Tellus (1999), 51A, 710–727
Printed in UK. All rights reserved

TELLUS
ISSN 0280–6495

Development and application of an operational, relocatable, mesogamma-scale weather analysis and forecasting system

By CHRISTOPHER DAVIS[1]*, THOMAS WARNER[1,3], ELFORD ASTLING[2] and JAMES BOWERS[2], [1]*National Center for Atmospheric Research*†*, Boulder, Colorado;* [2]*West Desert Test Center, Dugway Proving Ground, Dugway, Utah;* [3]*Program in Atmospheric and Oceanic Sciences, University of Colorado, Boulder, Colorado, USA*

(Manuscript received 14 September 1998; in final form 5 July 1999)

ABSTRACT

We report on the results from an operational forecast system built to predict local circulations forced by complex terrain and other variations in land-surface characteristics. The cornerstone of the prediction system is the Penn State University/National Center for Atmospheric Research (PSU/NCAR) mesoscale model, version 5 (MM5), a nonhydrostatic regional model. The specific application reported herein is for the region surrounding Dugway Proving Ground (DPG) in west- central Utah. The nature of the terrain requires a horizontal resolution of about 1 km in order to capture the important local features. This resolution is achieved by the use of grid nesting. To our knowledge, this resolution is finer than that being used in any other currently operational forecast system tasked with local and regional weather prediction. For verification purposes, forecasts are stratified according to season and mean-flow characteristics. Data for verification consist of DPG surface mesonet data. Root-mean-square errors (RMSE), time mean circulations and spatial anomaly correlation statistics are computed and composited for each hour of the day. These are compared with identical forecasts of lagged persistence, with the time lag being 24 h (MM5 forecasts are all initialized at 1200 UTC). For wind and temperature, the RMSE from MM5 is consistently lower than that of persistence. In addition, MM5 shows skill in predicting the time-mean circulations on the test range (variations of a few m/s and °Celsius). MM5 forecast errors grow slowly with time until around sunset, after which they decrease slightly, suggesting that local nighttime "forcing" dominates the error growth, as the surface layer decouples from the free atmosphere. Finally, spatial anomaly correlations suggest that the non-systematic, range-scale circulations exhibit low predictability.

1. Scientific motivation

Lorenz (1969) used a simplified barotropic model to propose a theoretical limit of atmospheric predictability as a function of the scale of motion. Scales of motion smaller than about 20 km were found to be predictable for not more than an hour or so. Though turbulence in this model was highly idealized, this basic notion of predictability on the meso-gamma scale (Orlanski 1975) has persisted. The dynamics in the above models was strongly controlled by initial conditions. In the case of operational forecasts, this translates into a need to observe fine details about the atmosphere on scales of a few km. Because measurements, especially of moisture and thermodynamic profiles, are generally not available on these scales, the practical realization of cloud-scale prediction falls far short of its theoretical limit, whatever that may be.

Mesoscale predictability that is dramatically

* Corresponding author: NCAR, PO Box 3000, Boulder, CO 80307-3000, USA. email: cdavis@ucar.edu.

† NCAR is sponsored by the National Science Foundation.

better than the Lorenz (1969) limit may be possible for the special case of flow over complex terrain or flows generated by surface-heating contrasts, due to the stationarity or periodicity of the forcing. In this case, the ambient flow and the local forcing can determine much of the dynamics. Hence, predictability on small scales may greatly exceed the limits set by internal dynamics alone, and can approach the predictability of the large-scale flow itself.

Relatively little is known about predictability on the meso-gamma scale and the extent to which stationary, local forcing can augment predictability. Case studies (Colle and Mass, 1996) have shown some dramatic examples of mesoscale model forecasts near complex terrain, but a systematic, long-term investigation of day-to-day forecast skill on these scales has only recently become feasible due to increased computer power. Such investigations are still largely impractical due to inadequate observations for model initialization and verification.

An opportunity to examine the utility of fine-scale forecasts near complex terrain, and hence learn about predictability in these environments, has recently emerged from a collaboration between NCAR and the Test and Evaluation Command (TECOM) of the US Army. TECOM operates several facilities across the US, each testing a number of different weapon, defense or surveillance systems. Each range is well instrumented, with anywhere from 15–20 surface mesonet stations, 1–2 wind profilers, radiosondes and other, specialized measurements (tethersondes, radar, etc.) covering an area of roughly 100 km by 100 km. The high density of instruments and the location of many of these ranges near complex terrain or significant land-surface contrasts makes attractive an investigation of model forecast skill. In addition, to the extent that models are skillful, the forecasts generated by such models are potentially of great benefit to planning and conducting range operations.

In this paper, forecasts generated by the PSU/NCAR mesoscale model, version 5.2 (MM5, Grell et al., 1994) for the area surrounding Dugway Proving Ground (DPG), in west-central Utah (Fig. 1) will be examined. The DPG is largely a desert environment. Relatively little is known about the meteorology of such environments, primarily because of limited observations. Thus, the fact that DPG is well instrumented allows detailed examination of one such desert environment. As is apparent from Fig. 1, DPG is surrounded by steep mountains whose peaks are 1–2 km above the desert floor. These naturally play a major role in diurnally forced mountain-valley circulations. In addition, there are salt flats, also known as playas, whose characteristics differ markedly from those of the desert or nearby arid grasslands. Playas have a larger thermal inertia and higher albedo than the desert (Physick and Tapper, 1990). In addition, Ulaby et al. (1975) suggest that the moisture fluxes from the playa over western Utah are substantially higher than from the surrounding desert. Hence, playa temperature changes lag those over the desert, and thermally direct circulations, analogous to the land-sea breeze, can result. Finally, DPG is near 40°N latitude and thus is strongly affected by synoptic-scale transients, especially in the cool season. Thus, much of the emphasis of this work, and part of the rationale for the design of the forecasting system (Section 2), involves the importance of synoptic-scale transients interacting with local physiographic features.

