAE | Mean bias* | Bias std dev | Intensity error (mm h$^{-1}$) | Basin MAE | Mean bias* | Bias std dev |
|---|---|---|---|---|---|---|---|
| +0.01 | 10.1 | −0.091 | 0.058 | −0.01 | 0.6 | 0.077 | 0.048 |
| +0.05 | 5.0 | −0.455 | 0.292 | −0.05 | 2.9 | 0.383 | 0.240 |
| +0.10 | 10.1 | −0.912 | 0.583 | −0.10 | 5.6 | 0.760 | 0.479 |
| +0.50 | 56.7 | −4.644 | 2.909 | −0.50 | 23.7 | 3.570 | 2.303 |
| +1.00 | 130.8 | −9.472 | 5.818 | −1.00 | 40.3 | 6.605 | 4.337 |
| +2.00 | 342.2 | −19.529 | 11.652 | −2.00 | 60.6 | 11.035 | 7.403 |
| +3.00 | 641.8 | −29.955 | 17.481 | −3.00 | 68.8 | 13.188 | 9.040 |
| +4.00 | 1023.7 | −40.632 | 23.285 | −4.00 | 72.1 | 14.225 | 9.944 |
| +5.00 | 1479.1 | −51.475 | 29.070 | −5.00 | 73.7 | 14.743 | 10.409 |
| Location error | Basin MAE | Mean bias* | Bias std dev | Timing errors | Basin MAE | Mean bias* | Bias std dev |
| Shift1NE | 7.8 | 1.075 | 1.022 | Wide peak | 35.7 | −9.897 | 8.876 |
| Shift1NW | 9.7 | 0.057 | 0.517 | Narrow peak | 15.2 | −4.455 | 4.123 |
| Shift1SE | 10.9 | 0.237 | 0.545 | Flash peak | 10.1 | −2.909 | 2.688 |
| Shift1SW | 6.5 | −0.800 | 0.825 | +3 h | 27.9 | 0.430 | 2.755 |
| Shift3NE | 30.6 | 3.757 | 3.585 | +6 h | 25.1 | 0.174 | 2.045 |
| Shift3NW | 24.3 | 1.175 | 1.598 | +12 h | 51.3 | 0.605 | 3.797 |
| Shift3SE | 44.6 | 1.173 | 1.999 | +18 h | 85.2 | 1.064 | 5.240 |
| Shift3SW | 17.6 | −1.589 | 1.936 | −3 h | 18.1 | 0.234 | 1.041 |
| Shift5NE | 73.5 | 6.239 | 5.717 | −6 h | 49.3 | 0.506 | 1.904 |
| Shift5NW | 35.1 | 3.092 | 2.822 | −12 h | 206.7 | 0.987 | 3.415 |
| Shift5SE | 106.9 | 2.286 | 3.185 | −18 h | 603.6 | 1.402 | 4.607 |
| Shift5SW | 25.6 | −1.003 | 2.068 | | | | |
| Shift10NE | 227.8 | 8.815 | 7.001 | | | | |
| Shift10NW | 53.3 | 7.844 | 5.897 | | | | |
| Shift10SE | 115.6 | 6.157 | 6.186 | | | | |
| Shift10SW | 40.1 | 0.755 | 2.447 | | | | |

* ft$^3$ s$^{-1}$.

(a)



