Comparison of the Diurnal Precipitation Cycle in Convection-Resolving and Non-Convection-Resolving Mesoscale Models

ADAM J. CLARK, WILLIAM A. GALLUS JR., AND TSING-CHANG CHEN

Department of Geological and Atmospheric Sciences, Iowa State University, Ames, Iowa

(Manuscript received 31 October 2006, in final form 11 January 2007)

ABSTRACT

The diurnal cycles of rainfall in 5-km grid-spacing convection-resolving and 22-km grid-spacing non-convection-resolving configurations of the Weather Research and Forecasting (WRF) model are compared to see if significant improvements can be obtained by using fine enough grid spacing to explicitly resolve convection. Diurnally averaged Hovmöller diagrams, spatial correlation coefficients computed in Hovmöller space, equitable threat scores (ETSs), and biases for forecasts conducted from 1 April to 25 July 2005 over a large portion of the central United States are used for the comparisons. A subjective comparison using Hovmöller diagrams of diurnally averaged rainfall show that the diurnal cycle representation in the 5-km configuration is clearly superior to that in the 22-km configuration during forecast hours 24–48. The superiority of the 5-km configuration is validated by much higher spatial correlation coefficients than in the 22-km configuration. During the first 24 forecast hours the 5-km model forecasts appear to be more adversely affected by model “spinup” processes than the 22-km model forecasts, and it is less clear, subjectively, which configuration has the better diurnal cycle representation, although spatial correlation coefficients are slightly higher in the 22-km configuration. ETSs in both configurations have diurnal oscillations with relative maxima occurring in both configurations at forecast hours corresponding to 0000–0300 LST, while biases also have diurnal oscillations with relative maxima (largest errors) in the 22-km (5-km) configuration occurring at forecast hours corresponding to 1200 (1800) LST. At all forecast hours, ETSs from the 22-km configuration are higher than those in the 5-km configuration. This inconsistency with some of the results obtained using the aforementioned spatial correlation coefficients reinforces discussion in past literature that cautions against using “traditional” verification statistics, such as ETS, to compare high- to low-resolution forecasts.

1. Introduction

One of the contributing factors to the difficulties in forecasting warm season rainfall in much of the central United States is the inability of current operational forecast models to correctly simulate the diurnal cycle of rainfall (e.g., Davis et al. 2003). This deficiency can be revealed by computing long-term averages of rainfall statistics in forecast models and comparing these averages to the corresponding observations. Recent work has shown that when these gross rainfall statistics are computed and displayed in a time–longitude format, systematic timing and longitude errors in models using cumulus parameterization schemes (CPSs) become apparent. These models fail to replicate the coherent propagating rainfall axis observed in the central United States (Davis et al. 2003) despite the intrinsic predictability suggested by the frequent occurrence of long-lived convective episodes during the warm season (Carbone et al. 2002) that contribute to much of the rainfall occurring along these axes.

The diurnal cycle depiction is especially poor in regimes characterized by weak forcing (Liu et al. 2006) and nocturnal convection (Davis et al. 2003). Typical errors include premature initiation of convection during the day over the central and eastern United States that can lead to a diurnal cycle that is almost exactly opposite in phase to that observed (Davis et al. 2003), overpredicted areal coverage of rainfall, and spurious westward propagating events (Liu et al. 2006).

The root cause of these errors is widely believed to be shortcomings associated with CPSs (e.g., Fritsch and Carbone 2004), which are still a necessity in operational models given current computational limitations. These
shortcomings include the difficulties associated with applying a CPS at grid spacings in the middle of a continuous spectrum of potentially highly energetic motion (Kain and Fritsch 1998). More recent studies have clearly shown that CPSs have difficulties simulating propagating convection, perhaps from the inability at grid spacing above 10 km to sufficiently resolve the downdrafts that lead to the cold pool dynamics associated with propagation (Davis et al. 2003). Bukovsky et al. (2006) note that CPSs are not specifically designed to simulate propagation because the schemes act independently in individual model columns. These findings imply that models using CPSs are not well suited for precipitation forecasts in areas like the Midwest where the majority of warm season rainfall comes from propagating mesoscale convective systems (MCSs; Fritsch et al. 1986). Other problems with CPSs can be linked to crude trigger functions (Liu et al. 2006) and the lack of mesoscale organization, a resolvable phenomena that initially develops from subgrid-scale processes that must be parameterized (Molinari and Dudek 1992). Also, CPSs have been linked to a rainfall intensity spectrum that is too narrow (Davis et al. 2006a).

Unlike the forecasts of other meteorological variables that have been improved through a refinement in resolution (e.g., Mass et al. 2002), it has been found that decreasing the grid spacing may have little benefit for forecasts of precipitation in mesoscale models using CPSs (Gallus 1999). Thus, many researchers believe that significant improvements in warm season rainfall forecasting will only be realized when operational models are run at a fine enough grid spacing to explicitly resolve convection (Fritsch and Carbone 2004). Advancements in computational technology may soon allow such fine grid spacing. The coarsest grid spacing needed to resolve MCSs has been shown to be around 4 km (Weisman et al. 1997).