An important aspect of a model evaluation study is a comparison with some other plausible method of generating forecasts. Some kind of persistence is often used as a benchmark for the skill of numerically generated local forecasts (Vislocky and Fritsch 1997). Because of the predominance of the diurnal forcing of circulations in desert environments, a logical benchmark is a lagged persistence forecast, with the lag being 24 h. This amounts to predicting that tomorrow will be identical to today. We will show that MM5 out-performs this benchmark for forecasts of 24 h. We will also show that MM5 provides better forecasts of low-level temperature than the Eta model from the National Centers for Environemntal Prediction (NCEP). RMSEs for winds are similar in the two models, although MM5 exhibits some skill at explicitly forecasting the recurring local range circulations on scales of 20–40 km owing to its finer resolution relative to Eta. The meso-gamma-scale deviations from the regularly occurring circulations are found to be unpredictable by MM5. These fluctuations will be shown

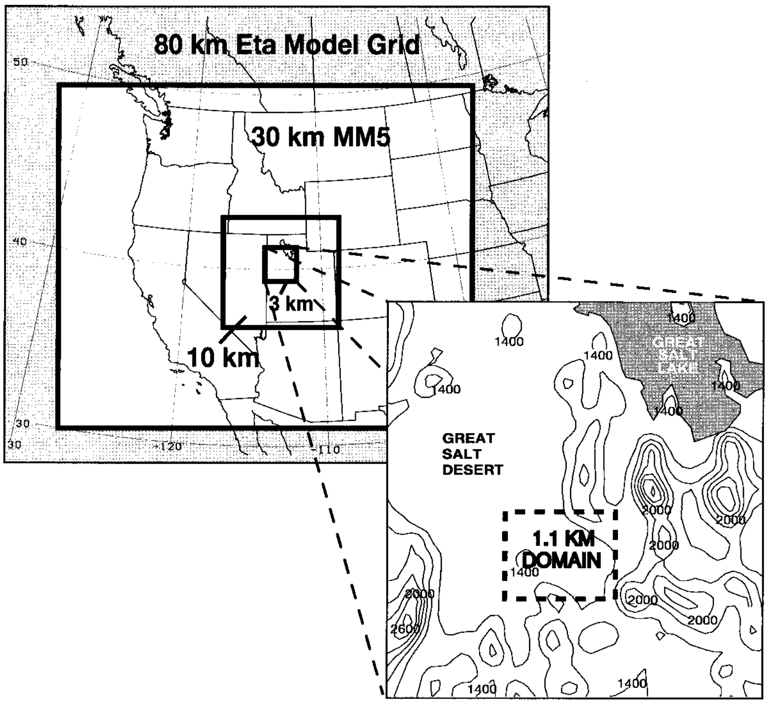

Fig. 1. MM5 configuration for DPG simulations. Domain 3 (3.3 km resolution) is expanded and terrain is contoured with an interval of 200 m.

to fall closer to the predictability/scale relationship noted by Lorenz.

2. Description of the forecast system

2.1. *Overall system concept*

Although we will focus on the application of MM5 at DPG, the larger goal is to design a relocatable forecast system that will run at each major TECOM test facility in the US. Each system must involve data collection, both from local range sources and the National Weather Service (NWS). Local data are subject to automated quality-control algorithms, which examine the spatial and temporal consistency of observations. Each system must have a display capability to permit viewing of both observations and model output. Each system must also perform the tasks of archiving and retrieving observations and model output, both for detailed analyses following tests and for the amalgamation of a range climatological data base. The system in its entirety has been named “Four-Dimensional Weather (4DWX)” in recognition of the need to depict meteorological conditions near test ranges at high resolution in both space and time (see Weingarten et al. (1996) for an overview of meteorological concerns at each test range and NCAR staff (1996) for a discussion of the different components of 4DWX).

We point out that there are predictive models within 4DWX other than MM5. These include an automated thunderstorm forecasting system (which functions independently from MM5) and a plume dispersion model, which is driven by MM5-produced winds. These components will be described in separate articles. In principle, other specialized models can be driven by MM5 output at other test ranges, depending on the local test-support needs.

2.2. *MM5 component of 4DWX*

The MM5 model integrates the fully compressible primitive equations with a semi-implicit scheme for integrating acoustic modes (Grell et al., 1994). Temporal differencing is accomplished with leapfrog steps with an Asselin filter. The vertical coordinate in the nonhydrostatic model is a terrain-following coordinate

$$\sigma = \frac{p_r - p_{top}}{p_s - p_{top}}, \quad (1)$$

where $p_r(z)$ is the reference-state pressure; p_{top}, the top of the model, is 50 mb; and p_s is the surface pressure. An upper-radiative boundary condition is imposed at the lid to reduce spurious reflection of gravity wave energy from the upper boundary.

At lateral boundaries, the model is relaxed to the background solution across a zone of five grid points. In this case, the background is provided by forecasts from the National Centers for Environmental Prediction (NCEP) Eta model (Black, 1994), obtained at 80 km resolution. The outer MM5 domain features a 30 km grid spacing. Three nest levels are inserted with horizontal resolutions of 10, 3.3 and 1.1 km (Fig. 1). On all domains, 31 vertical levels are used, stretched nearly uniformly from a vertical resolution of 80 m between the lowest two levels to just over 1 km in the lower stratosphere. Each nest feeds back onto the next coarser domain.