(b)
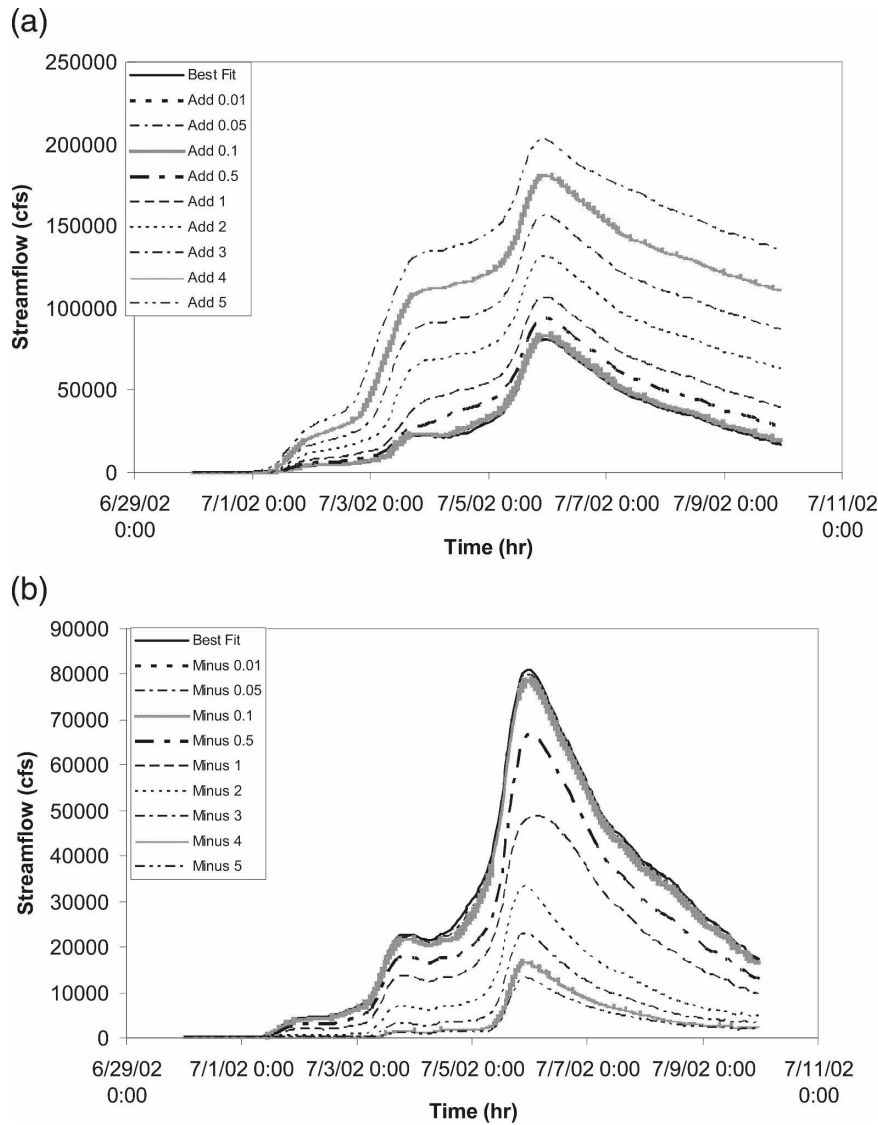


FIG. 15. Effect of (a) positive and (b) negative rainfall intensity errors (mm h$^{-1}$) on streamflow. Boldface line is the output from the best-fit WRF rainfall grid.

crophysics produces the most reasonable rainfall over the San Antonio River basin. Other cumulus schemes tend to underestimate rainfall intensity and shift the peak away from its true location. Contrary to modeling studies suggesting that domains with close grid spacing have the ability to resolve convective motions on their own, this research found that simulation of the 11-day convective storm requires parameterization of convection in the inner 4-km domain.

A 48-h initialization interval produces optimum performance in WRF, as determined by the high correlation values and low errors in both rainfall and streamflow. This interval allows WRF to successfully produce precipitation on its own, and the forecast is not sub-

