A Comparison Study of Convective Parameterization Schemes in a Mesoscale Model

WEI WANG* AND NELSON L. SEAMAN

Department of Meteorology, The Pennsylvania State University, University Park, Pennsylvania

(Manuscript received 21 September 1995, in final form 8 July 1996)

ABSTRACT

A comparison study of four cumulus parameterization schemes (CPSs), the Anthes–Kuo, Betts–Miller, Grell, and Kain–Fritsch schemes, is conducted using The Pennsylvania State University–National Center for Atmospheric Research mesoscale model. Performance of these CPSs is examined using six precipitation events over the continental United States for both cold and warm seasons. Grid resolutions of 36 and 12 km are chosen to represent current mesoscale research models and future operational models. The key parameters used to evaluate skill include precipitation, sea level pressure, wind, and temperature predictions. Precipitation is evaluated statistically using conventional skill scores (such as threat and bias scores) for different threshold values based on hourly rainfall observations. Rainfall and other mesoscale features are also evaluated by careful examination of analyzed and simulated fields, which are discussed in the context of timing, evolution, intensity, and structure of the precipitation systems.

It is found that the general 6-h precipitation forecast skill for these schemes is fairly good in predicting four out of six cases examined in this study, even for higher thresholds. The forecast skill is generally higher for cold-season events than for warm-season events. There is an increase in the forecast skill in the 12-km model, and the gain is most obvious in predicting heavier rainfall amounts. The model’s precipitation forecast skill is better in rainfall volume than in either the areal coverage or the peak amount. The scheme with the convective available potential energy–based closure assumption (Kain–Fritsch scheme) appears to perform better. Some systematic behaviors associated with various schemes are also noted wherever possible.

The partition of rainfall into subgrid scale and grid scale is sensitive to the particular parameterization scheme chosen, but relatively insensitive to either the model grid sizes or the convective environments.

The prediction of mesoscale surface features in warm-season cases, such as mesoscale pressure centers, wind-shift lines (gust fronts), and temperature fields, strongly suggests that the CPSs with moist downdrafts are able to predict these surface features more accurately.

1. Introduction

Precipitation is recognized as one of the most difficult parameters to forecast in numerical weather prediction. Despite substantial reductions in forecast errors for winds, temperatures, sea level pressures, and geopotential heights as models improve, progress in precipitation prediction has been slow (Olson et al. 1995). Difficulties exist in at least three areas. First, our understanding of precipitation processes is still quite limited. Second, data deficiencies often limit the accuracy of a model’s initial conditions, which are crucial to the prediction of precipitation systems. The third involves the representation of both resolved and subgrid-scale precipitation processes in a mesoscale model. The latter is known as the cumulus parameterization problem, and its challenge and complexity have been acknowledged for many years (e.g., Frank 1983; Emanuel and Raymond 1993).

A wide variety of cumulus parameterization schemes (CPSs) have been developed and incorporated into three-dimensional mesoscale models (e.g., Kuo 1974; Arakawa and Schubert 1974; Kreitzberg and Perkey 1976; Anthes 1977; Brown 1979; Fritsch and Chappell 1980; Molinari and Corsetti 1985; Betts and Miller 1986; Frank and Cohen 1987; Trembach 1990; Kain and Fritsch 1993; Grell 1993). Successful applications in mesoscale models have been reported in both semiprognostic tests (e.g., Grell et al. 1991) and fully prognostic tests (e.g., Zhang et al. 1989; Dudhia 1989; Grell 1993; Chen and Frank 1993; Kuo et al. 1996). However, most of these studies evaluated one scheme in a specific convective environment or in a limited number of case studies. Because no universal conceptual framework exists for cumulus parameterization at present (Arakawa 1993), the general applicability of any scheme is not obvious when it is applied in an environment other than those tested by the developer. Moreover, the large number of schemes and the wide variety of convective en-
environments have made a systematic intercomparison of parameterization performance an enormous task. This study presents an intercomparison of a few CPSs in a mesoscale numerical prediction model.

Specifically, there are three issues that provide motivation for this study. First, how well do CPSs perform in a mesoscale model under a variety of convective conditions, and do they have consistent performance in different environments? Second, can increased model resolution improve quantitative precipitation forecasting, and, if so, how much is the gain? Third, for schemes designed originally for large-scale models, what are the characteristic problems encountered in mesoscale applications? Do CPSs behave differently at different model resolutions in the range of 10–40 km? For example, Bougeault and Geleyn (1989) compared two Kuo-type CPSs in simulating two precipitation events at grid sizes of 160, 80, and 40 km, and showed that the schemes had clear grid-size dependency. As the grid size decreased, the difference between the schemes became larger.

An important related aspect of numerical modeling is that when the grid elements are fine enough (say, $\Delta x < 40$ km), some mesoscale dynamic features associated with convection can be simulated explicitly, even though the individual convective cells must be parameterized. These features include thunderstorm outflow boundaries, mesohighs and mesolows, rear-inflow jets, and midlevel vortices associated with mesoscale convective systems (MCSs). Numerous studies have demonstrated the ability of three-dimensional mesoscale models, using real data and parameterized convection, to simulate such observed mesoscale phenomena (e.g., Zhang et al. 1989; Cram et al. 1992; Stensrud and Fritsch 1994; Zheng et al. 1995). These studies indicate that parameterized, moist downdrafts are crucial for reproducing many of the observed mesoscale characteristics, as well as correct large-scale temperature fields. Therefore, verification of these mesoscale features is an important element of evaluating CPS performance.

In this study, we focus on three aspects of the cumulus parameterization problem in full-prognostic tests using the PSU–NCAR mesoscale model. First, we analyze the skill of several CPSs for predicting rainfall at mesoscale resolutions presently used in real-data research applications and those anticipated for near-future operational use. Second, we examine the characteristics of rainfall partitioning into subgrid scale and grid scale in the model. Third, we investigate how CPSs perform in terms of reproducing various surface features, including mesoscale dynamical structures and their propagation. The study is performed for a variety of midlatitude convective environments, including both warm- and cold-season cases.

The paper is arranged in the following manner. Section 2 describes the methodology used in the study. Brief descriptions of the precipitation events are provided in section 3. In section 4, the three-dimensional numerical model and the experiment design are described, while the results of the numerical experiments are discussed in section 5. The summary and conclusions are given in section 6.

2. Methodology

The adopted approach uses the PSU–NCAR mesoscale model as a common test framework to host four CPSs. The model is run for six cases with each CPS at grid sizes of 36 and 12 km. The predictions are then examined in terms of timing, evolution, intensity, and structure of the precipitation, and the related mesoscale dynamical and thermodynamic features.

a. Cumulus parameterization schemes

It is impractical to investigate all CPSs in this type of study. Four have been selected on the basis of their widespread use in numerical models and the representativeness of different closure assumptions and scale considerations. Another factor is the intent to include schemes representing a range of complexity and physical detail.

The four CPSs chosen for evaluation are those commonly referred to as the Anthes–Kuo scheme (AK: Anthes 1977; Grell et al. 1994), the Betts–Miller scheme (BM: Betts and Miller 1986), the Grell scheme (GR: Grell 1993; Grell et al. 1994), and the Kain–Fritsch scheme (KF: Kain and Fritsch 1993). All four CPSs have been introduced into the nonhydrostatic version of the PSU–NCAR model, known as MM5 (see section 4). The choice of these four convective parameterizations in no way implies that they are superior to others. Also, it should be noted that variations to a basic scheme tend to develop as it becomes more widely used. Therefore, the names attached to the schemes identify their “parentage” (and their fundamental closure assumptions), but wherever appropriate, we have identified the source of the individual version used in this study.

Brief descriptions of the selected convective schemes and their closure assumptions appear in appendix A.

b. Numerical experiments

It would be inappropriate to attribute all the errors in the precipitation forecasts solely to a CPS in a full-prognostic test. Inaccuracy in other physics parameterizations and uncertainties in the initial conditions can also contribute to model errors. Because of this complexity, it is important that all the experiments use the same mesoscale model and identical physics (other than the choice of the CPS) as a common framework, as well as the same initial conditions for each case. This design, to a certain extent, simplifies interpretation of the results and facilitates forecast intercomparison.

The choice of the grid resolutions reflects the current
Recognize that these CPSs were designed specifically for a 12-km model. In fact only GR and KF were developed and tested for resolutions as small as 20–25 km, with AK and BM originally applied at much coarser resolutions. Applications at 12 km (the meso-β scale) are relatively new, and the reliability of any CPS has yet to be proven. However, since the practical requirements of real-data modeling now include the 10–15 km scale, it is important to objectively assess the performance of a variety of schemes to facilitate future model design.