Standard MM5 physics options are employed, including: (1) the Grell cumulus scheme (Grell, 1993) on the 30 and 10 km domains; (2) the Medium Range Forecast (MRF) model Planetary Boundary Layer (PBL) scheme, which was obtained from NCEP (Hong and Pan, 1996); (3) the Dudhia radiation scheme (Dudhia, 1989); (4) a simple ice microphysical scheme (Dudhia, 1989). The Grell scheme is an adaptation of the Arakawa–Schubert (1974) parameterization in which convection occurs in proportion to the rate of destabilization by grid-resolved processes. It is a single-cloud scheme with updraft and downdraft fluxes, and has been tuned for use on grid sizes between 10 km and 30 km. The MRF PBL scheme is a 1st-order closure scheme which contains four stability-dependent regimes. The 4th, free convecting regime frequently characterizes the daytime continental PBL. In this regime, explicit vertical transfer of heat and moisture occurs between the bottom and top of the PBL. This is designed to mimick the effect of large, vigorous boundary-layer eddies that are not resolved on the grid scale. The Dudhia radiation package accounts for long-wave and short-wave radiative absorption and emission from clouds and clear air. No further spectral decomposition of radiation is performed. The scheme generally provides accurate surface radiation fluxes, especially in clear air. The Dudhia ice scheme is the crudest treatment of microphysics in MM5 which allows ice. Water mass and cloud mass are each characterized by a single prognostic variable, which converts from rain to snow or cloud water to cloud ice across the freezing point. Other multi-phase microphysical schemes are standard options within MM5. The choice of the simplest scheme was motivated by the paucity of precipitation near Dugway (25 cm per year) and the need for minimal computational expense. Sophisticated microphysical schemes can easily add 10%–20% or more to the run-time of the model and this was deemed unacceptable.

Some modifications to the standard MM5 were employed for the 4DWX system. A high-resolution US Geological Survey Earth Resource Observing System (USGS/EROS) data set was used, with vegetation type mapped at 1 km resolution (Loveland et al., 1995). Over 200 categories were represented in this dataset. However, within the 3.3 km and 1.1 km MM5 domains where these data are needed, only about 41 categories appear. Of these, only nine categories each populate more than 1% of the grid. Albedo, moisture availability (MA), emissivity, roughness length and thermal inertia were estimated from previously published data or direct measurements at DPG (Table 1). Although precipitation occurs infrequently at DPG, it does dramatically alter the albedo and moisture content of the soil. Because MM5 does not currently have a sophisticated land-surface model, a simple user-controlled switch was implemented to run the forecast with a set of "wet" or "dry" surface parameters (Table 1). One of the largest factors controlling surface physics in winter is the presence of snow cover. On the inner domains, snow cover is primarily a function of elevation; hence, the only observation used to generate a snow cover field is a user-defined snow level, a constant reset at the start of each forecast.

A sponge layer of the type discussed by Klemp and Lilly (1978) was introduced in the lower stratosphere to augment the damping of gravity waves by the upper radiative boundary condition described in Grell et al. (1994). The upper radiative boundary condition has been shown to greatly

Table 1. *Land use parameters for summer (15 April–15 October) and winter (15 October–15 April); numbers in parentheses indicate values used for wet conditions; units of thermal inertia are* $10^2\ J\ m^{-2}\ K^{-1}\ s^{-1/2}$

Vegetation type	(%)	Albedo (%)	MA (%)	Emissivity (% at 9 μm)	Roughness length (m)	Thermal inertia
Summer						
sparsely vegetated desert	34	16(13)	10(30)	92(94)	0.08	8(17)
mixed rangeland	20	15(13)	15(30)	92(92)	0.1	17
barren	13	50(20)	35(70)	91(97)	0.02	25
playa	9	60(20)	40(80)	90(98)	0.01	21(25)
grassland	5	20(15)	15(30)	92(94)	0.12	8(17)
water	5	8	100	98	0.0001	25
coniferous forest	4	18(15)	20(40)	94(96)	0.5	17
mixed grassland and shrubs	3	20(15)	15(30)	92(94)	0.2	13
desert shrubland	2	20(15)	15(30)	92(94)	0.15	13
Winter						
sparsely vegetated desert	34	19(13)	10(25)	92(94)	0.08	8(17)
mixed rangeland	20	18(15)	15(25)	92(94)	0.10	17(13)
barren	13	50(20)	35(70)	91(97)	0.02	21(25)
playa	9	60(20)	40(80)	90(98)	0.01	21(25)
grassland	5	23(19)	15(30)	92(94)	0.12	8(17)
water	5	8	100	98	0.00	25
coniferous forest	4	18(15)	20(40)	94(96)	0.50	21
mixed grassland and shrubs	3	20(15)	15(30)	92(94)	0.20	17
desert shrubland	2	20(15)	15(30)	92(94)	0.15	17

reduce reflection of gravity waves, but it does not perform perfectly in non-idealized flows. The sponge layer covered approximately the uppermost 6 km of the model domain.

Within the MRF PBL scheme, wind and temperature were extrapolated from the lowest model level to the 2 and 10 m levels above ground level (AGL) for comparison with observations at the same relative height. This extrapolation was done logarithmically in height according to similarity theory (Grell et al., 1994, p. 80.).