stantially degraded, as in longer forecasts. Investigation of various domain configurations suggests it is best to avoid placing any domain over oceans or other regions lacking sufficient data. The relative size and placement of the parent and nested grid(s) are also important; to best simulate the combined effects of synoptic, mesoscale, and local atmospheric forcings, the nest must be close to the area of flood interest while being kept a significant distance from the parent domain. Additionally, studies of domain grid spacing should be undertaken; Zhang et al. (2006) found that a finely gridded nest (3.3 km) did not produce the best simulation (using the MM5 model) compared to coarser nest grids. It would be useful to test alternate grid spacings to deter-
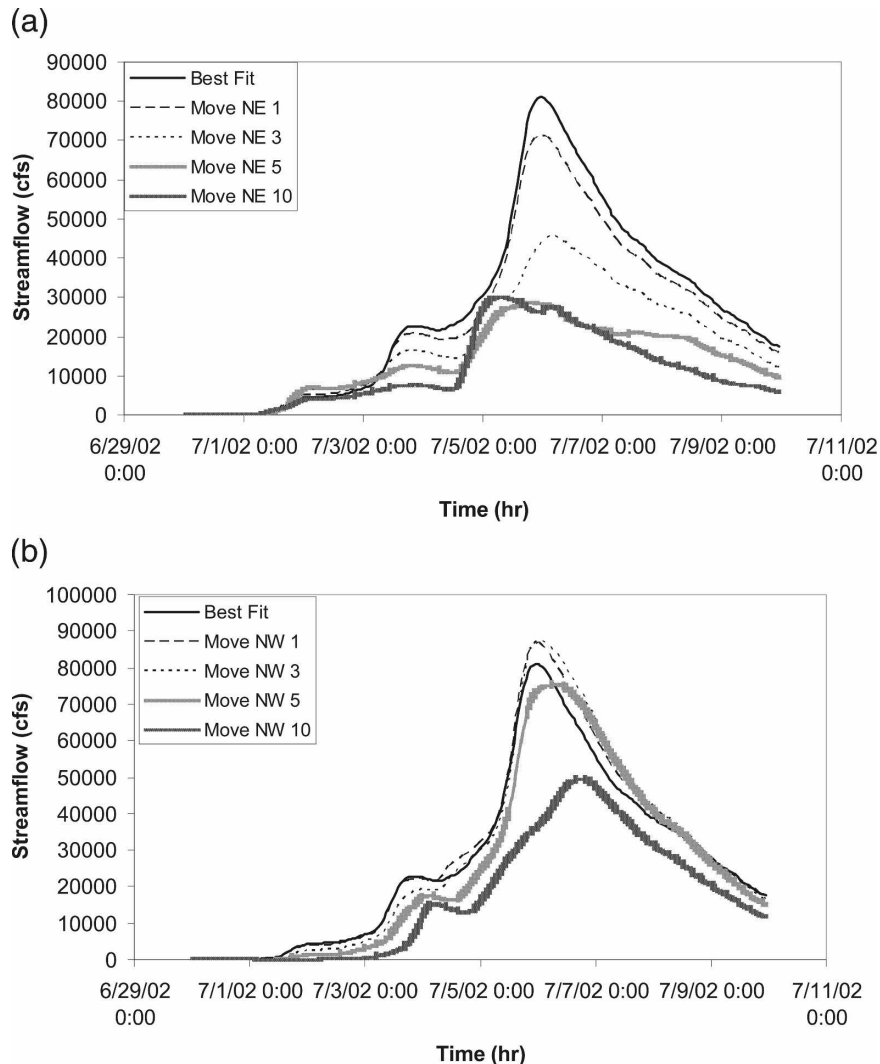
(a)



(b)



Fig. 16. Effects of rainfall location errors on streamflow. Units are number of grid cells that the best-fit output is shifted (a) northeast, (b) northwest, (c) southeast, and (d) southwest. Boldface line is the output from the best-fit WRF rainfall grid.

mine if forecast degradation occurs at high resolutions in WRF.