Results from experiments utilizing the explicit treatment of convection have been promising. Simulations using grids that explicitly resolve convection have been found to produce qualitatively good forecasts of the expected nature of severe weather (Bernardet et al. 2000). Comparing a sample of 4-km grid-spacing simulations explicitly representing convection to 10-km simulations that parameterize convection, Done et al. (2004) found that 4-km simulations predicted more MCSs that corresponded to those observed and had a better depiction of MCS structure than the 10-km grid-spacing model. In addition, the rainfall intensity spectrum has been found to be more realistic in convection-resolving models (Davis et al. 2006b) along with warm season rainfall coherence for a small number of cases (Liu et al. 2006). In contrast, other studies have shown that a refinement to convection-resolving resolution does not lead to the improvement of more traditional, gridpoint-based skill measures (e.g., Mass et al. 2002), implying that there is a point of diminishing returns for high-resolution forecasts when gridpoint-based measures are applied, even when a subjective evaluation of the high-resolution forecasts shows a distinct advantage. This behavior can result from large errors caused by small displacements in small areas of intense rainfall forecast by high-resolution models that are not resolved by coarser models (Baldwin et al. 2001; Davis et al. 2006a).

The goal of this paper is to compare the diurnal cycle of rainfall for a relatively large set of simulations during the period 1 April to 25 July 2005 in two configurations of the Weather Research and Forecasting (WRF) model (Skamarock et al. 2001; Michalakes et al. 2001) that explicitly resolve and parameterize convection. Part of the analysis is done by comparing diurnally averaged rainfall statistics displayed in a time–longitude format similar to the methodology utilized by Davis et al. (2003). Additional analysis is performed by computing diurnally averaged equitable threat scores (ETSs) and biases. Particular attention is given to the ability of each configuration to represent the coherent propagating axis of rainfall in the central United States and to the model rainfall patterns causing diurnal oscillations of ETS and bias. The paper is organized as follows: section 2 includes a model description, section 3 describes the data and methodology, section 4 examines the results, and section 5 offers conclusions and possibilities for future work.

2. Model description

Forecast rainfall data from 3-hourly accumulation periods of simulations conducted from 1 April to 25 July 2005 with 22- and 5-km grid-spacing configurations of the WRF model were examined. The 5- and 22-km grid-spacing configurations both had 38 vertical levels, used the nonhydrostatic mesoscale model (NMM) dynamic core (Janjic 2003), and are hereafter referred to as NMM5 and NMM22, respectively. The NMM5 simulations were conducted as part of the Developmental Testbed Center (DTC) NMM Spring Forecast Experiment, an extension of the DTC Winter Forecast Experiment (DWFE; Bernardet et al. 2005). More information on these experiments can be found online at http://www.dtcenter.org/projects/projects.php. The NMM22 simulations were conducted on the 64-processor computing cluster in the meteorology program at Iowa State University. It should be noted that simulations were also performed over the entire period...
with a 22-km version of the WRF model using the Advanced Research WRF (ARW) dynamic core. Despite the difference in dynamic core, the diurnal cycle representation in the ARW runs closely resembles that in the NMM runs, and further discussion will emphasize only the NMM22 runs.

Both configurations were run with a domain encompassing most of the continental United States (Fig. 1). All model integrations were initialized at 0000 UTC and integrated 48 h using initial and boundary condition data from 40-km grid-spacing Eta Model (now known as NAM; Mesinger 1998; Janjic 1994; Black 1994; Rogers et al. 1998) datasets interpolated from the 12-km grid-spacing operational run performed at the National Centers for Environmental Prediction (NCEP). The time frequency of the lateral boundary condition updates was 3 h. The physics parameterizations for both configurations consisted of the Ferrier microphysical scheme (Ferrier et al. 2002), the Mellor–Yamada–Janjic planetary boundary layer scheme (Janjic 1990, 1996, 2002; Mellor and Yamada 1982), and the Noah land surface model (Ek et al. 2003). NMM22 used the Kain–Fritsch CPS (Kain and Fritsch 1993); NMM5 did not use a CPS. Matching subsections of the model domains were extracted (Fig. 1) so that all areas of the domain corresponded to regions in which observed precipitation data were available. These subsections are also centered over the central United States, the region in which propagating MCSs are most frequent (Fritsch et al. 1986).

During the time period used for this study there were dates when data from one or both configurations were not available. Only the dates when data from both configurations were available (Fig. 2) are used in the construction of averaged rainfall plots and calculations of skill measures.

3. Data and methodology

The observations used for verification are 3-hourly observed precipitation fields derived from the 1-hourly NCEP stage IV multisensor analyses (Baldwin and Mitchell 1997). It should be noted that the stage IV multisensor data have been found to be wetter for rainfall amounts under 0.5 in. in 24 h than gauge-only data and drier for heavier amounts (Schwartz and Benjamin 2000). For calculations of skill measures the stage IV data were interpolated to both NMM5 and NMM22 grids using procedures utilized at NCEP that conserve the total amount of liquid in the domain. The stage IV analyses were interpolated to the grids of both model configurations, rather than interpolating all data to one independent grid, so that the effects of the differing grids would be retained.

Time–longitude, or Hovmöller, diagrams were constructed by computing meridional averages of forecast and observed 3-hourly accumulated precipitation between 29° and 49°N. This differs from similar previous works (e.g., Davis et al. 2003) that constructed Hovmöller diagrams using the frequency of rainfall events above a specified threshold. The methodology in this study was chosen to retain information on the actual rainfall amounts and still be able to infer information on timing and location. The Hovmöller format is useful
because collapsing one spatial dimension allows for a
time/direction of propagation depiction of phenomena
(Carbone et al. 1998). The distance coordinate used is
degrees of longitude because rainfall systems in the
Midwest typically propagate in the east–west direction.
This analysis technique is routinely used in climate di-
agnostics (e.g., Levey and Jury 1996; Black et al. 1996;
Murtugudde et al. 1996). Recently, these diagrams have
been used to study the life cycle of precipitation sys-
tems using Doppler radar (Carbone et al. 1998; Wilson
et al. 2001; Carbone et al. 2002) and to verify mesoscale
models (Davis et al. 2003). For diurnally averaged Hov-
möller diagrams, spatial correlation coefficients are
computed in Hovmöller space for each 24-h forecast
period (0–24 and 24–48) to obtain a more quantitative
measure of how well the diurnal cycle of each model
configuration corresponds to the observed diurnal
cycle.