2.3. *DPG weather and MM5 configuration*

Field tests at DPG primarily involve the release of smoke or other atmospheric obscurants and the subsequent monitoring of the downstream dispersion properties of the release. This naturally means that correct forecasts of low-level winds and static stability are of primary importance. Tests are often conducted at night, mainly because mixing is weaker. It is important to be able to predict the winds at night in order to place sensors in the proper position relative to the release point. Thus, drainage flows are a major concern to forecasters. In practice, the lead time for such predictions needs to be at least 12 h. Decisions are often made during late morning or early afternoon as to how to prepare for a test.

Synoptic-scale highs and lows, with attendant fronts, dramatically affect the local winds, especially during the daytime. Fronts are modified extensively by the local terrain and surface heating. Accurate prediction of the timing of fronts therefore requires an accurate prediction on the synoptic-scale as well as representation of how the front is modified by local surface heating and terrain.

A basic hypothesis governing the MM5 configuration used for DPG is that many of the local circulations result from the interaction of larger-scale weather systems with the local terrain. Thus, high resolution is needed near the complex terrain of interest, but elsewhere only enough resolution is needed to accurately predict the evolution of the large scale. This strategy is also a result of the practical limitation of computer resources. The platform available for these simulations is an 8-node Origin-2000 from Silicon Graphics Inc.

As shown in Fig. 1, we have 3 nests inside the 30 km domain. The location of each nest was chosen partly to minimize gradients of terrain normal to the nest boundary because large gradients of terrain can cause spurious behavior for flow crossing the boundary. Two forecasts per day are performed. The 1st is initialized at 0600 UTC and covers a 14-h period. This forecast, which provides initial guidance for the current day's weather, is available to forecasters by roughly 0600 MST. The second forecast is initialized at 1200 UTC and covers a 24-h period. In this forecast, the 1.1 km domain is switched on halfway into the simulation to cover the nocturnal period when testing is most likely. The 24-h forecast requires about 4.5 h on the 8-node Origin machine.

The first guess for the model initial conditions is supplied by forecasts from the 0000 UTC Eta model, interpolated to the 30 km MM5 outermost grid. Surface and rawinsonde observations are analyzed onto this first guess, resulting in a 30-km analysis. This 30-km analysis is then interpolated to the inner domains to provide initial conditions. Time dependent lateral boundary conditions on the 30 km domain are derived from Eta forecasts at 6-h intervals. The analysis is performed using a modified Cressman scheme in which the influence of an observation decays roughly quadratically with distance. The influence of observations is accentuated along the flow direction, primarily to account for the greater length scales associated with the along-stream direction near jets. We make no attempt to initialize small scale features on the inner domains owing to a lack of observations, especially aloft, on those scales. The initialization strategy used here emphasizes accurate depiction of larger scales in the initial state. The strong local forcing quickly generates realistic local circulations despite the smooth initial conditions.

There are two primary uses of model output. First, MM5 output is displayed using Vis5D, a graphics package which allows users to animate three-dimensional images of model output. Second, MM5 output is interpolated to the location of each of the surface mesonet stations (hereafter known as Surface Atmospheric Measurement Systems, SAMS). These interpolated data will be the focus of the verification effort (Section 3).

3. Verification of MM5 forecasts

Two major elements of model verification are case studies and statistics compiled on a large sample of cases. In this section, we emphasize both daytime and nocturnal range-scale circulations and the model's ability to capture them. Statistics are presented based on either the entire set of 1200 UTC initialized forecasts that have been run to date (about 170), or on subsets of these forecasts defined roughly by season. We concentrate on surface observations, because most of the phenomena of interest are near-surface circulations. Also, the surface data form a more complete record for the compilation of long-term statistics. The locations of the SAMS stations used for verification are shown in Fig. 2.

3.1. Case of 9 September 1997

The synoptic-scale flow on this particular day featured a quasi-stationary ridge at upper levels

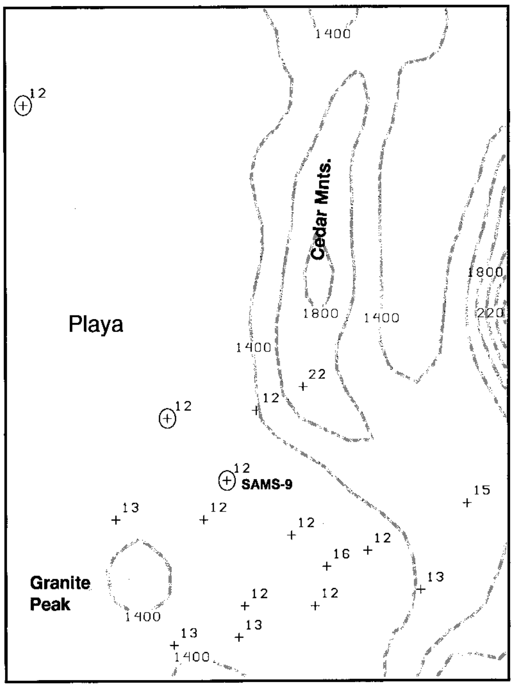

Fig. 2. Locations of SAMS stations, with station elevation given in 100's of meters, superposed on MM5 terrain on Domain 3. Circled station identifiers indicate stations over the playa. Note that actual station elevations are not always consistent with MM5 terrain.

over the western US, with a developing trough over the central and eastern US (not shown). The low-level geostrophic wind over DPG was light southerly at around 0000 UTC 10 September. Local circulation features prevailed during this day. The dominant local feature on this day appears as an enhanced temperature gradient in the simulation in Fig. 3 (at 2200 UTC 9 September), with an along-front length scale of only a few tens of kilometers and a cross-front scale of about 10 km. The general area of surface divergence over the northwestern quadrant of the plot, and the mesoscale front, result from the existence of the salt flats in the model. The properties of this playa land surface strongly contrast with the properties of the surrounding arid grass and shrub land. For instance, the moisture availability and thermal inertia are three times larger on the playa than on the surrounding surface. As a result, a diurnal, thermally direct circulation sets up which is reminiscent of the sea breeze. This feature is analogously called the salt breeze (Physick and Tapper, 1990).