The best-fit WRF model accurately reproduces precipitation patterns and intensity over the warm season, convective central Texas storm of interest. Comparisons of observed and modeled wind vectors at three different atmospheric levels suggest that errors in horizontal winds for the poor-performing WRF configurations may be a major factor in producing insufficient precipitation. There are several implications from the success of WRF in simulating the July 2002 storm and its subsequent streamflow. For data-sparse areas or International River Basins (IRB) nations that do not have access to upstream data, due to either a lack of infrastructure or to a lack of data sharing between coinci-

dent riparian nations, WRF output can provide a proxy for such data. However, WRF may not increase the lead time enough to significantly improve flood forecasting in these regions. Additionally, WRF output can be used as the driving meteorological forcing data for regional- and watershed-scale hydrological models.

Current numerical weather models categorically contain small errors associated with initial and boundary conditions, which result in errors in the precipitation forecast. This research investigates numerous perturbations to the WRF rainfall grid and their effects on streamflow output after running through a hydrological model. Results indicate that 1) error dampening from rainfall to streamflow is greatest when the initial rainfall perturbation caused an increase in rainfall relative
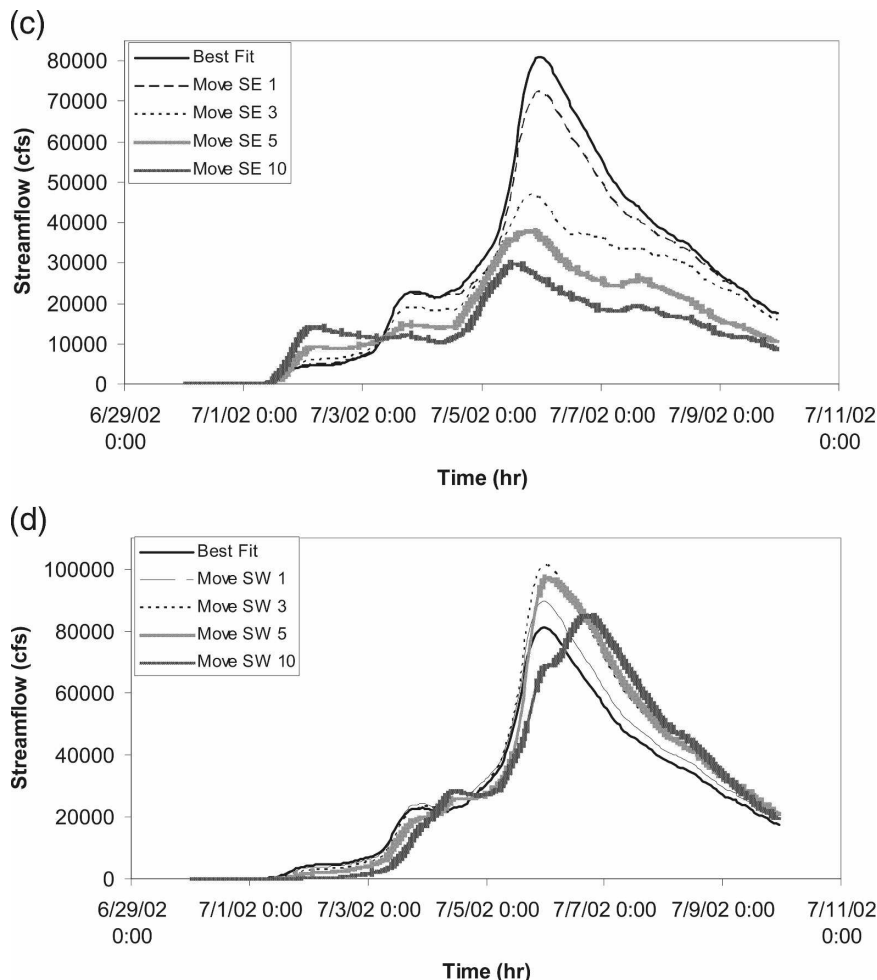
(c)



(d)



FIG. 16. (*Continued*).

to the control and 2) translation of error from precipitation to streamflow is strongly affected by surface characteristics such as topography, land use, and soils. Intensity and timing errors propagate fairly linearly to streamflow error, which therefore may be simpler to predict. Simulations that result in location error are most important because the effects of erroneous precipitation locations on streamflow are dependent on the pattern of observed precipitation. Streamflow effects from these errors are nonlinear and difficult to determine for operational flood prediction.