All CPSs have certain “tunable” parameters that may be adjusted to enhance their performance for a given case. That approach is inappropriate for operational models and is not allowed in this study. Each CPS is used in exactly the same form for all cases and at both resolutions.

c. Verification datasets

The observations used to assess precipitation predictions are the hourly station reports collected by the National Climate Data Center and archived at NCAR. This dataset consists of about 2500 stations with an average separation of about 40 km east of the Rocky Mountains, providing high temporal and spatial resolution compatible with the coarse grid (36 km) and the timescale of convective systems. The station precipitation data are analyzed to the model grids following Haagenson et al. (1992).

d. Evaluation methods

Evaluation of the precipitation predictions focuses on comparison of several key aspects of the simulations: timing, evolution, areal coverage, and intensity of the rainfall. The intent is to identify the performance characteristics of each CPS and to identify any systematic errors. To achieve these goals, it is important to combine both subjective and objective measures of skill.

First, the simulated prognostic fields (temperatures, winds, and sea level pressures) are verified against objective analyses to ensure that the model has adequate predictive skill to reproduce the synoptic-scale environments in which the precipitation events took place. Then, precipitation forecasts are compared to the observations and are evaluated quantitatively using statistical skill scores (such as the threat and bias scores; Anthes 1983) for several threshold values (typically at 0.0254, 0.254, 0.635, 1.27, and 2.54 cm). The scores are interpreted in conjunction with the forecast and analysis maps. Mean errors are also calculated to examine the skill of volumetric predictions. In this paper, we will focus on 6-h accumulated rainfall predictions to assess the model’s ability to forecast the initiation and evolution of mesoscale precipitation systems. Mesoscale dynamical features associated with convective precipitation are also examined and compared to surface observations wherever appropriate.

Note that the evaluation methods outlined here do not examine all aspects of a convective system. For example, it is beyond the scope of this study to examine or validate the net heating profiles associated with convection for each scheme and each case. Such an effort would require very special datasets that are difficult to obtain for a wide variety of environments.

3. Precipitation events

To limit the scope of this work, care was taken to select a number of cases that represent a variety of continental extratropical convective conditions. Three cold-season and three warm-season cases are selected. Table 1 lists the case number, the periods of interest for the 36-km experiments, the type of environment, the duration of the precipitation events, and the maximum 6-h rainfall rates for each case. The large-scale environment for the cold-season events was characterized by approaching upper-air troughs over the Rockies and southward moving surface cold fronts in the Plains. Ahead of the upper-level troughs, broad southerly flow brought abundant moisture from the Gulf of Mexico to the coastal states. In the case of 15–17 February 1993 (case 1), a secondary front was located across central Texas, separating the northward-moving warm and moist air from the Gulf of Mexico and cool air ahead of a polar front. It was along this front that severe convection broke off. As the polar front caught up the southern front around 0600 UTC 16 February, more convection occurred ahead of the front in the warm sector, while stratiform precipitation took place along a nearly west–east-oriented stationary front. In a similar large-scale environment, the initial convection in case 2 occurred just ahead of a north–south-oriented dryline in west-central Texas. Detailed descriptions of synoptic conditions for cases 2 and 3 appear in Locatelli et al. (1995) and Wang et al. (1995). As the surface lows marched eastward through the southern states in these wintertime events, precipitation was heavy during most of the simulation periods. Maximum observed 6-h accumulated rainfall exceeded 2.54 cm for 30 h in case 1, 12 h in case 2, and 36 h in case 3, respectively.

The warm-season events are chosen exclusively for MCS-type systems. This is because these systems con-
tribute significantly to the seasonal rainfall totals (Fritsch et al. 1986) and provide the most challenge for both research and operational models (e.g., Fritsch and Maddox 1981). Detailed descriptions of synoptic conditions can be found in the studies by Brandes (1990) for case 4, Smull and Augustine (1993) and Trier and Parsons (1993) for case 5, and Rutledge et al. (1988) and Johnson and Hamilton (1988) for case 6. Distinct surface features associated with the two squall-line cases (4 and 6) include leading troughs ahead of the surface gust fronts, mesohighs associated with thunderstorms, and trailing wake lows. Surface precipitation was also heavy and long-lasting for these organized warm-season cases, with 6-h accumulated rainfall exceeding 2.54 cm through much of the life cycles of these systems.

### 4. Numerical model and experiment design

The mesoscale modeling system used in this study is the nonhydrostatic PSU–NCAR MM5 described by Dudhia (1993) and Grell et al. (1994). The MM5 is a three-dimensional, limited-area, primitive-equation, nested-grid model with a terrain-following \( \sigma \) (nondimensionalized pressure) vertical coordinate. In addition to the CPSs (section 2a), the important physics of the model include a prognostic surface energy budget equation for ground temperature, the Blackadar medium-resolution planetary boundary layer parameterization (Zhang and Anthes 1982), and an explicit prediction scheme for rainwater and cloud water with simple ice physics (Dudhia 1989). The MM5, like many mesoscale models, uses a hybrid precipitation representation. That is, an explicit moisture scheme (EX) is used to simulate resolved-scale precipitation, while a CPS is used simultaneously to represent subgrid-scale precipitation.

The initial time for each of the 36-km experiments is chosen so that the initiation of the precipitation system of interest takes place 6–12 h into the simulation, and the maximum precipitation occurs 18–30 h into the integration (Table 1). This enables the model to “spin up” critical environmental characteristics, such as convergence, vertical motions, and cloud liquid water, prior to the convective outbreaks. All 36-km experiments are run for 36 h. The 12-km experiments start at 6–12 h into the 36-km simulations in a one-way nesting mode with lateral-boundary conditions provided by the 36-km model at 1-h intervals (Table 2). To minimize any residual spinup problems in the 12-km domains, the vertical velocity, cloud water, and rainwater fields are included in the initial and boundary conditions. Each 12-km experiment is run for 24 h.

### Table 1. Summary of cases.

<table>
<thead>
<tr>
<th>Case</th>
<th>Date/time for 36-km experiments</th>
<th>Case description</th>
<th>Duration of precipitation event</th>
<th>Max. 6-h rainfall (cm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1200 UTC 15 Feb–0000 UTC 17 Feb 1993</td>
<td>Arctic front, with convection south of front</td>
<td>Entire period</td>
<td>8.6</td>
</tr>
<tr>
<td>2</td>
<td>1200 UTC 8 Mar–0000 UTC 10 Mar 1992</td>
<td>Convection ahead of cold front associated with occluding cyclone</td>
<td>1800 UTC 8 Mar–0000 UTC 10 Mar</td>
<td>6.1</td>
</tr>
<tr>
<td>3</td>
<td>1200 UTC 9 Mar–0000 UTC 11 Mar 1992</td>
<td>Continuation of case 2, with new convection along Gulf Coast</td>
<td>Entire period</td>
<td>7.6</td>
</tr>
<tr>
<td>4</td>
<td>1200 UTC 6 May–0000 UTC 8 May 1985</td>
<td>Single MCS</td>
<td>1800 UTC 6 May–0000 UTC 8 May</td>
<td>6.6</td>
</tr>
<tr>
<td>5</td>
<td>0000 UTC 3 Jun–1200 UTC 4 Jun 1985</td>
<td>Four MCSs</td>
<td>1200 UTC 3 Jun–1200 UTC 4 Jun</td>
<td>6.3</td>
</tr>
<tr>
<td>6</td>
<td>1200 UTC 10 Jun–0000 UTC 12 Jun 1985</td>
<td>MCS/squall line</td>
<td>1800 UTC 10 Jun–1800 UTC 11 Jun</td>
<td>5.8</td>
</tr>
</tbody>
</table>

### Table 2. Attributes of the 12-km experiments.

<table>
<thead>
<tr>
<th>Case</th>
<th>Initial time</th>
<th>Domain center</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1800 UTC 15 February 1993</td>
<td>33.6°N, 88.5°W</td>
</tr>
<tr>
<td>2</td>
<td>2100 UTC 8 March 1992</td>
<td>34.7°N, 88.5°W</td>
</tr>
<tr>
<td>3</td>
<td>1800 UTC 9 March 1992</td>
<td>32.2°N, 91.3°W</td>
</tr>
<tr>
<td>4</td>
<td>2100 UTC 6 May 1985</td>
<td>35.4°N, 99.3°W</td>
</tr>
<tr>
<td>5</td>
<td>1200 UTC 3 June 1985</td>
<td>36.4°N, 98.5°W</td>
</tr>
<tr>
<td>6</td>
<td>1800 UTC 10 June 1985</td>
<td>38.0°N, 100.7°W</td>
</tr>
</tbody>
</table>
Initial and lateral boundary conditions for the 36-km runs are prepared using only routine National Weather Service (NWS) observations and following the standard objective analysis procedures of the PSU–NCAR modeling system (Manning and Haagenson 1992). Although special observations are available for most of the cases, none of these nonstandard data are used to enhance the initial or lateral boundary conditions. This requirement more closely simulates an operational setting. It should be noted that initial conditions are crucial for the successful simulation of precipitation events (e.g., Zhang and Fritsch 1986; Kuo and Low-Nam 1990; Fritsch and Heideman 1989); this issue will not be addressed in the present study.