The forecast and corresponding surface observations are shown on the 1.1 km domain in Fig. 4. The timing of the salt breeze was 2–3 h early in the model, but the strength and movement of the predicted salt breeze front resembled closely that of the observed front. A vertical cross section (Fig. 5) reveals a structure which closely resembles a gravity current. The updraft at the leading edge is about 0.5–0.6 m/s only a few hundred meters above the ground. Static instability results over the lowest 200 m as the cooler air is heated from below. Finally, the frontal circulation significantly increases the stratification aloft as subsidence

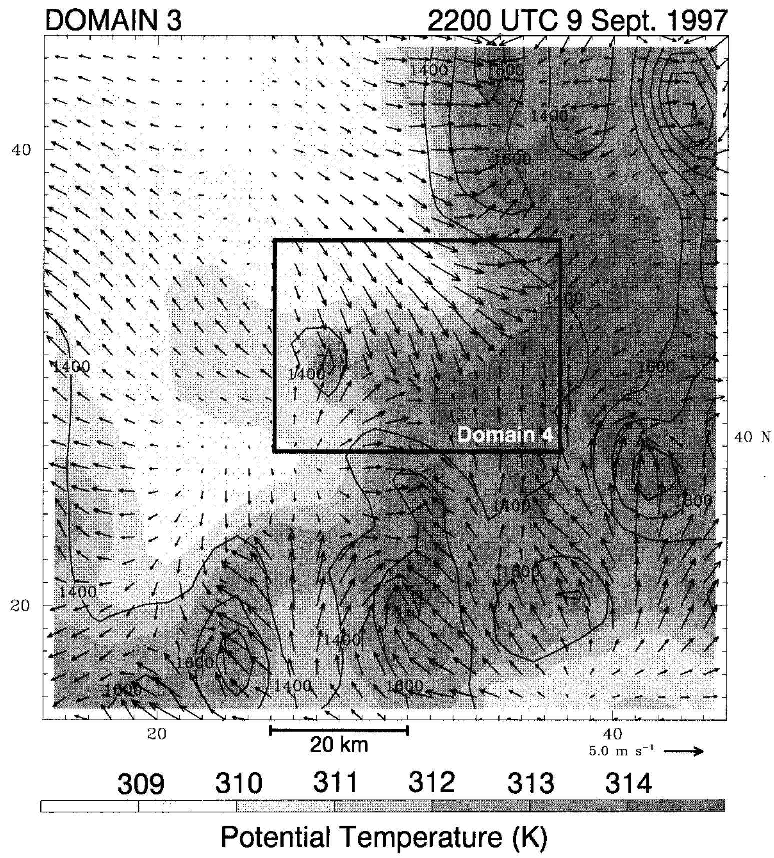

Fig. 3. 10-h forecast on Domain 3 for 2200 UTC 9 September 1997. Displayed are surface ($\sigma = 0.996$) wind and potential temperature (shaded). Numerical axis labels on left and bottom boundaries refer to model grid points.