The study of error in WRF may help to establish limits on acceptable rainfall error for flood forecasting of large events. For example, WRF rainfall errors less than or equal to $\pm 0.5$ mm h$^{-1}$ appear to have a negligible effect on the quality of a streamflow forecast, as do timing errors of 6 h or less. This type of research can serve as a guideline for error estimation and correction in operational modeling with the WRF model. Further research is necessary, using a wider range of perturbation experiments, to better define the relative effects of different error types. Future publications on this subject should include more quantitative analysis, such as the growth rate of error for each perturbation type. In addition, a finescale precipitation product such as NOAA stage III should be used for more detailed verifications of results.
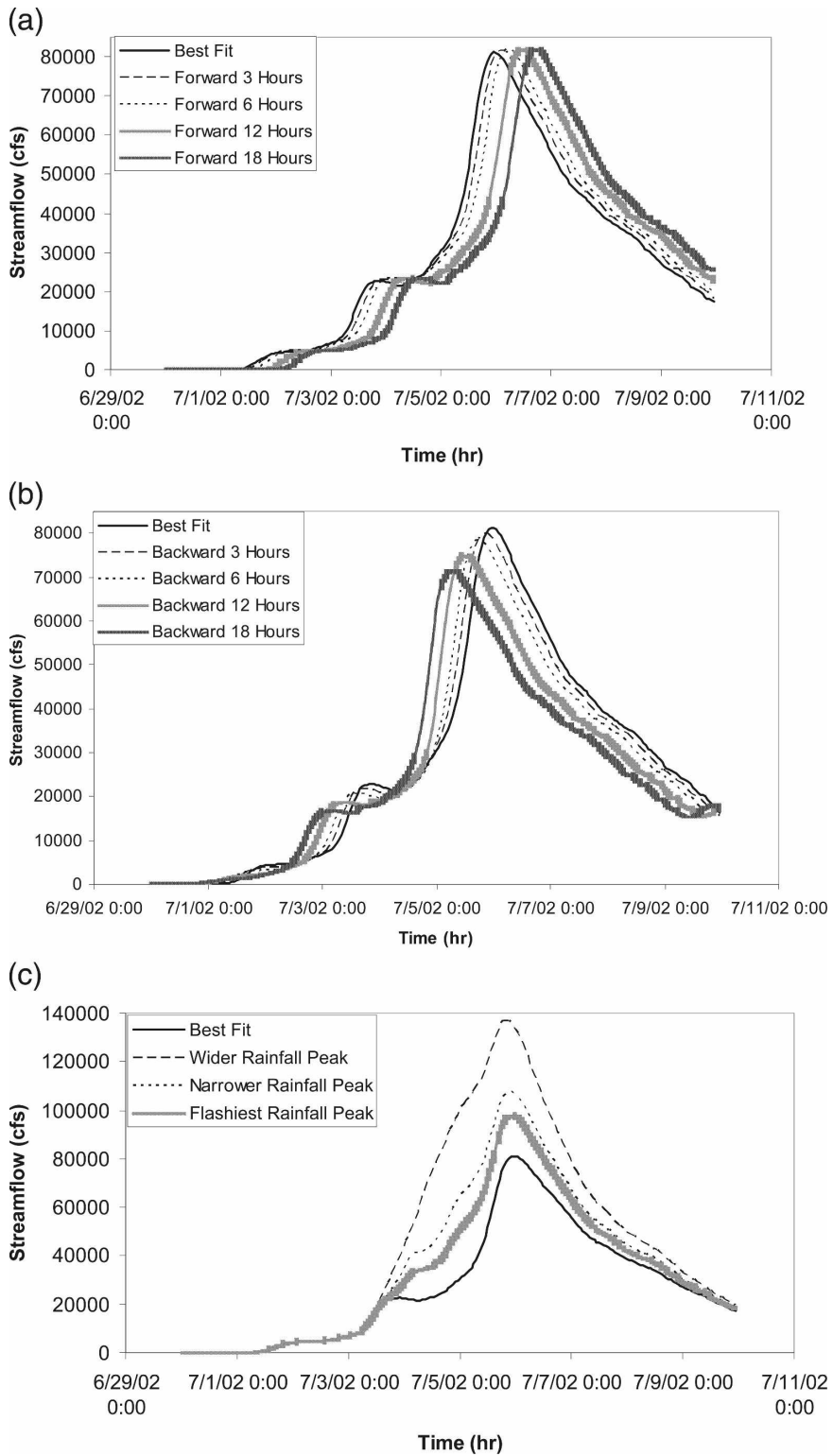
(a)