ETS (Schaefer 1990) and bias are also used to verify
the 3-hourly forecasts, where

\[
ETS = \frac{(CFA - CHA)}{F + O - CFA - CHA},
\]

(1)

\[
CHA = O \frac{F}{V},
\]

(2)

and

\[
\text{bias} = \frac{F}{O}.
\]

In the above expressions, the variables indicate the
number of grid points at which (i) rainfall was correctly
forecasted to exceed the specified rainfall threshold
(CFA), (ii) rainfall was forecasted to exceed the thresh-
old (F), (iii) rainfall was observed exceeding the thresh-
old (O), and (iv) a correct forecast would occur by
chance (CHA), where V is the total number of evalu-
ated grid points. ETSs range from \( -\frac{1}{2} \) to 1; scores be-
low 0 have no skill and 1 represents a perfect score. Bias
values range from 0 to infinity. Values of bias signifi-
cantly higher (lower) than 1 indicate that the model
notably overpredicted (underpredicted) areal coverage.
Average ETSs and biases can be calculated by averaging
scores from daily forecasts or summing contingency
table elements from all the forecasts and computing the
scores from the summed elements. The first method
gives equal weight to each forecast while the second
gives more weight to larger precipitation events. Both
methods were used, with the only notable differences
occurring in the ETSs, which had higher scores using
the summed contingency table elements (differences
not shown). These higher scores imply that larger, more
widespread precipitation events were associated with
higher ETSs. Only the summed contingency table ele-
ment scores are used in this paper.

4. Results

a. Diurnally averaged Hovmöller diagrams

To subjectively compare the diurnal cycle represen-
tations in the NMM5 and NMM22 to those computed
from stage IV observations, diurnally averaged Hov-
möller diagrams were constructed (Figs. 3a–c). Note
that because both configurations significantly overpre-
dicted rainfall amounts, the scale for the observed rain-
fall amounts was made to be half of the forecast
amounts to more easily compare the axes of nonpropa-
gating and propagating rainfall. The Hovmöller dia-
gram of observed rainfall (Fig. 3c) clearly depicts two
modes of rainfall that are characteristic to this domain
during the warm season. The first is a coherent propa-
gating rainfall axis extending from 105°W to about
92°W longitude and the second a nonpropagating axis

---

**Fig. 2.** Dates for which data from both the NMM5 and NMM22 model configurations were available. The days that are shaded gray
specify days for which data from one or both model configurations were not available. Data were available from both model configu-
rations on 89 days.
that is strongest from 94°W to the eastern edge of the analysis domain (80°W longitude).

During forecast hours 0–24 it is relatively unclear which configuration performs better from a subjective evaluation of the diurnally averaged Hovmöller diagrams (Figs. 3a–c), although spatial correlation coefficients of 0.62 and 0.45 for the NMM22 (Fig. 3a) and NMM5 (Fig. 3b), respectively, indicate the NMM22 has an advantage. During these times, although there is evidence of a propagating signal in the western portion of the domain in the NMM22, the signal is weaker and less coherent in the NMM22 than in the NMM5. In addition, the NMM22 simulates the afternoon rainfall maximum in the eastern part of the domain about 3 h too early. The NMM5 forecasts feature a propagating axis of rainfall that is shifted later in time and to the east of the observed axis. Although both configurations overpredict rainfall, the overprediction is worse in the NMM5. Note that because of the scale difference mentioned above, this overprediction is even greater than what the shading on the plots implies. It is possible that the higher spatial correlation coefficients in the NMM22 are because of the faster “spinup” time in the NMM22 relative to the NMM5 (discussed in detail later). This faster spinup time in the NMM22 should give an advantage over the NMM5 because of the large amounts of observed precipitation occurring in the domain during forecast hours 3-6. In addition, the displacement of the propagating rainfall axis in the NMM5 should hurt the NMM5 spatial correlation coefficients more than the displacement of a less defined and weaker propagating rainfall axis hurts the NMM22 spatial correlation coefficients.

The overprediction in the NMM5 is consistent with past research using models that explicitly resolve convection at grid spacing comparable to this study (e.g., Weisman et al. 1997, 2004) and is likely attributable to the inability of the model to remove convective insta-

![Hovmöller diagrams of diurnally averaged rainfall (in.) from the (a) NMM22, (b) NMM5, and (c) stage IV observations. The numbers on the left-hand side of (a) and (b) are spatial correlation coefficients computed in Hovmöller space between the model and stage IV rainfall corresponding to the 0–24- and 24–48-h forecasts. The domain over which the Hovmöller diagrams were computed is displayed in (a)–(c). The diagonal (horizontal) white lines in (c) mark the major propagating (nonpropagating) rainfall axes. The scale for the stage IV rainfall is exactly half of the observed rainfall. (d) Bias and (e) ETS at the 0.10-, 0.25-, and 0.50-in. rainfall thresholds valid at the times indicated by the y axis of plots (a)–(c).]
bility by sub-cloud-scale eddies not resolvable on the grid scale (Bryan et al. 2003, Molinari and Dudek 1992). Bryan et al. (2003) argues that even at 1-km grid spacing the resolution is insufficient to resolve these sub-cloud-scale eddies. To confirm if this is part of the problem in this study, thermodynamic profiles at grid points that experienced heavy rainfall would need to be examined, but in the present study only rainfall data were available so this was not possible.