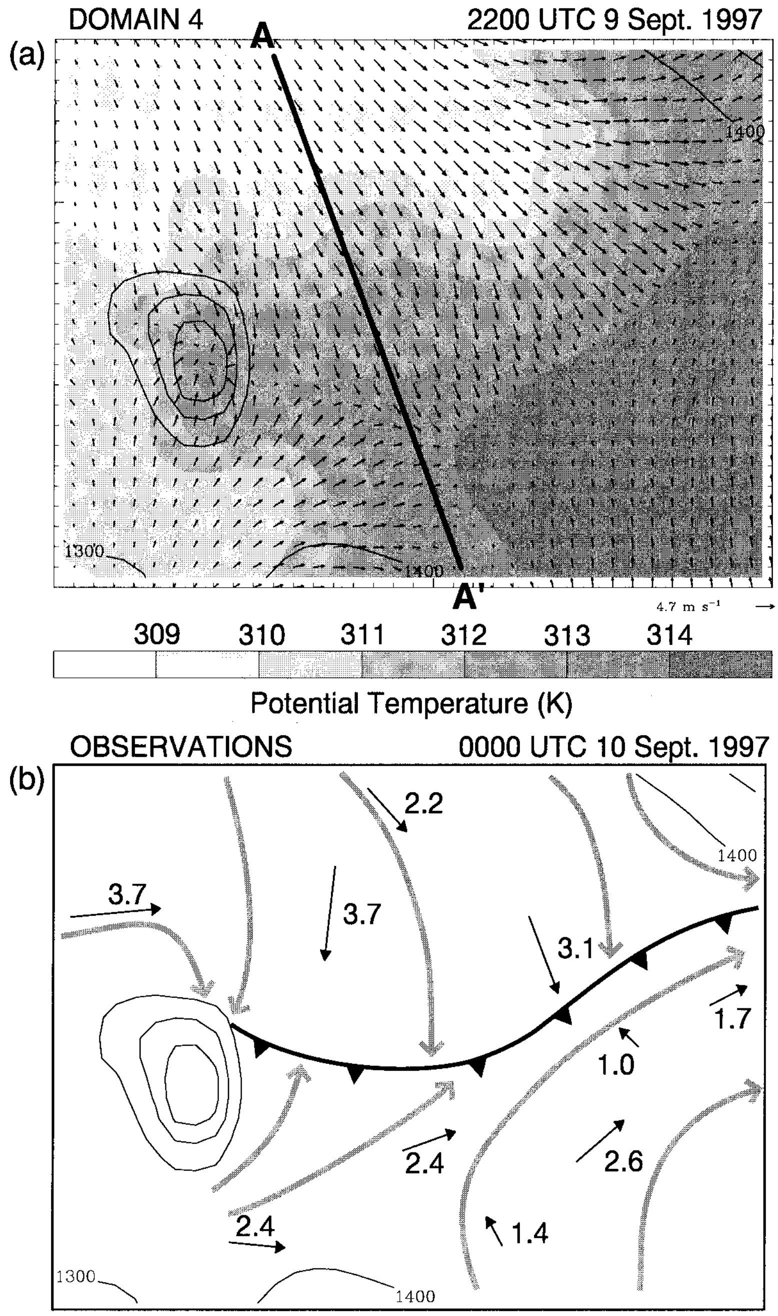

Fig. 4. (a) As in Fig. 2, but for Domain 4, the 1.1 km resolution domain. Terrain is contoured in 100 m intervals; (b) Hand streamline analysis of observed conditions at 0000 UTC 10 September, mapped onto the 1.1 km domain.

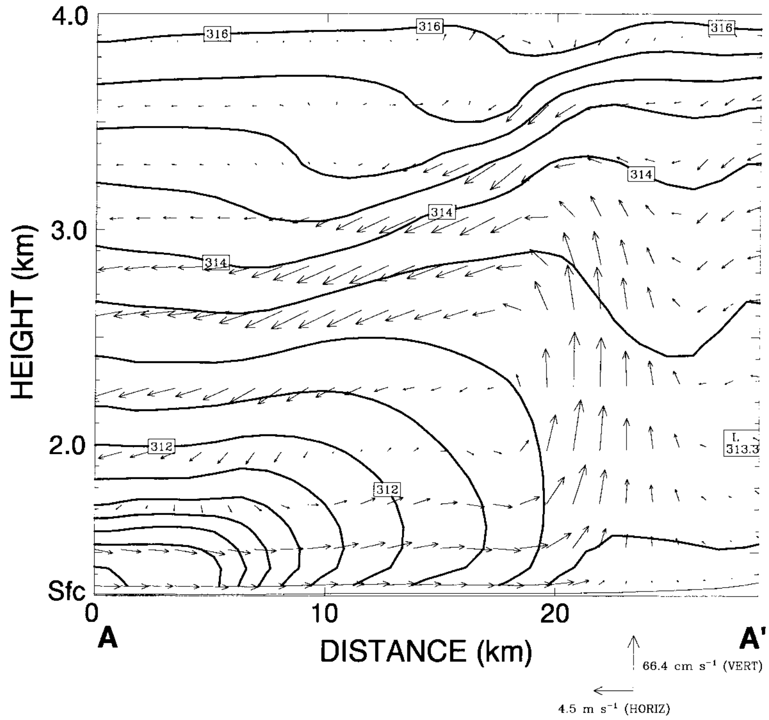

Fig. 5. Vertical cross section (see Fig. 4a for location) showing predicted potential temperature and wind parallel to the section at 2200 UTC 9 September.

occurs at about 3 km MSL or 1.7 km AGL behind the leading updraft. This increase in stratification appears distinct from that at lower altitudes that is due mainly to horizontal advection of cooler air at the surface.

As a test of the importance of horizontal resolution for capturing this type of local circulation, we performed a sensitivity simulation in which domains 3 and 4 were removed, leaving 10 km as the finest resolution. Perhaps as would be expected, the 10 km simulation (Fig. 6) is not simply a coarser version of the full simulation. Over much of the Great Salt Desert, the 10 km resolution is adequate to depict mesoscale divergence at the surface, as the daytime temperature rise over the playa lags that of the desert and a thermally direct circulation develops. However, a secondary divergence center, located south of Granite Peak (location in Fig. 2), is absent in the 10 km simulation. This results in the erroneous generation of a northerly component of surface wind over the southern part of the test range. Many aspects of the local mountain-valley circulations are also absent at 10 km resolution. The coarser depiction of the terrain does not capture much of the local physiography that is important for generating the local flows, but the mesocale thermal contrast between desert and playa is reasonably represented.

3.2. Model performance by season and flow regime

We now present statistics based on the entire sample of MM5 forecasts for which observations were available, which totals about 170 forecasts for the period 20 August 1997 to 31 July 1998. Only forecasts initialized at 1200 UTC are considered. The system was not operational from the period 20 December 1997 to 1 March 1998 due to the transfer of the computing platform from NCAR to DPG. Since early April, the system failure rate has been less than 10%; most of the problems have involved a failure to properly transfer the Eta model output from NCEP, or power interruptions at DPG.