For each case five numerical experiments are performed at the two grid sizes. These five experiments include four in which a hybrid precipitation scheme is used, each with a different CPS (AK, BM, GR, or KF) and EX scheme, plus an additional experiment in which only the EX scheme is used (no CPS). Therefore, there are 10 experiments for each case, or a total of 60 numerical experiments in the study. The EX scheme requires grid-box saturation before precipitation can be initiated. While it is not anticipated that the EX scheme is appropriate to be used alone at these resolutions and for all convective environments, the additional experiment is performed to aid in the interpretation of the results from other experiments using a CPS.

When referring to individual experiments in the set, they will hereafter be identified by the two-character code for the CPS, the grid length, and the case number, so that the 36-km Anthes–Kuo experiment for case 1 is designated as AK36-1. When referring to an averaged response for a group of experiments having the same CPS and the same resolution, the shorter two-character code will be used alone.

5. Results

Verification of the model-simulated wind, temperature, and sea level pressure fields using conventional objective measures (such as root-mean-square error, mean error, and $S_1$ scores), indicates that simulations with various CPSs do not have significant differences, especially in the middle and upper levels (not shown). For the winter cases, where synoptic-scale low pressure centers exist, the storm tracks and intensities differ very little. This implies that, as anticipated, the large-scale features of these cases are not dictated by the parameterized convection. However, mesoscale pressure features in the vicinity of precipitation often show appreciable differences among the experiments for a given case (e.g., Zhang et al. 1989; Kuo et al. 1996), since they are closely associated with the details of a precipitation forecast and the dynamics of the precipitation systems. This is particularly apparent in the case of warm-season MCSs. This response will be examined in greater detail in section 5c.

a. Precipitation prediction skills

Analysis of simulated rainfall from the 36-km experiments reveals that each of the CPSs can predict all the precipitation events with a certain degree of success.
However, predictive skill varies from cold season to warm season and from case to case. As expected from the performance of typical operational models, the precipitation simulations are generally better for cold-season events and for systems in which synoptic-scale dynamical forcing is dominant. The important aspects of this performance are described in the following sections.

As noted in section 2d, the statistical calculations of model skill are performed for 6-h accumulative precipitation. This allows the evaluations to focus on the evolution of mesoscale convective systems and embedded convective structures within synoptic-scale systems. Although such a procedure is clearly appropriate, considering the objectives of the paper, it means that the skill scores may appear lower than those based on precipitation totals summed over longer periods (e.g., 24-h accumulation). It should be noted that while the statistical verifications may provide helpful measures of how a scheme performs, this approach to rainfall verification can obscure details that are easily detected visually. Some attempt is made in sections 5b and 5c to discuss other aspects of the rainfall prediction.

1) PRECIPITATION IN COLD-SEASON EVENTS: 36-km RUNS

The characteristic light-precipitation forecast at 36-km resolution for the three cold-season events (cases 1–3) is revealed in Fig. 2. The figure shows the averaged threat scores (TS) and bias scores (BS) as a function of time computed for the 0.0254-cm threshold on a subdomain of 39 × 39 points centered over the primary precipitation areas (the subdomain coincides with the location of the 12-km domain for each case). Notice that although there is a good deal of embedded convection in these synoptically dominated events, the positions of the precipitation systems measured by the TS (Fig. 2a) are predicted quite well by all the CPSs and EX throughout the integration periods. In fact, when averaged over the complete 36-h period (indicated at the right side of the figure), the range of the TS among the five precipitation schemes is only about 0.03. The TS rises to 0.59 near the time of maximum intensity of the primary precipitation systems (about 24 h in these experiments). Predictive skill is the lowest early in the period when the model is still undergoing “spinup” (TS ~ 0.37), and the observed precipitation systems were developing. The skill is also lower late in the period when the storms were decaying, and the area of observed precipitation began to decrease rapidly (TS ~ 0.43).

Areal precipitation prediction skill is examined using the BS. Figure 2b shows the average cold-season BSs computed for the 0.0254-cm threshold. The figure indicates that areal predictions also have fairly good skill, with the time-averaged BSs (at the right) ranging from 1.18 for EX runs (best) to 1.72 for AK runs (worst). The scatter among the five types of experiments is large initially, but becomes smaller as the storm systems mature. The low BS in the EX runs at 6-h forecast (0.58) reflects the fact that when a resolvable scheme is used alone at 36-km grid, the model can severely underpredict precipitation during the early period of a forecast. The runs with subgrid-scale CPSs, on the other hand, can produce precipitation in unsaturated conditions and hence have smaller low-bias error.

For cold-season cases, Fig. 2 shows that the AK scheme has a marginally greater skill in terms of the TS, but it tends to overpredict the areal coverage (large BS). Conversely, when the explicit moisture prediction is used alone (EX), it performs worse than the CPS runs for the TS, but better than the CPS runs for the BS. The generally good skill found with the EX experiments is not surprising in cold-season cases. Under the influence of strong large-scale forcing in a relatively stable atmosphere and with ample moisture, grid-scale vertical motion is often sufficient to lift the moist air and produce precipitation. However, it will be shown later in section...
TABLE 3a. Six 6-h period averages over three cold-season cases of threat scores for various precipitation thresholds.

<table>
<thead>
<tr>
<th>Threshold (cm)</th>
<th>KF36</th>
<th>GR36</th>
<th>BM36</th>
<th>AK36</th>
<th>EX36</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0254</td>
<td>0.457</td>
<td>0.463</td>
<td>0.437</td>
<td>0.466</td>
<td>0.443</td>
</tr>
<tr>
<td>0.254</td>
<td>0.395</td>
<td>0.390</td>
<td>0.380</td>
<td>0.390</td>
<td>0.348</td>
</tr>
<tr>
<td>0.635</td>
<td>0.285</td>
<td>0.287</td>
<td>0.273</td>
<td>0.315</td>
<td>0.241</td>
</tr>
<tr>
<td>1.27</td>
<td>0.187</td>
<td>0.147</td>
<td>0.141</td>
<td>0.168</td>
<td>0.142</td>
</tr>
<tr>
<td>2.54</td>
<td>0.057</td>
<td>0.053</td>
<td>0.048</td>
<td>0.028</td>
<td>0.068</td>
</tr>
</tbody>
</table>

TABLE 3b. Six 6-h period averages over three cold-season cases of bias scores for various precipitation thresholds.

<table>
<thead>
<tr>
<th>Threshold (cm)</th>
<th>KF36</th>
<th>GR36</th>
<th>BM36</th>
<th>AK36</th>
<th>EX36</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0254</td>
<td>1.449</td>
<td>1.587</td>
<td>1.345</td>
<td>1.719</td>
<td>1.182</td>
</tr>
<tr>
<td>0.254</td>
<td>1.249</td>
<td>1.008</td>
<td>1.072</td>
<td>1.169</td>
<td>0.866</td>
</tr>
<tr>
<td>0.637</td>
<td>1.020</td>
<td>0.947</td>
<td>1.112</td>
<td>1.087</td>
<td>0.934</td>
</tr>
<tr>
<td>1.27</td>
<td>0.831</td>
<td>0.887</td>
<td>0.890</td>
<td>0.810</td>
<td>1.114</td>
</tr>
<tr>
<td>2.54</td>
<td>0.240</td>
<td>0.447</td>
<td>0.189</td>
<td>0.201</td>
<td>1.051</td>
</tr>
</tbody>
</table>

5a(5) that the spinup problem associated with the EX runs during the early simulation periods can cause overprediction later in the forecasts.

To summarize precipitation forecast skill at all thresholds for the cold-season cases, Table 3 lists TSs and BSs averaged for the entire 36 h for each scheme. The TSs show that the precipitation forecast at higher thresholds still possesses some skill from the CPS runs. The BSs show that the model rainfall forecast bias is small for moderate rainfall thresholds of 0.254 and 0.635 cm (closer to 1), better than the BSs for the light precipitation (0.0254-cm threshold). Based on the verifications represented by the TS and BS, none of the schemes evaluated here show notably superior skill for these cold-season events.