The 3-h phase difference for the nonpropagating rainfall axis in the eastern portion of the domain is also a signal that has been identified in past studies using models with CPSs. For example, Baldwin et al. (2001) noted that the Betts–Miller–Janjic (BMJ; Betts 1986; Betts and Miller 1986; Janjic 1994) CPS within the NCEP Eta Model had a tendency to remove capping inversions that are typical during the daytime in the Great Plains, because of its shallow mixing parameterization. This was discussed in Davis et al. (2003) who noted similar behavior from simulations of the WRF–ARW model that used the BMJ CPS. Also, Dai and Trenberth (2004) noticed that moist convection over land, simulated by version 2 of the Community Climate System Model (CCSM2), was initiated about 4 h prematurely.

It appears that the shift of the propagating rainfall axis later in time and to the east of the observed axis is related to the lack of rainfall during the first six forecast hours caused by the time it takes the model to spin up microphysical variables from zero. The use of initial and lateral boundary conditions that have been degraded from 12- to 40-km grid spacing should affect the spinup time because the model is forced to spin up a larger portion of the kinetic energy spectrum than when using the 12-km data. Also, in theory, the forecasts should be better, especially in the short term, when using the 12-km data because the 12-km data is able to resolve smaller-scale features than the 40-km data. To investigate the potential impacts of the grid degradation a sensitivity test was conducted by running 5-km grid spacing WRF–NMM simulations over a large sub-domain of that used in the NMM5 using both the 12- and 40-km initial and lateral boundary conditions for one case in which widespread convection was present at the time of initialization and for much of the 48-h forecast period. There were only subtle differences in the placement and amounts of precipitation in the 5-km forecasts produced from each set of initial and lateral boundary conditions, with neither set of forecasts appearing to be superior. In addition, while the simulations using the 12-km initial and lateral boundary conditions were able to spin up areas of precipitation slightly faster than the simulations using the 40-km data, the differences in the areas of precipitation forecast were very small relative to the observed areas. For example, at forecast hour 3 the simulation using the 12-km (40-km) data was predicting 1.93% (1.10%) of the areal coverage of observed precipitation above 0.25 in. and at forecast hour 6 the simulation using the 12-km (40-km) data was predicting 9.62% (8.84%). The results of this sensitivity test imply that the errors caused by the coarsened grid are small.

More significant improvements during the spinup time may be obtained by using initialization data in which microphysical variables are analyzed (i.e., “hot” start analyses). One such hot start method is the Local Analysis and Prediction System (LAPS; Jian et al. 2003). Jankov et al. (2007) showed that a 12-km grid-spacing version of the WRF model initialized with LAPS was able to spin up the grid-resolved component of precipitation much faster than the same version using the same 40-km initialization data used in the current study. Further investigation is needed to examine whether the use of hot start initialization would result in an improvement in the location and magnitude of the propagating rainfall axis for a large number of cases during the first diurnal cycle.

During the model spinup time, approximately 0000–0600 UTC, the peak in the diurnal mode of observed rainfall frequency typically occurs from about 105° to 95°W in the central high plains (Knievel et al. 2004). This is also the time and location of the strongest propagating signal of rainfall frequency in the United States. MCSs begin to organize during this time, usually with the help of a strengthening low-level jet, and propagate to the east across the western high plains. For the current sample of data, a maximum in observed rainfall centered around 103°W at 0300 UTC moves east at a speed of roughly 20 m s⁻¹ and ends up centered around 95°W by 1200 UTC (Fig. 3c). The area of rainfall appears to reach its peak intensity around 0600–0900 UTC. The implications of this strong propagating rainfall signal are that the mesoscale dynamics (i.e., upscale development of mesoscale circulations, formation of downdrafts leading to propagation attributable to cold pools, interaction with low-level jets, etc.) taking place while the model is spinning up starting at 0000 UTC are very different from the dynamics that would be occurring while a model that was initialized at 1200 UTC was spinning up. By initializing the model at 0000 UTC the challenge of resolving developing/ongoing propagating MCSs, systems that are already difficult to predict without having to worry about spinup issues, is exacerbated.

For the reasons mentioned above, the problems in the NMM5 observed during the times when spinup oc-
curs are not unexpected. However, it is puzzling that the axis of propagation, which was shifted later in time and to the east of the observed axis, extends into the hours at which there should be an observed minimum in rainfall (forecast hours 15 to 18) related to the dissipation of nighttime MCSs. This shift is significant because it means that the spinup effects may not be limited to the times during which spinup occurs; areas downstream from where convection forms during the first 3 to 6 h of the forecast also may be indirectly affected later because of timing and placement errors. Also puzzling is the observation that systems initiating too late also dissipate too late. To determine precisely why the dissipation of rainfall is delayed is beyond the scope of this study. Future research should verify that rainfall systems are in fact dissipating too late (as opposed to anomalous generation of distinct new convection at this time), and possible mechanisms for the late dissipation should be investigated. One possible cause of this behavior may be the improper simulation of the low-level jet. Past studies (e.g., Maddox 1983) have shown that the weakening of the low-level jet during the morning hours caused by the inertial oscillation (Bonner et al. 1968) is a major factor in the dissipation of MCSs during this time. If the low-level jet is properly simulated, the model may not be properly representing the mesoscale dynamics and circulations within MCSs that lead to dissipation. If systems are dissipating properly and new convection is initiating prematurely, future research should investigate problems in the planetary boundary layer parameterization. If too much moisture or heat is present leading to overpredicted instability, this could lead to the premature initiation of convection.