Forecasts were grouped into seasons as follows:

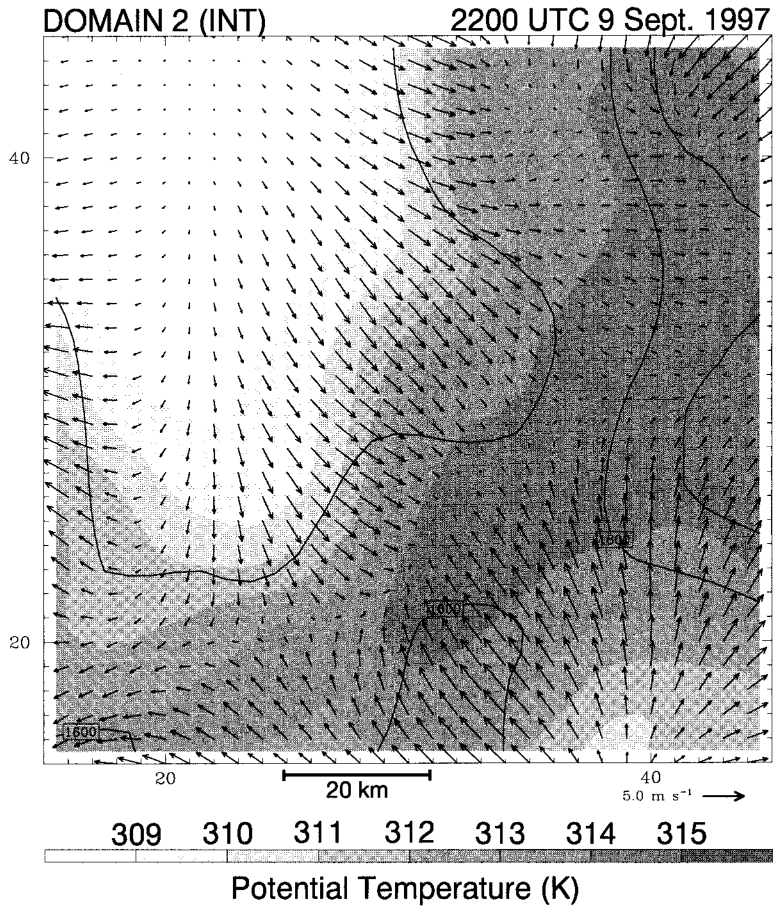

Fig. 6. As in Fig. 3, but for a forecast at 10 km resolution interpolated to domain 3.

(1) Autumn, 1 September to 10 November 1997; (2) Winter, 11 November 1997 to 19 December 1997; (3) Spring, 1 March to 31 May 1998; and (4) Summer, 1 June to 31 July 1998. The numbers of forecasts during these periods are 20, 25, 74 and 55, respectively. For comparison of forecasts with SAMS data, model output at 10 m (wind) and 2 m (temperature) was interpolated bi-linearly to the location of each SAMS. No correction was made for the fact that the elevation of some SAMS stations differed from the height of the surface in the model. Using a fuzzy-logic algorithm based partly on the work of Shafer and Hughes (1996), the SAMS data were quality-checked. A set of three influence functions were created to map each instrument's step, range, and persistence values over time into a set of weights which were in turn combined into a final weight in continuous [0, 1] space. Values of this (fuzzy) weight less than 0.5 were grounds for rejecting the observation.

Three types of statistics were calculated for the wind and temperature. First, the bias and root-mean-square difference (RMSD) between the model forecasts and the SAMS data were computed. Second, the time averages at individual stations were calculated and displayed on a DPG-area map so that the ability of MM5 to predict systematic local circulations could be assessed. Finally, we computed anomaly correlations. To define an anomaly, we first removed the time average (the average of conditions for a given time of day) from each station. The range mean of these departures from the time average was computed, and the anomaly was defined as the deviation from this range mean value.

The profound effect of the diurnal cycle in desert regions like DPG influenced our methodology in two ways. First, all statistics were computed as a function of hour of the day. Because all forecasts included in the verification were initialized at 1200 UTC, each forecast hour corresponds to a given hour of the day. Second, one of the benchmarks against which model forecasts were measured is a lagged persistence which assumes that the time trace of wind and temperature for the previous day is the same as for the current day.

3.2.1. Bias and RMSD. Fig. 7 shows the composite MM5 biases for the entire period of record. The model biases shown in Fig. 7 differ somewhat from those reported for MM5 in previous studies. For example, Manning and Davis (1997) analyzed 60 MM5 forecasts over the western and central US and found a pronounced bias in the temperature diurnal cycle (amplitude too small) and winds that were too strong at the surface. Fig. 7 shows diurnal cycles of temperature and wind that are fairly close to those observed, but slightly weaker. The main feature of the temperature behavior shown in Fig. 7 is a cold bias of 1–2°C. We also note that there is a bias in the zonal (*u*) wind component of about 1 m/s. In addition, forecasts of the meridional (*v*) wind component exhibit a bias that varies with the diurnal cycle, with a negative bias in the afternoon and early evening and a positive bias late at night. A two-sided *t*-test was performed on the biases. Any bias of more than about 0.25 m/s or 0.25°C is significant at the 99% confidence level. Thus, all the major features described above are statistically significant. Seasonal statistics (not shown) indicate that the westerly bias occurs mainly in spring and summer, and that the bias in the diurnal cycle is largest during autumn. Unlike results from Manning and Davis (1997) and another (unpublished) verification of MM5 forecasts (C. Mass, 1998, personal communication), the mean surface wind speed in MM5 is too weak when compared with the SAMS data. The primary negative wind speed bias occurs in the early evening, around the time when the PBL transitions from an unstable to a stable regime.