2) PRECIPITATION IN WARM-SEASON EVENTS: 36-KM RUNS

In the three cases of strong summertime convection (cases 4–6), rainfall prediction varies more noticeably from case to case and among the CPS runs. Apparent differences exist in the predictions for the timing and locations of initial convection. In the case of squall line simulations (cases 4 and 6), the prediction of the line positions strongly affects the precipitation predictive skills. The TS is generally low in the summer events even for a low threshold (0.02–0.34 at the 0.0254-cm threshold; Fig. 3a). Figure 3a also shows that there is very large scatter in TSs at hour 6. Even though the precipitation events of interest did not start until 6–12 h into the simulations, the low and scattered TSs during the first 6-h period reflect the model’s inability to handle precipitation forecast during this spinup period. It also suggests there is a greater sensitivity among the CPSs for initiating rainfall during this period. The BM and KF schemes appear to handle this situation better than others. The averaged TSs during the final 30 h for these CPSs show similar statistical skills. However, this is somewhat misleading. A visual inspection of the predicted and observed rainfall fields (not shown) indicates that forecasts with some schemes (most often KF and GR) actually are considerably better than others in terms of timing, location, and structure of the rainfall prediction (see section 5c). The TSs for these forecasts at low precipitation threshold (0.0254 cm), however, are penalized by overprediction of the areal rainfall, which also degrades the BSs (see Fig. 3b). This apparent contradiction demonstrates that a more thorough examination of the forecasts is necessary to form an accurate assessment of the predictive skill of the CPSs.

Comparing Figs. 2b and 3b, it is evident that the characteristics of the BSs also have large seasonal differences. In the warm-season cases BSs tend to be larger (greater than 1), and there is a greater scatter among the experiments. The rather large BSs for the GR and KF
and over 80% was related to thunderstorms. The scattered nature of thunderstorms suggests that there may be a large variation in the convective potential of the environmental conditions across a region during warm seasons. Under such situations, mesoscale forcing and moisture distribution can become very important for initiating convection at the right time and location (Zhang and Fritsch 1986; Bresch 1994). Because these finescale features are not resolved well by the conventional observation network, models often have difficulty correctly initiating a precipitation event when initialized from synoptic-scale data. When large-scale dynamic forcing is weak and the convective initiation is in serious error, it is not clear whether the parameterized physics can correct the subsequent evolution and structure of the convective system.

3) 12-KM VERSUS 36-KM RUNS

Analysis of precipitation forecasts from the 12-km experiments reveals an overall improvement in the skill scores, compared to the 36-km runs, especially at moderate and heavy rain thresholds. The TSs and BSs for five threshold values averaged for the three 6-h periods (which overlap with the 6-h periods of the 36-km experiments) are provided in Tables 5 and 6. When compared with the scores in the corresponding three 6-h periods for the 36-km runs (not shown), the TSs for the 12-km model are consistently higher (shown in italic). There is also some improvement in the BSs at higher threshold values for the cold-season cases. For the warm-season events, only the averaged BSs for KF and GR runs at lower thresholds improve.

To illustrate this improvement more clearly, the TS and BS at the 1.27-cm threshold averaged over the three

<table>
<thead>
<tr>
<th>Table 4a. Six 6-h period averages over three warm-season cases of threat scores for various precipitation thresholds.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Threshold (cm)</td>
</tr>
<tr>
<td>0.0254</td>
</tr>
<tr>
<td>0.254</td>
</tr>
<tr>
<td>0.635</td>
</tr>
<tr>
<td>1.27</td>
</tr>
<tr>
<td>2.54</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 4b. Six 6-h period averages over three warm-season cases of bias scores for various precipitation thresholds.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Threshold (cm)</td>
</tr>
<tr>
<td>0.0254</td>
</tr>
<tr>
<td>0.254</td>
</tr>
<tr>
<td>0.635</td>
</tr>
<tr>
<td>1.27</td>
</tr>
<tr>
<td>2.54</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 5a. Six 6-h period averages over three cold-season cases of threat scores for various precipitation thresholds for 12-km model runs. The scores in italic are better than those in the 36-km runs.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Threshold (cm)</td>
</tr>
<tr>
<td>0.0254</td>
</tr>
<tr>
<td>0.254</td>
</tr>
<tr>
<td>0.635</td>
</tr>
<tr>
<td>1.27</td>
</tr>
<tr>
<td>2.54</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 5b. Three six 6-h period averages over three cold-season cases of bias scores for various precipitation thresholds for 12-km model runs. The scores in italic are better than those in the 36-km runs.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Threshold (cm)</td>
</tr>
<tr>
<td>0.0254</td>
</tr>
<tr>
<td>0.254</td>
</tr>
<tr>
<td>0.635</td>
</tr>
<tr>
<td>1.27</td>
</tr>
<tr>
<td>2.54</td>
</tr>
</tbody>
</table>

experiments at this threshold indicate a tendency to overpredict the area of light precipitation. This may be explained in part by the closure assumptions used by the two schemes and the design of their trigger functions (Kain and Fritsch 1992). The KF scheme determines whether convection can occur based on existence of convective available potential energy (CAPE), while the GR scheme does so by computing buoyant energy generated by grid-scale fields. Both schemes are therefore sensitive to thermal forcing at the surface, which is strong during summertime. At this light-rain threshold, EX and BM have the best overall areal prediction skill (BS closer to 1). However, for the EX experiments, the spinup effect leads to very low BSs that persist through the first 12 h of a forecast. Therefore, for the light-rain threshold in warm-season events, the TS and BS indicate that the BM scheme has the greatest skill among the schemes evaluated in this study.

The averaged TSs and BSs over the six 6-h periods for the three warm-season events at all thresholds are provided in Table 4. Similar to the scores at the lowest threshold value, the TSs and BSs at higher thresholds also show less skill than their counterparts for the cold-season events. In particular, the BSs for EX and BM at thresholds 1.27 and 2.54 cm show that these schemes tend to overpredict the area of heavy rainfall.

The results shown here agree with many previous studies in that the model's precipitation forecast skill is generally lower in the summertime (e.g., Olson et al. 1995), when large-scale dynamic forcing is relatively weak and thermal forcing at the surface is relatively strong. The difficulty that models have in accurately predicting warm-season precipitation is well documented by Fritsch and Heideman (1989). In their study, about half of the summertime precipitation was associated with organized mesoscale systems and processes, and over 80% was related to thunderstorms. The scattered nature of thunderstorms suggests that there may be a large variation in the convective potential of the environmental conditions across a region during warm seasons. Under such situations, mesoscale forcing and moisture distribution can become very important for initiating convection at the right time and location (Zhang and Fritsch 1986; Bresch 1994). Because these finescale features are not resolved well by the conventional observation network, models often have difficulty correctly initiating a precipitation event when initialized from synoptic-scale data. When large-scale dynamic forcing is weak and the convective initiation is in serious error, it is not clear whether the parameterized physics can correct the subsequent evolution and structure of the convective system.

The results shown here agree with many previous studies in that the model's precipitation forecast skill is generally lower in the summertime (e.g., Olson et al. 1995), when large-scale dynamic forcing is relatively weak and thermal forcing at the surface is relatively strong. The difficulty that models have in accurately predicting warm-season precipitation is well documented by Fritsch and Heideman (1989). In their study, about half of the summertime precipitation was associated with organized mesoscale systems and processes, and over 80% was related to thunderstorms. The scattered nature of thunderstorms suggests that there may be a large variation in the convective potential of the environmental conditions across a region during warm seasons. Under such situations, mesoscale forcing and moisture distribution can become very important for initiating convection at the right time and location (Zhang and Fritsch 1986; Bresch 1994). Because these finescale features are not resolved well by the conventional observation network, models often have difficulty correctly initiating a precipitation event when initialized from synoptic-scale data. When large-scale dynamic forcing is weak and the convective initiation is in serious error, it is not clear whether the parameterized physics can correct the subsequent evolution and structure of the convective system.

3) 12-KM VERSUS 36-KM RUNS

Analysis of precipitation forecasts from the 12-km experiments reveals an overall improvement in the skill scores, compared to the 36-km runs, especially at moderate and heavy rain thresholds. The TSs and BSs for five threshold values averaged for the three 6-h periods (which overlap with the 6-h periods of the 36-km experiments) are provided in Tables 5 and 6. When compared with the scores in the corresponding three 6-h periods for the 36-km runs (not shown), the TSs for the 12-km model are consistently higher (shown in italic). There is also some improvement in the BSs at higher threshold values for the cold-season cases. For the warm-season events, only the averaged BSs for KF and GR runs at lower thresholds improve.

To illustrate this improvement more clearly, the TS and BS at the 1.27-cm threshold averaged over the three
cold-season events are shown in Fig. 4. The bars on the left for a corresponding time are for the 36-km runs, and the bars on the right are for 12-km experiments. The hours marked on the abscissa correspond to the end of the 6-h accumulation periods in the 36-km runs. The increase in TSs for the 12-km model is apparent for all the three periods and all schemes. The average TS rises from 0.13 for the 36-km runs to 0.21 for the 12-km runs. The change due to grid resolution in the BS is generally small, with the exception of the EX runs, which have relatively large increase (degradation) in BS as the grid size decreases.