A subjective evaluation of the diurnally averaged Hovmöller diagrams during forecast hours 24–48 clearly shows that the representation of the diurnal cycle is better in the NMM5 (Fig. 3b) than in the NMM22 (Fig. 3a). The propagating axis of rainfall in the western portion of the domain in the NMM5 has shifted so that it is much more aligned with the observed axis than it was during the 0–24-h forecast period, and the NMM5 continues to accurately represent the timing and location of the nonpropagating axis in the eastern portion of the domain as it did during forecast hours 0–24. Also, in the NMM22 the evidence for a propagating signal in the western portion of the domain is even less than that during forecast hours 0–24, and the NMM22 continues to simulate the afternoon rainfall maximum in the eastern part of the domain about 3 h too early, similar to the behavior during forecast hours 0–24. Generally, there was little change in the diurnal cycle representations between the 0–24- and 24–48-h forecast periods in the NMM22, evident by correlation coefficients of 0.62 and 0.69, respectively (Fig. 3a). However, there was major improvement in the diurnal cycle representations between the 0–24 and 24–48-h forecast periods in the NMM5, evident by correlation coefficients of 0.45 and 0.82, respectively (Fig. 3b). It will be shown in a later section that this improvement was because of the ability of the NMM5 to correctly simulate the timing and placement of the propagating axis of rainfall in the western part of the domain.

The improvement in the timing and placement of the propagating rainfall axis in the NMM5 during the second diurnal cycle (forecast hours 24–48; Fig. 3b) is related to model spinup no longer being an issue at these times. To see if high-resolution models can continue to accurately represent the diurnal cycle after forecast hours 24–48, it would be useful to study forecasts going out to at least 72 h, encompassing one more diurnal cycle.

b. Time series of daily Hovmöller diagrams

More information can be gained about the forecasts by constructing time series of Hovmöller diagrams over the entire several-month period for the forecast periods 0–24 and 24–48 h. The forecasts are separated into these two periods because of the differences in the diurnal cycle representations, as discussed in the previous section. Analysis of these time series can show when the errors observed in the averaged plots occurred (such as displaced propagating rainfall axes), and these errors can be matched to the large-scale weather regimes that were occurring at the time. The full set of plots constructed using each model configuration for each forecast period can be viewed online at http://mtarchive.geol.iastate.edu/misc/Hovmoller_diagrams. An example is displayed in Fig. 4.

In the forecast and observations, eastward propagating areas of rainfall are easily identified by diagonal streaks. Multiple occurrences of precipitation “episodes” are observed, defined by Carbone et al. (2002) as time–space clusters of heavy precipitation that often result from sequences of organized convection. The episodes consist of slow, eastward propagating precipitation areas within which there are faster propagating rainfall areas. An example during the time period of this study occurred 3–7 June 2005 (Fig. 4).

It is possible to distinguish between periods in which diurnal or synoptic forcing were likely the dominant forcing mechanisms in generating convection. The periods in which synoptic-scale forcing dominates feature areas of rainfall that move across the domain from west to east over the span of 2 or 3 days with little or no daily variation in rainfall amounts, whereas the periods in
Fig. 4. Hovmöller diagram of rainfall (in.) for (a) 1–15 Jun 2005 and (b) 16–30 Jun 2005, for the forecast hours 0–24. The shading represents the NMM5 forecasts and the contours represent the stage IV observations. The three contour levels match the three levels of shading.
which diurnal forcing dominates feature areas of rainfall that propagate from west to east over a smaller part of the domain and dissipate within 12–24 h on a daily basis. There are also periods when both forcing mechanisms help generate rainfall, similar to the aforementioned precipitation “episodes.”

An example of a time period in which diurnal forcing is the dominant mechanism generating rainfall is 4–11 July 2005 in the longitudinal corridor −108° to −94°W (Figs. 5 and 6). During this period and other similar periods the 0–24 h forecasts from the NMM5 badly underestimate or completely miss areas of propagating rainfall while the 24–48-h forecasts are often markedly better, with forecast propagating rainfall streaks corresponding very well with observed streaks. During these periods when diurnal forcing dominates, the forecasts from the NMM22 differ the most from the NMM5 forecasts, especially during forecast hours 24–48. This difference during forecast hours 24–48 is evident from spatial correlation coefficients of 0.51 (Fig. 5b) and 0.38 (Fig. 6b) in the NMM5 and NMM22, respectively. In the NMM22, during the 24–48-h forecast period, there is very little correspondence between forecast and observed propagating rainfall streaks (Fig. 6).

Generally, during many time periods the time series from both configurations appear very similar. Many apparently propagating rain streaks are observed in the NMM22 and the NMM5 forecasts. This is not surprising: past studies (e.g., Bukovsky et al. 2006) have shown that models using CPSs can appear to simulate propagating MCSs. However, the propagating signal in the diurnally averaged plots is much more clearly seen in the NMM5 forecasts. Davis et al. (2003) noted similar findings in 22-km grid-spacing WRF model simulations.
using a CPS and concluded that the rain streaks observed in the forecasts were not phase locked to the diurnal cycle like they are in the real atmosphere. In other words, because of timing and placement errors, when the averaging was performed the propagating signal was drowned out. Results in this study indicate that the rainfall forecasts from the NMM5 appear to be phase locked to the diurnal cycle because the long-term averaging does not drown out the propagating signal, but this does not appear to be the case in the NMM22 forecasts.

c. Comparison of Hovmöller diagrams for the north and south portions of the domain

Because observations show that different rainfall regimes occurred within the domain in this study, the domain was split into northern and southern halves to identify the features in the forecasts and observations associated with the different regions (Figs. 7 and 8, for the northern and southern halves, respectively). The latitude 38°N was chosen to separate the domain into two halves because the corridor in which propagating rainfall systems occurred most frequently was roughly the northern half of the full domain. The tendency for propagating systems to concentrate in latitudinal corridors has been documented by Tuttle and Davis (2006). This partitioning led to diurnally averaged Hovmöller diagrams of observed rainfall with the propagating (nonpropagating) rainfall axis having the higher amplitude in the northern (southern) portion of the analysis domain. Also, from observations in the northern region (Fig. 7c), two axes of propagation can be identified; one emanating at about 104°W and another weaker axis emanating around 96°W, approximately the same longitude where the strongest part of the western propagation axis ends. This second, weaker axis is not observed in the southern part of the domain.