(b)



(c)



FIG. 17. Effect of rainfall timing errors on streamflow: (a) forward time error, (b) backward time error, and (c) peak shape error. Boldface line is the output from the best-fit WRF rainfall grid.

necessarily reflect the policies and views of DHS, DOE, or ORISE. Software support for the WRF model was provided by UCAR user support. The authors thank Guo-Yue Niu, Jeff Lo, and Xiaoyan Jiang for their insight and suggestions on the WRF model and experimental design. We also appreciate the comments of two anonymous reviewers that helped to improve the original manuscript.

REFERENCES

Baik, J.-J., M. DeMaria, and S. Raman, 1991: Tropical cyclone simulations with the Betts convective adjustment scheme. Part III: Comparisons with the Kuo convective parameterization. *Mon. Wea. Rev.,* **119,** 2889–2899.

Bélair, S., and J. Mailhot, 2001: Impact of horizontal resolution on the numerical simulation of a midlatitude squall line: Implicit versus explicit condensation. *Mon. Wea. Rev.,* **129,** 2362–2376.

Bonner, W. D., 1968: Climatology of the low-level jet. *Mon. Wea. Rev.,* **96,** 833–850.

Déry, S. J., E. F. Wood, and C. Kerr, cited 2007: Comparing NARR, ERA-40, and observed Hudson Bay river discharge. Preprints, *North American Regional Reanalysis Workshop,* San Diego, CA, Amer. Meteor. Soc. [Available online at http://www.emc.ncep.noaa.gov/mmb/rreanl/.]

Emanuel, K. A., and M. Zivkovic-Rothman, 1999: Development and evaluation of a convection scheme for use in climate models. *J. Atmos. Sci.,* **56,** 1766–1782.

Ferrier, B. S., 1994: A double-moment multiple-phase four-class bulk ice scheme. Part I: Description. *J. Atmos. Sci.,* **51,** 249–280.

Fovell, R. G., 2006: Impact of microphysics on hurricane track and intensity forecasts. Preprints, *Seventh WRF Users' Workshop,* National Center for Atmospheric Research, Boulder, CO, 3.2. [Available online at http://www.mmm.ucar.edu/wrf/users/workshops/WS2006/abstracts/Session03/3_2_Fovell.pdf.]

Fritsch, J. M., and R. E. Carbone, 2004: Improving quantitative precipitation forecasts in the warm season. *Bull. Amer. Meteor. Soc.,* **85,** 955–965.

Gallus, W. A., Jr., and M. Segal, 2001: Impact of improved initialization of mesoscale features on convective system rainfall in the 10-km Eta simulations. *Wea. Forecasting,* **16,** 680–696.

Giorgi, F., and L. O. Mearns, 1999: Introduction to special section: Regional climate modeling revisited. *J. Geophys. Res.,* **104,** 6335–6352.

——, and C. Shields, 1999: Tests of precipitation parameterizations available in latest version of NCAR Regional Climate Model (RegCM) over continental United States. *J. Geophys. Res.,* **104,** 6353–6375.