As an example of the warm-season events, Fig. 5 shows a similar improvement in case 6 for the TS and BS at the 0.0254-cm (light rain) threshold due to increased grid resolution. In this squall-line simulation, the improvement is most evident in the TS during the second period ending at 18-h of the 36-km run time, when experiment KF12-6 produces a TS = 0.56, which is high for a 6-h period rainfall skill forecast in a summer event. During this period, the observed squall line propagated from the eastern end of the Oklahoma panhandle to central Oklahoma (see Figs. 12 and 16 of Zhang et al. 1989). An increase in the TS at the 12-km resolution occurs for all five of the schemes because the squall-line propagation rate is predicted more accurately, especially in the Texas panhandle region. Figure 6 demonstrates this improvement in the KF and BM experiments by comparing the 6-h rainfall predictions at 36- and 12-km grid sizes. Similar changes occur for the other schemes as well. A possible explanation for the improved skill in precipitation forecast is that as the grid size decreases, convergence and vertical motion tend to be better resolved. This, in turn, helps improve precipitation forecast.

Note that the various CPSs overpredict the area of light rain in case 6 at both grid sizes (most BSs are greater than 1.0 in Fig. 5b), and that the GR and AK schemes seem to have the most difficulty in this respect. For heavier rainfall thresholds (≥ 1.27 cm), the EX and BM schemes generally produce greater overprediction of precipitation area in the warm season cases (Table 6b). The most consistent change in warm-season BSs due to resolution is produced by the EX runs. Over a range of lower thresholds (≤ 0.635 cm), in almost all instances the characteristic of the areal prediction in EX experiments changes from underprediction at 36-km to overprediction at 12-km resolution.

The results imply that an increase in model resolution alone cannot overcome all of the difficulty that the model experienced at a coarser grid size. Other factors, such as the details in a CPS (e.g., a trigger function) and the model’s initial conditions, may have significant impact on whether the precipitation forecast will be improved or not.

4) Case-to-case variations

Because a variety of convective environments is chosen, it is not surprising to find variations in the performance statistics among the cases. In general, the CPSs perform better in cold-season events and in cases in which synoptic forcing is prominent. To illustrate this point, Fig. 7 shows the TSs and BSs at the 0.0254-cm threshold for the 6-h period ending at the 24-h forecast in the 36-km model. This is the period when the individual precipitation events are near the mature stage and the model’s precipitation forecast skill tends to be the best. Among the six cases, two cold-season cases, 1 and 3, are simulated best (highest TSs), while two warm-season cases 4, and 6, are also simulated quite well. This is confirmed by examining precipitation maps subjectively (not shown). The skill scores are the worst for case 5. A detailed numerical study by Bresch (1994) showed that this case is particularly sensitive to the initial moisture distribution, and the KF scheme had difficulty triggering the high-cloud-base convection. Apparently, similar difficulty is experienced by other schemes. Figure 7 also indicates that the CPSs performance in the cold-season events is more uniform than in the warm-season cases. Comparison among the 12-km model’s warm-season runs reveals more uniform performance among the CPSs than in the 36-km model (not shown).

5) Volumetric predictions

The prediction of total precipitation volume in a given event is also an important indicator of a successful forecast. This is particularly so for soil hydrology, agricultural interests and flood forecasts. It is also an important measure for the total amount of latent heating the model receives from condensation. A useful measure of the volumetric prediction is the mean error (ME; see appendix B for definition). Figure 8 shows the six-case.

<table>
<thead>
<tr>
<th>Threshold (cm)</th>
<th>KF12</th>
<th>GR12</th>
<th>BM12</th>
<th>AK12</th>
<th>EX12</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0254</td>
<td>0.288</td>
<td>0.314</td>
<td>0.276</td>
<td>0.338</td>
<td>0.314</td>
</tr>
<tr>
<td>0.254</td>
<td>0.195</td>
<td>0.183</td>
<td>0.196</td>
<td>0.223</td>
<td>0.198</td>
</tr>
<tr>
<td>0.635</td>
<td>0.090</td>
<td>0.094</td>
<td>0.114</td>
<td>0.116</td>
<td>0.109</td>
</tr>
<tr>
<td>1.27</td>
<td>0.032</td>
<td>0.041</td>
<td>0.053</td>
<td>0.027</td>
<td>0.056</td>
</tr>
<tr>
<td>2.54</td>
<td>0.003</td>
<td>0.002</td>
<td>0.001</td>
<td>0.002</td>
<td>0.005</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Threshold (cm)</th>
<th>KF12</th>
<th>GR12</th>
<th>BM12</th>
<th>AK12</th>
<th>EX12</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0254</td>
<td>1.662</td>
<td>1.950</td>
<td>1.382</td>
<td>2.057</td>
<td>1.325</td>
</tr>
<tr>
<td>0.254</td>
<td>1.810</td>
<td>1.414</td>
<td>1.573</td>
<td>1.435</td>
<td>1.335</td>
</tr>
<tr>
<td>0.635</td>
<td>1.884</td>
<td>1.743</td>
<td>2.250</td>
<td>1.620</td>
<td>1.919</td>
</tr>
<tr>
<td>1.27</td>
<td>1.313</td>
<td>1.837</td>
<td>2.426</td>
<td>1.710</td>
<td>2.483</td>
</tr>
<tr>
<td>2.54</td>
<td>1.280</td>
<td>2.317</td>
<td>4.215</td>
<td>2.094</td>
<td>3.875</td>
</tr>
</tbody>
</table>
averaged MEs for the 6-h rainfall predictions over the areas where rain rates exceed 0.0254 cm (0.01 in.).

The volumetric precipitation evaluations indicate that all the CPSs are reasonably accurate for this key forecast variable. These schemes start with negative mean errors for the first 6-h period when the model is undergoing spinup and the observed precipitation is light and more scattered. Consistent with results shown above, the EX runs exhibit the most serious underforecast at this time. Subsequently, all schemes are able to simulate the observed volume of rain quite well, although the mean errors for EX runs become somewhat greater (over-
forecast of rain) after 18 h than those of other experiments that use a CPS. Apparently, the resolvable-scale rain mechanism of EX is rather inefficient for removing convective instability in a 36-km resolution model. It often delays precipitation, but later overpredicts heavier rainfall amounts over a greater area than both in the observation and in the runs with a CPS. The relatively large negative ME for AK runs at hour 24 indicates that this scheme tends to underestimate total rainfall when the observed rain rates are the highest. Recalling that the model with these CPSs tends to overpredict areas of light rain (BSs > 1.0), the overall reasonable skill as measured by MEs suggests that the overprediction for light rain is compensated by underprediction of heavy amounts. This also implies that the total latent heating from condensation in the model is not excessive.
In contrast to the improvements found in TS and BS as the grid resolution is reduced to 12 km, the mean error becomes slightly greater at the finer resolution (not shown). A possible explanation to this apparent discrepancy is that the observed data density, which is about 40 km, is too coarse to allow accurate verification of the 12-km model forecasts. For example, when a comparison was made between precipitation characteristics from MCSs using hourly rain gauge data alone (McAnelly and Cotton 1986) and a combined dataset that included both 24-h precipitation data from the National Centers for Environmental Prediction’s (NCEP) Heavy Precipitation Branch and hourly data (Kane et al. 1987), it was noted that the latter study indicated much higher total system precipitation (J. M. Fritsch 1994, personal communication). This suggests that the hourly rain data, used for the 6-h rainfall verification in the present study, may not be fully adequate for evaluating the prediction of the total precipitation amounts, especially for models that have higher resolution than the observation. This point remains to be clarified with a better verification dataset than is currently available.
6) PREDICTION OF MAXIMUM RAINFALL

The ability of the CPSs to predict the position and intensity of extreme precipitation amounts can be crucial for flash-flood forecasts. In the experiments that include a CPS, comparisons of the model’s convective and resolvable-scale precipitation indicate that much of the heavy precipitation is produced by the explicit scheme rather than the CPS itself. However, the CPS is a vital factor that often has a great impact on the accuracy of the model-simulated maxima and whether spurious maxima are generated.

No single skill score adequately represents the ability of the precipitation schemes to predict maxima accurately. An obvious useful measure is the mean error in the 6-h precipitation. However, the mean absolute error (MAE; see appendix B for definition) is also important because it discloses information about the typical departure from the mean. Furthermore, the placement of the maxima is crucial, so a mean position error (MPE) is calculated. Also, we calculate the percentage of observed maxima that are simulated by the model (PMS), and the number of spurious maxima (NSM) that are generated but not observed. For this study, observed maxima are defined to be distinct when they are separated by at least 200 km and have 6-h totals of at least 1.5 cm. These distinctions are somewhat arbitrary. Furthermore, limitations in the observation density mean that some maxima may be measured inaccurately or even missed entirely. However, the impact of these limitations is minimized for a large sample size. For cases 1–6 a total of 89 observed maxima were found, with 6-h precipitation averaging over 3 cm of the set.