**Fig. 6.** Same as in Fig. 5 but for the NMM22.
The NMM5 forecasts correctly differentiate the two different precipitation regimes in both the northern and southern regions, with the only exception occurring during forecast hours 0–24 in the northern region when spinup issues are likely affecting the intensity of the propagating axis. As in observations (Figs. 7c and 8c), the NMM5 forecasts show that the propagating (non-propagating) signal in the western (eastern) part of the domain has its highest amplitude in the northern (southern) region (Figs. 7b and 8b), with the only exception mentioned above.

The NMM22 correctly differentiates the different precipitation regimes only in the southern region where the NMM22 appears to simulate a coherent propagating axis of rainfall in the western portion of domain that corresponds reasonably well with observations (albeit, with a speed that is too slow), especially during forecast hours 0–24, and a stronger nonpropagating axis in the eastern portion of the domain (Fig. 8a). In the northern region there do not appear to be any axes of propagating rainfall in the western portion of the domain, although it appears the NMM22 may be able to simulate the weaker secondary propagating rainfall axis noted in the eastern portion of the domain. The amplitudes of the NMM22 rainfall axes in both the western and eastern regions are comparable (Fig. 7a). Thus, the NMM22 forecasts in the northern region do not correspond to the observations showing the propagating axis of rainfall in the western portion of the domain having a higher amplitude than the nonpropagating axis in the eastern portion of the domain (Fig. 7c). The largest rainfall amounts in the northern region of the NMM22 occur in a nonpropagating axis during forecast hours 24–30 centered around 94°W longitude. This nonpropagating axis corresponds well with a nonpropagating axis in the observations occurring between the two propagating axes marked in Fig. 7c.

In the NMM22, a comparison of spatial correlation coefficients between the northern and southern regions agrees with what was concluded from the subjective comparison. During forecast hours 0–24 (24–48) these values were 0.46 and 0.81 (0.66 and 0.79) for the northern and southern regions, respectively, which agrees with the subjective comparison showing that the southern region has a better diurnal cycle representation than the northern region in the NMM22. Comparing spatial correlation coefficients for the western and eastern halves of the NMM22 forecasts (Table 1) reveals...
that the NMM22 had the most difficulty in the western half of the domain in both northern and southern regions, evident by the lower spatial correlation coefficients in the west than in the east. This difficulty in the western half of the domain is not unexpected because of previous studies discussed earlier indicating that models using CPSs have trouble simulating propagating convection, which is most frequent in the western half of the domain.

In the NMM5, for both northern and southern regions, values of spatial correlation coefficients were lower (higher) in the western half than the eastern half of the domain during forecast hours 0–24 (24–48; Table 1). The most dramatic difference occurs in the northern region where spatial correlation coefficients improve from 0.14 to 0.87 from the 0–24- to the 24–48-h forecast period in the western region (Table 1). This reversal reflects the improvement in the timing and placement of the propagating rainfall axis in the western half of the domain from forecast hours 0–24 to 24–48.

d. Traditional skill measures

1) Bias

The biases from the NMM5 simulations were less than 1.0 initially and slowly increased to above 1.0 by the 6–9-h period (Fig. 3e), reflecting the model spinup of microphysical variables from zero. Skamarock (2004) also found, using 4- and 10-km grid-spacing convection-resolving configurations of the WRF model, that the mesoscale portion of the kinetic energy spectrum reaches a fully developed state between 6 and 12 h into the forecast. For the 0.10-in. rainfall threshold,
the bias scores approached or exceeded 1.0 at forecast hour 9, while for the 0.25-in. thresholds and above this occurred at forecast hour 6 (Fig. 3e). Apparently, it takes the model slightly longer to generate areas of light rainfall that are comparable in scale to the observed areas of light rainfall than it does to create the smaller areas of heavy rainfall.

In the NMM22 simulations, at the 0.10- to 0.50-in. thresholds (Fig. 3e), the biases start relatively high and drop off to lower values at forecast hour 6 before increasing again at forecast hour 9. This behavior is likely evidence of a spurious gap between the convective precipitation generated by the cumulus parameterization scheme and nonconvective precipitation that is being resolved on the grid scale. A phenomenon that has been observed in models using CPSs (Molinari and Dudek 1992). Although not available for all cases, the contributions of convective and grid-resolved components of the total rainfall for a subset of five cases support the suggestion that this spurious gap occurred. In all of these cases the domain-averaged convective precipitation amounts dropped from higher values at forecast hour 3 to lower values at forecast hour 6 while grid-resolved precipitation amounts increased from forecast hours 3 to 9. This trend occurs because CPS formulations allow the schemes to quickly activate and begin generating precipitation in moist, unstable environments favorable for convection. So, while the CPS can quickly generate areas of rainfall, the grid-resolved component still needs time to spin up microphysics variables. At thresholds above 0.50 in. (not shown) there is less evidence of this spurious gap, and the behavior of the biases are similar to the NMM5, except the rate of increase is much lower in the NMM22 than in the NMM5, implying that the CPS has trouble generating heavy precipitation during forecast hours 3–9.