Gochis, D. J., W. J. Shuttleworth, and Z.-L. Yang, 2002: Sensitivity of the modeled North American monsoon regional climate to convection parameterization. *Mon. Wea. Rev.,* **130,** 1282–1298.

Higgins, R. W., W. Shi, E. Yarosh, and R. Joyce, 2000: Improved United States precipitation quality control system and analysis. *NCEP/Climate Prediction Center Atlas 7,* 40 pp.

Hirschboeck, K. K., 1987: Catastrophic flooding and atmospheric circulation anomalies. *Catastrophic Flooding,* L. Mayer and D. Nash, Eds., Binghamton Symposia in Geomorphology: International Series, No. 18, Elsevier, 23–56.

Hong, S. Y., and J. O. J. Lim, 2006: The WRF single-moment 6-class microphysics scheme (WSM6). *J. Kor. Meteor. Soc.,* **42,** 129–151.

Janjić, Z. I., 1994: The step-mountain Eta coordinate model: Further development of the convection, viscous sublayer, and turbulence closure schemes. *Mon. Wea. Rev.,* **122,** 927–945.

Jankov, I., and W. A. Gallus Jr., 2004: MCS rainfall forecast accuracy as a function of large-scale forcing. *Wea. Forecasting,* **19,** 428–439.

——, ——, M. Segal, B. Shaw, and S. E. Koch, 2005: The impact of different WRF model physical parameterizations and their interactions on warm season MCS rainfall. *Wea. Forecasting,* **20,** 1048–1060.

Juang, H.-M. H., and S.-Y. Hong, 2001: Sensitivity of the NCEP Regional Spectral Model to domains size and nesting strategy. *Mon. Wea. Rev.,* **129,** 2904–2922.

Kain, J. S., and J. M. Fritsch, 1998: Multiscale convective overturning in mesoscale convective systems: Reconciling observations, simulations, and theory. *Mon. Wea. Rev.,* **126,** 2254–2273.

Kessler, E., 1969: *On the Distribution and Continuity of Water Substance in Atmospheric Circulation. Meteor. Monogr.,* No. 32, Amer. Meteor. Soc., 84 pp.

Knebl, M. R., Z.-L. Yang, K. Hutchison, and D. R. Maidment, 2005: Regional scale flood modeling using NEXRAD rainfall, GIS, and HEC-HMS/RAS: A case study for the San Antonio River Basin summer 2002 storm event. *J. Environ. Manage.,* **75,** 325–336.

Leung, L. R., Y. Qian, and X. Bian, 2003: Hydroclimate of the western United States based on observations and regional climate simulation of 1981–2000. Part I: Seasonal statistics. *J. Climate,* **16,** 1892–1911.

Liang, X.-Z., K. E. Kunkel, and A. N. Samel, 2001: Development of a regional climate model for U.S. Midwest applications. Part I: Sensitivity to buffer zone treatment. *J. Climate,* **14,** 4363–4378.

——, L. Li, A. Dai, and K. E. Kunkel, 2004a: Regional climate model simulation of summer precipitation diurnal cycle over the United States. *Geophys. Res. Lett.,* **31,** L24208, doi:10.1029/2004GL021054.

——, ——, and K. E. Kunkel, 2004b: Regional climate model simulation of U.S. precipitation during 1982-2002. Part I: Annual cycle. *J. Climate,* **17,** 3510–3529.

Lin, Y. L., R. D. Farley, and H. D. Orville, 1983: Bulk parameterization of the snow field in a cloud model. *J. Climate Appl. Meteor.,* **22,** 1065–1092.

Marchok, T., R. Rogers, and R. Tuleya, 2007: Validation schemes for tropical cyclone quantitative precipitation forecasts: Evaluation of operational models for U.S. landfalling cases. *Wea. Forecasting,* **22,** 726–746.

Mesinger, F., and Coauthors, 2006: North American Regional Reanalysis: A long-term, consistent, high-resolution climate dataset for the North American domain, as a major improvement upon the earlier global reanalysis datasets in both resolution and accuracy. *Bull. Amer. Meteor. Soc.,* **87,** 343–360.