If the volumetric prediction of a precipitation scheme is reliable, then the mean error of the maxima may be the least important of the five skill scores used here. In practice, adjustments for systematic mean errors in the forecasted maxima can be made quite easily during model postprocessing, based on model performance statistics. In fact, synoptic-scale and mesoscale models are expected to underestimate the heaviest local rainfall because their grid resolutions naturally limit the predictability of finer-scale convective systems and their associated extreme precipitation amounts. However, reliability in the detection of the maxima (related to the PMS and NSM) and their correct placement (the MPE) are crucial because they may be more difficult to compensate for using statistical methods. Therefore, the MPE, PMS, and NSM are important scores for model evaluation.

Table 7 summarizes the performance of the five schemes in the cold and warm seasons for the 36-km model results and for the five skill scores related to maximum precipitation predictions. The 12-km results are generally similar, but the maxima tend to be higher due to the increased resolution (not shown). At first glance the distribution of the asterisks, indicating greatest skill for an individual type of score, reveals no obvious pattern. Notice that the MEs indicate that, on av-
average, the warm-season maxima have been overpredicted by the EX scheme. This is consistent with earlier results showing that this scheme tends to overpredict the volume of rainfall after the spinup period (see Fig. 8). While all the other ME scores indicate that the maxima have been underpredicted (i.e., negative MEs), it is the KF scheme with the most complex physics that has the greatest underprediction in the warm-season cases. As mentioned above, however, the mean error is not an especially critical measure for evaluating the accuracy of the maximum predictions, since this scheme does simulate the total precipitation volume quite well.

For the other four scores concerning rain-maxima predictions (MAE, MPE, PMS, and NSM), the KF scheme performs quite well and does not have large errors in most of the measures. This demonstrates general consistency and reliability. In contrast, the large MAE for the AK scheme suggests that individual maximum predictions were seriously in error more often than for other schemes. The AK scheme also had the greatest (worst) MPE and tended to produce more spurious maxima for the warm-season events. By most measures, GR and BM schemes performed quite well, but GR tended to have fairly large position errors (MPE), while BM had a large error in NSM in the summer cases. In summary, the evaluation of maximum precipitation statistics shows that no scheme has greatly superior performance, but that the KF scheme may have a modest advantage if one considers a variety of measures of skill.

b. Partition of parameterized and grid-scale precipitation

As is typical in the evaluation of operational modeling systems, the verifications presented thus far have focused on total precipitation without regard to whether the rainfall is produced through resolved-scale processes or through parameterized convection. In a numerical model that uses both subgrid-scale and grid-scale precipitation parameterization schemes, a reasonable partition between parameterized and grid-scale rainfall is important to produce realistic numerical simulations of precipitation events and to prevent numerical moist static instability from occurring. This is because the sub-grid-scale convective parameterization represents a more rapid upward transport of heat and moisture, and hence can stabilize a column faster. In some precipitation systems, such as MCSs, the characteristics of the precipitation can change from being intensely convective during the early developmental stages to more stratiform as the systems mature (e.g., Maddox 1980; Smull and Houze 1985; Rutledge et al. 1988; Bartels and Maddox 1991). Numerical modeling studies of midlatitude MCSs indicate that correct simulation of the transformation from a predominantly convective system to a resolvable mesoscale system is crucial for reproducing many of the observed features and dynamics of a MCS (Dudhia 1989; Zhang et al. 1989; Zhang and Gao 1989; Kain 1994).

In view of the importance of a reasonable partition between convective and grid-scale precipitation, the evolution of the partitioning in the model is examined for case 4 (a MCS case) and case 1. Figure 9 presents a time series of observed hourly rain rates, together with the simulated precipitation rates from the 12-km experiments for case 4 (the 36-km counterparts are omitted because they are very similar except for somewhat smaller amounts). The rainfall rates are computed for the areas where hourly rain rates exceed 0.0254 cm. The evolution of the observed precipitation is typical for an MCS (Maddox 1980; McAnelly and Cotton 1986; Kane et al. 1987). The system began with intense convective activity shortly after 0000 UTC 7 May, which is reflected in the large rainfall rates between 0000 and 0400 UTC 7 May in Fig. 9. This activity was followed by a brief period of relatively light rain, which was in turn followed by a longer period of moderate rain (0600–1500 UTC 7 May) associated with the maturation of the MCS and the development of the stratiform rain regime (Brandes 1990). The intense convection occurred in a
very localized region, while the moderate rain took place over a much broader area (not shown). The model’s simulated hourly rainfall rates do not reproduce the early intensive localized rainfall well, but nevertheless the solution progresses to the mature stage of the MCS by producing greater rainfall rates for an extended period of time centered around 0900 UTC 7 May. The KF, GR, and BM schemes have solutions that closely resemble the observations, while AK significantly underpredicts the precipitation, and EX overproduces rainfall from 0400 to 0900 UTC and from 1800 to 2100 UTC 7 May.

The hourly ratios of convective to total precipitation from the 36-km and 12-km experiments are presented in Fig. 10. The most striking feature in Fig. 10 is that the partitioning of model-predicted precipitation is vastly different from one scheme to another. At both grid sizes, in one extreme, the AK scheme produces nearly all of the precipitation through parameterized convection throughout the history of the MCS. At the other extreme, the experiment using the GR scheme quickly switches from convective to resolved-scale dominance before the MCS matures around 20 (or 0800 UTC May 7). The KF and BM schemes produce a rainfall partitioning somewhere between that of GR and AK. Overall, the ratio of convective to total precipitation is slightly smaller as grid size decreases, with the exception of AK runs. The large change in the partitioning of the rainfall with respect to resolution for BM may be explained in part by the choice of the relaxation timescale $\tau$. In principle, $\tau$ should be reduced in BM as the grid size is decreased so that appropriate subsaturation is maintained (Betts and Miller 1993). However, in this study the same value of $\tau = 50$ min is used at both 36- and 12-km resolutions.

To further illustrate the precipitation characteristics associated with different CPS, 6-h accumulations of the model’s convective and resolvable-scale rainfall during the early stage of case 4 are presented in Fig. 11 for the 36-km runs. Different partitioning is very apparent in two areas. Over the Oklahoma–Texas panhandle region, where the observed precipitation occurred, KF and BM runs produce significant amounts of convective rainfall, whereas there is very little convective rain in GR and AK runs. In southern and southeastern Kansas north of the surface stationary frontal zone, precipitation is produced almost completely via convective parameterization in the BM and AK runs, while nearly all the rain is generated by the explicit scheme in the GR run. During the subsequent 6 h, as the southern portion of the precipitation system moves into central Oklahoma, some resolvable-scale rain develops in the KF and BM runs. In the northern portion, rainfall over Kansas becomes predominantly nonconvective in the KF run, and some resolvable-scale rain also develops in the BM run. A large portion of the rainfall in all areas of the GR run still is produced via the explicit scheme during this period, but precipitation in the AK run is nearly all convective (not shown). Thus, the reduction in the percentage of the convective precipitation in the KF and BM runs between hours 18 and 24 shown in Fig. 10 is mostly due to the development of resolvable-scale rain in the northern portion of the system, though some resolvable-scale rain also develops in the southern system. The distribution of convective and resolvable-scale rain is similar for the 12-km runs despite considerable small-scale variations in the precipitation fields.

In a very different convective environment in case 1 for the cold season, precipitation occurred along a frontal zone and in the warm sector south of a surface cyclone. Figure 12 depicts the hourly ratio of convective to total precipitation from the 36-km and 12-km runs for this case. Notably, the convective percentage of the precipitation is considerably lower than in case 4 for all four schemes. This is primarily because a significant amount of system total rainfall was occurring along the stationary frontal zone. The rainfall distribution maps (not shown) depict that in the warm sector, precipitation is largely produced via CPSs for the KF, BM, and AK runs, while the explicit scheme is responsible for producing much of the rain along the stationary frontal zone. The convective portion is fairly small everywhere in the GR run. The subtle differences in the way the model handles precipitation via “convective” or “explicit” schemes result in different surface pressure fields (not shown). The location of the surface cyclone center at 0600 UTC February 6 is reproduced well in the KF and BM runs, while the low center in the AK run is about 100 km behind the observed storm center. The GR scheme produces double centers, which are not readily verified.
Although the partitionings of rain by the KF, GR, BM, and AK schemes have been shown to be very different from one another in the cases examined here, Figs. 10 and 12 show that the ratios of convective to total precipitation are similar as the grid size decreases from 36 km to 12 km in the runs with the same scheme. This suggests that the characteristics of these CPSs, although very different from one another, are relatively insensitive to the mesoscale grid resolution. Similar findings are reported in Kuo et al. (1996) for an oceanic cyclogenesis case.