A distinct diurnal oscillation exists in the biases from both configurations (Fig. 3e). In the NMM5 these oscillations have a higher amplitude at higher rainfall thresholds, while the amplitude of the oscillations in the NMM22 forecasts is fairly constant. The NMM5 biases generally peak around forecast hours 24 and 48 although the peaks occur in the range of forecast hours 21–27 and 45–48 (1500–2100 LST). The NMM22 biases generally peak around forecast hours 15–18 and 39–42 (0900–1200 LST), 6 h earlier than the peak observed in the NMM5. The reasoning for these relative maxima in each configuration is discussed in the next section.

Mean biases over the entire forecast period (not shown) reveal that the NMM5 tends to overestimate precipitation more at higher rainfall thresholds. As discussed earlier, the overprediction in the NMM5 may be related to the 5-km grid not being able to resolve subcloud-scale processes that limit instability.

2) ETS

Although ETSs are so low as to imply no skill at any time for heavier rainfall thresholds, it is insightful to explore if similar oscillatory behavior as that observed in the biases is present in the ETSs. Similar to biases, the ETSs do exhibit oscillatory behavior, and the effects of spinup can be observed during the first few forecast hours (Fig. 3d). In the NMM5 forecasts the ETSs attain their highest values at forecast hours 6–9 with a trend for the heavier rainfall thresholds to have their highest scores earlier in the forecast. A definitive exploration of the cause of these trends in both the biases and the ETSs is beyond the scope of this study. However, future work should see if this behavior is caused by differences in dynamical forcing. Perhaps when weak forcing is present the model may generate areas of light rainfall that take a relatively long time to spin up, while under strong forcing the model quickly generates areas of heavy rainfall that are more likely to be associated with the stronger forcing. It is also possible that the model fails to generate stratiform rainfall areas within organized MCSs, or that organized MCSs may be simulated well, but other areas of light rain are often completely missed.

In the NMM22 simulations the ETSs at the lightest rainfall threshold shown (0.10 in.) generally attain their highest values at forecast hour 9 (Fig. 3d), similar to those in the NMM5. This implies that the relatively large areas of rainfall likely generated by the CPS during the first 3 h of the forecast (indicated by the higher bias scores for the 0.10- and 0.25-in. rainfall thresholds at these times), are not corresponding well to observed areas. The highest skill is likely not attained until the grid-resolved component has spun up. At the higher rainfall thresholds [0.50 in. and above (not shown)] the biases in the NMM22 are relatively low (less than 0.5) and the ETSs are also very low (less than 0.1) so that little can be inferred regarding the forecast quality at these times.

After initial peaks, the ETSs from both configurations tend to peak again around forecast hours 30–33 (0600–0900 UTC). Past studies (e.g., Hamill 1999) have indicated that high ETSs are often associated with high biases. Because the highest ETSs in the present study do not occur at the times of the highest bias, it is likely the model truly does have more skill at forecasting rainfall at these times compared to other times. Also, at virtually all forecast hours and rainfall thresholds the ETSs are higher in the NMM22. At least during the 24–48-h period, this is contradictory to what was
found from the subjective evaluation and calculations of spatial correlation coefficients in previous sections, which clearly indicated that the NMM5 had a better diurnal cycle representation during forecast hours 24–48. Past studies have shown that there may be a point of diminishing returns when applying traditional gridpoint verification methods to high-resolution forecasts (e.g., Mass et al. 2002; Gallus 2002; Fritsch and Carbone 2004) because coarser-grid models, unlike fine-grid models, do not get penalized by having finescale details that may be displaced slightly, lowering the ETSs (e.g., Baldwin et al. 2001; Davis et al. 2006a).

e. Relation between diurnally averaged plots of rainfall and skill measures

The positioning of the ETS and bias time series alongside the diurnally averaged Hovmöller diagrams of forecast and observed precipitation make it easy to infer information on what features in each model configuration are causing the observed behavior in the skill measures (Fig. 3). In addition, further information can be obtained from an examination of diurnally averaged x–y plots of forecast and observed precipitation for the forecast hours at which relative peaks in the skill measures were observed (Fig. 9).

The NMM5 biases peak near forecast hours 24 and 48, corresponding to times near or shortly after peak heating and when the nonpropagating component of rainfall in the eastern part of the domain is near its maximum amplitude (Fig. 3). This nonpropagating rainfall maximum is due mainly to rainfall in the southeast United States (Figs. 9j,l). The NMM5 appears to accurately depict the timing of this rainfall maximum occurring at forecast hours 24 and 48, especially during the second diurnal cycle simulated by the model (Fig. 3b). Thus, it can be inferred that the high biases are simply a result of the NMM5 overpredicting rainfall at the maximum phases of the nonpropagating component of the diurnal cycle.

The biases for the NMM22 are highest around forecast hours 18 and 42 corresponding to times at which the propagating signal in the west and nonpropagating signal in the east are at minimum amplitudes (Fig. 3). It is inferred from the diurnally averaged Hovmöller diagrams of forecast precipitation from the NMM22 (Fig. 3a) along with the x–y plots of diurnally averaged precipitation at forecast hours 18 and 42 (Figs. 9m,o) that the relatively high biases at these times are the result of the NMM22 simulating the late-afternoon nonpropagating rainfall maximum too early. Thus, the high biases were caused mainly by this phase difference, as opposed to the overprediction observed in the NMM5.

ETTs for both configurations are generally highest at or around forecast hours 6–9 and 30–33 (Fig. 3d), corresponding to the times at which the propagating component in the western part of the domain is at its maximum amplitude (Fig. 3c). Because the NMM5 accurately depicts the timing and longitude of both rainfall maxima (propagating in the west and nonpropagating in the east), it is difficult to ascertain why the ETSs were highest around 0600–0900 UTC in this model configuration, corresponding to the maximum amplitude of the propagating component. It is likely that in the southeast United States the nonpropagating rainfall maximum is associated with convection that is unorganized, short-lived, and chaotic in nature. In the western high plains the propagating rainfall maximum is associated with long-lived and organized MCSs that are inherently more predictable (Carbone et al. 2002). Thus, small errors in the location of rainfall areas will penalize the ETSs more in the areas with more random and chaotic convection.