Michalakes, J., J. Dudhia, D. Gill, T. Henderson, J. Klemp, W. Skamarock, and W. Wang, 2005: The Weather Research and Forecast model: Software architecture and performance. *Proc. 11th ECMWF Workshop on the Use of High Performance Computing in Meteorology,* W. Zwieflhofer and G. Mozdzynski, Eds., World Scientific, 156–168.

Mitchell, M. J., R. W. Arritt, and K. Labas, 1995: A climatology of

the warm season Great Plains low-level jet using wind profiler observations. *Wea. Forecasting,* **10,** 576–591.

Nielsen-Gammon, J. W., F. Zhang, A. Odins, and B. Myoung, 2005: Extreme rainfall events in Texas: Patterns and predictability. *Phys. Geogr.,* **26,** 340–364.

Nigam, S., and A. Ruiz-Barradas, 2006: Seasonal hydroclimate variability over North America in global and regional reanalysis and AMIP simulations: Varied representation. *J. Climate,* **19,** 815–837.

Olson, D. A., N. W. Junker, and B. Korty, 1995: Evaluation of 33 years of quantitative precipitation forecasting. *Wea. Forecasting,* **10,** 498–511.

Patton, P. C., and V. R. Baker, 1977: Geomorphic response of central Texas stream channels to catastrophic rainfall and runoff. *Geomorphology in Arid Regions,* D. O. Doehring, Ed., SUNY Binghamton Publications in Geomorphology, Elsevier, 189–217.

Seth, A., and F. Giorgi, 1998: The effects of domain choice on summer precipitation simulation and sensitivity in a regional climate model. *J. Climate,* **11,** 2698–2712.

Smedsmo, J. L., E. Foufoula-Georgiou, and V. Vuruputur, 2005: On the vertical structure of modeled and observed deep convective storms: Insights for precipitation retrieval and microphysical parameterization. *J. Appl. Meteor.,* **44,** 1866–1884.

Thompson, G., R. M. Rasmussen, and K. Manning, 2004: Explicit forecasts of winter precipitation using an improved bulk microphysics scheme. Part I: Description and sensitivity analysis. *Mon. Wea. Rev.,* **132,** 519–542.

Ting, M., and H. Wang, 2006: The role of the North American topography on the maintenance of the Great Plains summer low-level jet. *J. Atmos. Sci.,* **63,** 1056–1068.

Vaidya, S. S., and S. S. Singh, 2000: Applying the Betts–Miller–Janjić scheme of convection in prediction of the Indian monsoon. *Wea. Forecasting,* **15,** 349–356.

Wang, D., G. Wang, and E. N. Anagnostou, 2005: Use of satellite-based precipitation observation in improving the parameterization of canopy hydrological processes in land surface models. *J. Hydrometeor.,* **6,** 745–763.

Wang, W., and N. L. Seaman, 1997: A comparison study of convective parameterization schemes in a mesoscale model. *Mon. Wea. Rev.,* **125,** 253–278.

Warner, T. T., and H. M. Hsu, 2000: Nested-model simulation of moist convection: The impact of coarse-grid parameterized convection on fine-grid resolved convection. *Mon. Wea. Rev.,* **128,** 2211–2231.

Westrick, K. J., P. Storck, and C. F. Mass, 2002: Description and evaluation of a hydrometeorological forecast system for mountainous watersheds. *Wea. Forecasting,* **17,** 250–262.

Xu, J., and E. E. Small, 2002: Simulating summertime rainfall variability in the North American monsoon region: The influence of convection and radiation parameterizations. *J. Geophys. Res.,* **107,** 4727, doi:10.1029/2001JD002047.

Zhang, F., A. M. Odins, and J. W. Nielson-Gammon, 2006: Mesoscale predictability of an extreme warm-season precipitation event. *Wea. Forecasting,* **21,** 149–166.