Examination of the convective to total precipitation ratio in the other cases (not shown) reveals some similarities to the ones presented here. In particular, the GR scheme is not convectively active enough in both winter- and summertime environments, especially at the coarser grid size (36 km in this case). The runs with GR typically start with mostly convective precipitation and quickly change to predominantly resolvable-scale rain. Additional experiments show that the low convective to total precipitation ratio for GR is primarily due to the precipitation efficiency parameter in the scheme being too small. Adjustment of this parameter can lead to significant increase in the convective rainfall. On the other hand, the AK scheme is generally too active, and even at the 12-km grid size, it does not allow the explicit scheme to play a significant role. In fact the AK scheme becomes even more active as the grid size decreases from 36 to 12 km, in contrast to other schemes. An explanation for this behavior has to do with its closure assumption. Since the AK scheme relies on grid-scale moisture convergence to determine whether convection can occur, and increased model resolution generally leads to greater moisture convergence, this in turn leads to the simulation of more convective rain. The percentage of convective rainfall in AK runs is the lowest in case 1 (about 50%, Fig. 12d) and is generally above 80% in the other five cases. The AK runs tend to have the highest percentage of convective rain among all the schemes. In contrast to these two extremes, the con-
vective to total precipitation ratios for the KF and BM schemes are somewhere in between. These characteristics appear to be insensitive to model grid sizes in the range 12–36 km and to convective environments.

c. Predictions of surface mesoscale features

As discussed in section 1, evaluations of CPS performance should include weather parameters other than precipitation that are related to the dynamics of convection. In this section, predictions of meso-β-scale surface features associated with CPSs are compared to observed characteristics of those features. Two warm-season cases are examined.

Figure 13 shows the observed surface analysis at 0600 UTC 11 June 1985 for case 6. Figures 14 and 15 display the simulated SLP, hourly rainfall, surface winds, and temperatures at the corresponding time (18-h forecast) using the four different CPSs in the 36-km model. The heavy dashed lines in Fig. 14 indicate the simulated squall-line positions at 0000, 0300, 0600, and 0900 UTC 11 June, respectively. It is
apparent from the figures that there are significant differences among the predictions of the surface fields. In particular, the simulated squall line propagates at different speeds, the elongated mesoscale squall-line trough immediately ahead of the precipitation area is located at different places, and the sea level pressure pattern is dissimilar in different CPS runs. Figure 13 indicates that the observed trough is located from northeastern to southern Oklahoma at this time. In Fig. 14, the KF scheme places the trough from northern to southwestern Oklahoma, or about 140 km behind the observed position. The squall-line positions analyzed at earlier times suggest that the error in the location largely results from the error in the early development stage prior to 0000 UTC 11 June. The squall line is located in western Kansas in the KF run at 0000 UTC 11 June, while it is placed in eastern Colorado in the other runs. The KF scheme is also able to reproduce mesoscale high and low pressure centers that resemble those observed (cf. Figs. 13 and 14a). The GR and BM runs produce strong mesoscale high pressure centers, but fail to produce a
significant mesoscale wake low. In the AK run both the mesoscale high and low are indistinct. Analysis at subsequent times shows that mesolows developed later in the GR and BM runs at forecast hours 19 and 22, respectively.

It should be noted that the KF and GR schemes were previously applied to case 6 at 25-km grid spacing in a hydrostatic version of the PSU–NCAR mesoscale model (Grell 1993; Kain 1994). Both predicted the squall-line position very close to the observed location at this time. However, the “frozen” versions of these CPSs in MM5 are different from the ones used in those case studies, in part due to the developers’ intent to cover a wider range of convective environments. Therefore, the ability of the KF scheme to maintain a fairly small position error for the squall line, despite the change of resolutions and use of the frozen version, represents a significantly favorable result.

The position of the simulated squall line is better illustrated in Fig. 15, where the sharp wind-shift line and strong temperature gradient represent the leading edge of the downdraft outflow boundary. It is again very apparent that the KF run produces the best location of the squall line with the tightest temperature gradients. Also, notice the difference in the temperature fields just behind the simulated squall line. Experiments that use schemes with parameterized moist downdrafts (KF and GR) are able to produce a broad area of cooling, with surface temperatures less than 18°C, especially in western Kansas. This is in agreement with the observations. The BM scheme produces some cooling behind the squall line, probably due to activation of the explicit scheme (and hence explicit cooling by evaporation). The AK scheme, however, fails to produce any appreciable cooling immediately behind the squall line, resulting in a weaker temperature gradient. As shown in section 5b, the GR scheme is not very active in the 36-km simulations. Experiment GR36-6 processes precipitation mostly through the explicit scheme, which develops the stable rain too slow-
ly due to more severe spinup problems at coarser resolutions. As a result, evaporative cooling is delayed. This may help to explain why GR fails to move the squall line at the correct speed. Admittedly, at this resolution all CPS experiments exhibit errors in the speed of this fast-moving squall line, but clearly the KF experiment produces a solution that is closest to the observations. The results from the 12-km experiments demonstrate significant improvement in simulating the squall-line position in all CPS runs, as well as in the run with EX only, although the cooling behind the squall line in the AK run is still minimal (not shown).

Another illustration of the different abilities of the CPSs to predict surface weather parameters is found in the simulations for case 4. Figure 16 shows the observed winds and mesoscale pressure perturbations at 0900 UTC 7 May 1985. Figure 17 shows the predictions of wind and hourly rainfall at the corresponding time (21-h forecast) from the 36-km simulations using the four CPSs. Again it is very apparent that the schemes designed with moist convective downdrafts (KF and GR) are able to advance the outflow boundary and precipitation fields into central Oklahoma much faster than the other two schemes (BM and AK). The KF and GR schemes are also able to produce wind shifts and temperature falls behind the squall line (not shown). Comparison of the 36-km results from the KF and GR schemes indicates, however, that GR propagates the squall line too rapidly in this case, while KF simulates approximately the correct position for the mesoscale features.

Prediction of a variety of surface weather parameters is very important for providing accurate and useful forecasting services to the public. These results demonstrate that not only are the more sophisticated CPSs capable of producing better warm-season rainfall forecasts, but they also better predict the mesoscale surface weather features that accompany them.

6. Summary

The purpose of this study is not to discard one CPS or to favor another. The intent has been to identify characteristic responses and systematic errors associated with mesoscale applications of these CPSs in continental environments. This information should be useful for the design of future numerical prediction systems and for improvement of parameterization schemes intended for operational use at these scales.

It is noted that part of the precipitation forecast error does not result directly from a given CPS, but may be related to uncertainties in the model’s initial conditions. Such details can be very important for obtaining a good forecast, especially when large-scale dynamic forcing is weak. Some errors may be related to deficiencies in internal model components or to the coupling of a CPS with other model physics. We also recognize that the evaluation process becomes more difficult when dealing with fine-resolution models. Statistical verification scores do not always provide a fully “objective” assessment of a forecast. Development of better evaluation methods for the mesoscale is still very much needed.

Bearing in mind the limitations and assumptions made in this study, much useful information can be extracted from the results presented here.

1) The MM5 6-h rainfall predictions with the four CPSs show generally good skill, even for the higher threshold amounts in four out of six precipitation events examined in this study. The skill is well reflected in the conventional verification statistics, as well as examination of individual precipitation maps. As expected, the CPSs perform better in the cold-season events than in the warm-season cases, and the predictive skill of each CPS has a fairly large case-to-case variation in the warm-season events. None of the schemes consistently outperforms the others by a wide margin or in all measures of skill. The predictions of precipitation volume are generally better than either the areal predictions or the maximum rainfall predictions, indicating that the model receives about the right amount of latent heating from condensation. The model tends to overpredict the area of light rainfall and underpredict the heavy amounts.

2) There is, in general, a small increase in forecast skill for most measures of accuracy as the model’s grid size decreases from 36 to 12 km. The largest gain in
the predictive skill for the fine-resolution model is for 
the heavier rainfall thresholds. The increased model 
resolution also enables the explicit scheme to be more 
active (except for the AK runs).

3) Some systematic errors that may be identified
with these four CPSs are the following. Although the 
GR and KF schemes perform well in predicting the 
life cycles and the total volume of precipitation 
events, they have a tendency to overpredict the area 
of light rain. The low precipitation efficiency in the 
GR scheme should be addressed to correct the prob-
lem associated with underprediction of convective 
precipitation. The BM scheme has generally better 
skill for the light-rain areal coverage and the maxi-
mum rain rates, but it has a tendency to overpredict 
the areas of moderate-to-heavy rainfall in warm-sea-
son events. The AK scheme has the most difficulty 
in predicting the warm-season events at both 36- and 
12-km grid sizes. This supports the results from pre-
vious studies (e.g., Grell 1993) that the moisture-con-
vergence closure for a CPS can lead to underprediction of rainfall amounts in a mesoscale model. The AK scheme also severely limits the involvement of the explicit scheme in the precipitation process, even at 12-km grid size. Furthermore, the AK scheme tends to produce a large variance around the mean of the rain maxima (large MAE), suggesting lower reliability for warm-season heavy rain than the other schemes. All four schemes apparently have difficulty in predicting high-cloud-base convection in case 5.