In the NMM22, the ETS maxima likely occur at hours similar to the NMM5 for the same reason because of the enhanced predictability of the rainfall systems that occur at these times. Also, the skill in the NMM22 forecasts is likely degraded more during the afternoon hours than in the NMM5 forecasts because of the timing errors discussed earlier.

5. Summary and future work

This paper examined the representation of the diurnal cycle by a 5-km grid-spacing configuration of WRF–NMM that did not use a CPS and compared it with a 22-km grid-spacing configuration of WRF–NMM that used a CPS. The historic lack of progress in improving quantitative precipitation forecasts (QPFs) is often blamed on CPSs because CPSs are likely unable to simulate the mesoscale dynamics leading to propagating convection (i.e., the unrealistic treatment of downdrafts and resulting poor representation of cold pools in models with grid spacing above 10 km; Davis et al. 2003). The major findings from this study are summarized below.

Because of model spinup the axes of propagating rainfall from the NMM5 forecasts in both regimes were displaced later in time and east of what was observed so that indirect effects of spinup were probably present up to 24 h into the forecasts. At forecast hours 24–48 the NMM5 corrected the issues associated with spinup and depicted the timing and location of propagating and nonpropagating areas of rainfall extremely well. The NMM22 had problems during both forecast periods simulating a coherent propagating rainfall axis in the western portion of the domain and simulated the non-
propagating rainfall maximum in the eastern portion of the domain too early. Because of these features it was unclear from a subjective evaluation which model configuration had the better diurnal cycle depiction during forecast hours 0–24, although spatial correlation coefficients were slightly higher in the NMM22. However, it was clear that the NMM5 had a superior diurnal cycle representation during forecast hours 24–48.

Fig. 9. Average 3-hourly accumulated rainfall ending at forecast hour 9 for the (a) NMM5 and (b) stage IV observations; forecast hour 30 for the (c) NMM5 and (d) stage IV observations; forecast hour 18 for the (e) NMM22 and (f) stage IV observations; forecast hour 24 for the (g) NMM22 and (h) stage IV observations; forecast hour 27 for the (i) NMM5 and (j) stage IV observations; forecast hour 42 for the (k) NMM5 and (l) stage IV observations; forecast hour 18 for the (m) NMM22 and (n) stage IV observations; and forecast hour 42 for the (o) NMM22 and (p) stage IV observations.
Time series of Hovmöller diagrams covering monthly time periods revealed that the most notable difference between forecast hours 0–24 and 24–48 occurred during time periods in which most of the rainfall appeared to be generated through diurnal forcing mechanisms. Carbone et al. (2002) noted that the dominance of diurnal forcing corresponds to low skill in the dynamical prediction of convective precipitation. Over this period, at forecast hours 0–24, the NMM5 had trouble generating areas of precipitation; however, at forecast hours 24–48 there was major improvement. The NMM5 forecasts appeared to be much better during the 24–48-h period than the NMM22 forecasts.

The NMM5 was generally able to differentiate the two different precipitation regimes observed in the northern and southern portions of the model domain (greater amplitude to propagating signal in western half of northern region, greater amplitude to nonpropagating signal in eastern half of southern region), separated by 38°N, while the NMM22 was only able to differentiate the two regimes observed in the southern region of the domain.

Distinct oscillations are present in the diurnally averaged ETS and bias skill measures. In the NMM5, peaks in bias are caused by the overprediction of rainfall during the times at which the nonpropagating rainfall signal in the eastern part of the domain is at its maximum amplitude. In the NMM22 the peaks in bias are caused by a difference in phase between the times at which the maximum amplitude of the nonpropagating component in the eastern part of the domain is simulated and the times it is observed.

The ETSs from both model configurations peak at about the same times, which correspond to the times at which the propagating component of the diurnal cycle in the western part of the domain is at its maximum amplitude. It is speculated that this occurs because the organized MCSs that are responsible for the propagating signal are inherently more predictable than the more random, chaotic convection responsible for the nonpropagating signal.

Overall, the ETs from the NMM22 were higher at virtually all forecast hours and rainfall thresholds. Because the diurnally averaged Hovmöller plots suggest that the NMM5 had a much better representation of the diurnal cycle, this study is further evidence of how traditional verification measures can be misleading when applied to high-resolution forecasts.

It should be noted that the year that this study was conducted was not a typical year with respect to rainfall. Much of the Midwest, especially portions of Iowa, Illinois, and Missouri, experienced a severe drought. It will be necessary to see if the results from this study are consistent with those obtained using simulations conducted during other years. Future work should also focus on the mechanisms that are leading to what appears to be timing errors associated with the dissipation of MCSs during the first 24 h of the NMM5 forecasts and whether or not “hot” start initialization procedures can improve the location and magnitude of the major propagating rainfall axis during the first diurnal cycle. Also, future studies should investigate what is causing heavier areas of rainfall to spinup faster than lighter areas.

Acknowledgments. This research was funded by Baker Endowment Fund 497-01-78-3803 and NSF Grant ATM-0537043. Data from the NMM22 were obtained from simulations run on the 64 processor computing cluster in the meteorology program at Iowa State University. Data from the NMM5 were obtained from the NCAR Mass Store with the assistance of Wei Wang and Jim Bresch.

REFERENCES


——, 2002: Nonsingular implementation of the Mellor-Yamada level 2.5 scheme in the NCEP Mesos Model. NCEP Office Note 437, 61 pp.