4) Despite underpredicting the rain maxima, the KF scheme performs well consistently for the six cases examined here at both 36- and 12-km grid sizes. The scheme does well in predicting the occurrence and location of the maxima (PMS and MPE scores), and in minimizing the generation of spurious maxima (the NSM score). This suggests that the convective available potential energy-based closure assumption as used in the KF scheme may have some advantage in the mesoscale model. Inclusion of a more sophisti-
The KF and GR schemes, which are designed for mesoscale applications and include parameterized moist downdrafts, demonstrate considerable skill in predicting warm-season precipitation. These schemes are also capable of predicting the position and timing of important surface weather parameters, such as wind shifts, temperature gradients, mesohighs, and mesolows. The BM and AK schemes, however, do not simulate the mesoscale surface features well because they have no parameterized downdrafts.

6) The partition of precipitation between convective and resolved-scale (explicit) processes is dramatically different from one scheme to another, although it is generally consistent in 36- and 12-km runs with the same scheme. The characteristics of the partition for each scheme do not appear to be highly sensitive to model grid sizes in the ranges considered here, nor to the convective environments. Based on some observational evidence of convective and stratiform precipitation and their transition in MCS-type convective systems, the ratio of convective to total rainfall does not appear to be realistic for the AK scheme (nearly all convective) or the GR scheme (nearly all explicit) in these cases.

7) The underprediction of precipitation events at the beginning of the 36-km simulations and overprediction of moderate-to-heavy rainfall at both 36- and 12-km grid sizes from EX runs suggest that it is not an appropriate scheme to be used by itself at these model resolutions.

In conclusion, improving mesoscale quantitative precipitation forecasts remains a very challenging problem, even when a model is equipped with a sophisticated subgrid-scale convective scheme. Overall, this study demonstrates that, for a variety of warm- and cold-season cases, there are significant advantages to a CPS, which directly represents moist convective downdrafts, such as the KF and GR schemes. Much work remains to be done to adequately understand how a CPS interacts with other components of the model, especially in meso-β-scale applications. Convective triggering mechanisms within the more sophisticated CPSs are certainly an area requiring further study, which eventually may reduce the likelihood of serious errors in precipitation predictions. Improvement of the model’s initial conditions is also crucial for reducing errors in precipitation forecasts.

**Acknowledgments.** The authors express appreciation for many helpful discussions with Drs. J. S. Kain, J. M. Fritsch, and Y.-H. Kuo. Thanks also go to Tom Black of NCEP for providing the Fortran source codes of the Betts–Miller scheme, and Dr. Yu-Bao Liu for providing graphic software. We would like to thank two anonymous reviewers for their constructive criticism, which helped to improve the presentation of this study. This research was funded by the U.S. Department of Energy through Grant DE-FG02-90ER61071 and by the Federal Aviation Administration and National Science Foundation (NSF) through Grant ATM-9203317. Computer resources were made available through the Earth System Science Center of The Pennsylvania State University and through NCAR, which is supported by NSF.

**Appendix A**

**The Cumulus Parameterization Schemes**

Special attention is given to how each CPS produces subgrid precipitation and identification of the key closure assumptions relating the convection to the resolved-scale environment. Details of the CPS formulations appear in the referenced papers.

**a. Anthes–Kuo scheme**

The AK scheme is a variation of the original Kuo scheme (Kuo 1974). The scheme uses the column-integrated moisture convergence $M_t$ to determine where convection will occur and how intense it will be. Convection is initiated when conditional instability exists (determined by parcel theory) and $M_t$ exceeds a threshold value in the column. The moisture convergence is partitioned into moistening of the column and con...

![Fig. 16. Observed surface winds (m s$^{-1}$) and mesoscale pressure perturbations (mb) for case 4 at 0900 UTC 7 May 1985 (reproduced from Zheng 1993). A full barb is 10 m s$^{-1}$, and the contour interval for pressure is 1 mb.](image)
vective precipitation, such that the precipitation is defined as

$$PR = (1 - b)M,$$  \hspace{1cm} (A1)$$

where $b = 2(1 - \text{RH})$ and $\text{RH}$ is the mean relative humidity of the column. This scheme does not include a cloud model nor parameterized convective downdrafts.

b. Betts–Miller scheme

The BM scheme has recently received considerable testing at the European Centre for Medium-Range Weather Forecasts and the NCEP (Betts and Miller 1993; Janjic 1994). The version of BM applied here is similar to that currently used in the operational eta-coordinate model at NCEP. Unlike the other three CPSs, the BM scheme is a lagged convective adjustment scheme. It adjusts the model’s thermal and moisture
structures toward specified reference profiles that reflect the quasi-equilibrium state established by deep convection (Betts 1986). The parametric precipitation resulting from this adjustment is defined as

$$PR = \int_{p_a}^{p_r} \frac{q_s - q}{\tau} \frac{dp}{g},$$

(A2)

where \( q \) is the model's specific humidity, \( q_s \) is the reference-profile specific humidity (a function of height), \( \tau \) is the timescale over which the adjustment occurs, and \( p_a \) and \( p_r \) are pressures at cloud top and bottom, respectively. The parameters used to specify the temperature and moisture profiles and the adjustment timescale are those used by NCEP, which are different from those originally employed by Betts and Miller (1986): \( \tau = 50 \) min, \( \alpha = 0.9 \), \( S_b = -38.75 \) mb, \( S_m = -58.75 \) mb, and \( S_t = -18.75 \) mb, where \( \alpha \) is the stability parameter (\( \alpha = 1 \) gives moist adiabats) and \( S \) is the subsaturation parameter defining the pressure difference between the level at which an air parcel becomes saturated and its originating level (Betts 1986). The subscripts \( b \), \( m \), and \( t \) of \( S \) denote the cloud base, freezing level, and cloud top, respectively.

This version of the BM scheme does not explicitly consider the effect of convective downdrafts, although the energy conservation constraint in the scheme may allow boundary layer temperature to be reduced (hence implying downdraft activity). A more recent version has included this effect (see Betts and Miller 1993). The BM scheme also parameterizes the effects of shallow convection, but its impact on precipitation in the cases presented here is minimal.

c. Grell scheme

The Grell scheme is a one-cloud version of the Arakawa–Schubert scheme (1974) with parameterized downdrafts and was originally applied in the PSU–NCAR model. The closure used by the GR scheme is the quasi-equilibrium assumption, as proposed by Arakawa and Schubert (1974). The subgrid precipitation is calculated as

$$PR = I_1 m_b (1 - \beta),$$

(A3)

where \( I_1 \) is the integrated condensate in the updraft, \( m_b \) is the cloud-base mass flux of the updraft, and \((1 - \beta)\) is the precipitation efficiency, which is assumed to be a function of the mean wind shear in the lower troposphere. The effects of the convective-scale downdrafts are parameterized in this scheme.

d. Kain–Fritsch scheme

The KF scheme is similar to the Fritsch–Chappell scheme (1980), but with improvements in detrainment effect and cloud model. It was also applied originally in the PSU–NCAR model. The closure assumption in KF parameterization is the same as that used in the Fritsch–Chappell scheme. That is, the convection is determined by CAPE at a grid point. Once convection is triggered, CAPE is assumed to be removed in a grid column within an advective time period (Kain and Fritsch 1993). The KF scheme utilizes an improved cloud model, which is mass-conservative, allows cloud–environment interaction, includes parameterized moist downdrafts, and has a detailed representation of cloud physics, including entrainment and detrainment at cloud edge (Kain and Fritsch 1990). The convective precipitation is computed as

$$PR = ES,$$

(A4)

where \( E \) is the precipitation efficiency, and \( S \) is the sum of the vertical fluxes of vapor and liquid at about 150 mb above the lifting condensation level.

APPENDIX B

Definition of Mean Error and Mean Absolute Error

A mean error is defined as

$$ME = \frac{1}{N} \sum_{n=1}^{N} (P_m - P_o),$$

(B1)

where \( P_m \) and \( P_o \) are model and observed precipitation, respectively, and \( N \) are grid points where either observed or simulated precipitation exceeds a threshold amount. A mean absolute error is defined similarly as

$$MAE = \frac{1}{N} \sum_{n=1}^{N} |(P_m - P_o)|.$$

(B2)

REFERENCES


——, and ———, 1993: The Betts–Miller scheme. The Representation


Trier, S. B., and D. B. Parsons, 1993: Evolution of environmental