The RMSD statistics in Fig. 8 indicate that the

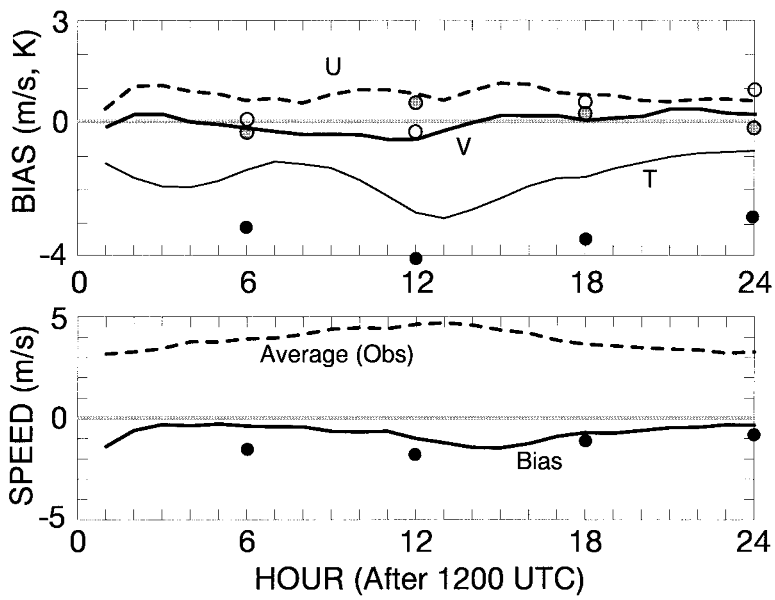

Fig. 7. Composite MM5 biases for all seasons. Top panel shows temperature and zonal and meridional wind component as a function of hour from 1200 UTC (0500 MST); bottom panel shows mean wind speed (dashed) and MM5 speed bias (solid). Black, gray and open circles in the top panel denote Eta model biases in 2-m temperature, meridional wind and zonal wind, respectively. Black circles in the bottom panel denote Eta model wind speed bias.

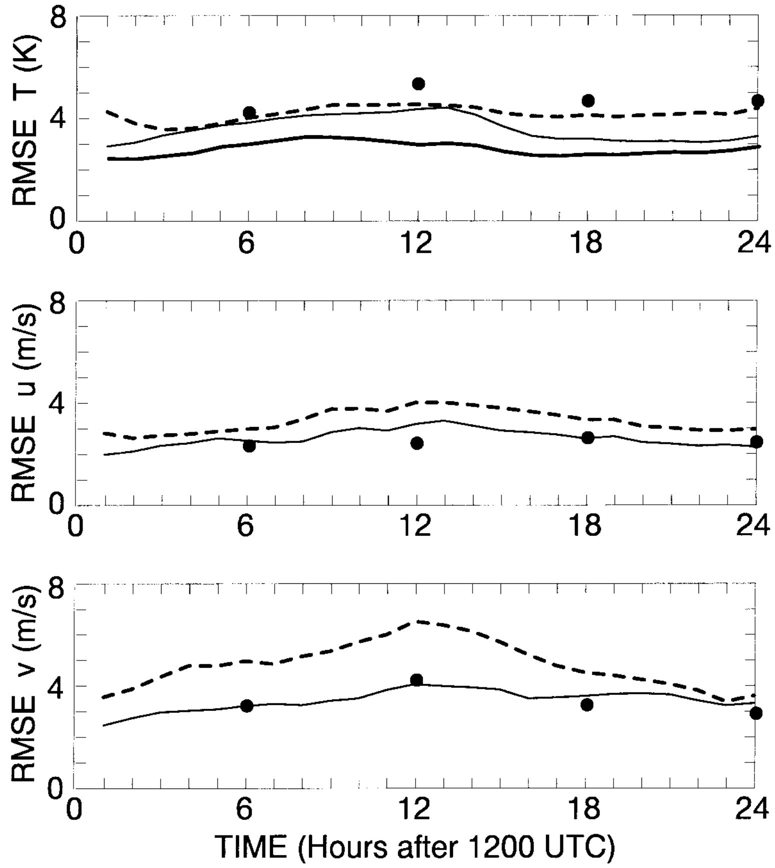

Fig. 8. Root-mean-square errors (RMSE) derived from all forecasts for temperature (top), zonal wind (middle) and meridional wind (bottom). Dashed lines represent persistence, thin solid lines represent MM5 forecasts and heavy solid line represents MM5 RMSE with bias (Fig. 7) removed. Black circles denote Eta model RMSEs.

model systematically performs better than a 24-h-lagged persistence forecast. For temperature we have displayed the resulting improvement if the time-mean bias from the model is removed. In this case, temperature RMSDs are 1–2°C smaller than those from persistence. The largest improvement of any variable relative to persistence is in the meridional wind component, especially during the afternoon, when the model RMSD is only about 60% of the RMSD for persistence.

In addition to lagged persistence forecasts, we computed biases and RMSDs for all Eta forecasts initialized at 1200 UTC (Figs. 7, 8). Note that the boundary and initial conditions for MM5 forecasts used the previous (0000 UTC) Eta forecast as a first guess, so this comparison may favor Eta slightly. The collection of Eta forecasts (about 100) did not correspond exactly to dates of the MM5 forecasts, but it is unknown how this may have affected the results. Because of the coarse resolution grid (80 km) on which Eta was received, verification of Eta forecasts was restricted to the grid point nearest the center of DPG. Hence, no information about local circulations over the test range can be discerned from Eta. The comparison of Eta with MM5 therefore mainly concerns prediction of synoptic-scale features.

In general, wind biases for Eta were comparable in magnitude to those of MM5. Eta showed a large cold bias in temperature, also noted by Cook (1998) in a separate verification of Eta forecasts centered near Salt Lake City. The strong cold bias in Eta is probably responsible for part of the cold bias in MM5. RMSD values for Eta were comparable to those of MM5, except for temperature, whose larger RMSD is a result of the large bias.

The most notable feature of Fig. 8 is the fact that model errors do not grow systematically with time. Model errors do grow for about the first 12 h, but then remain nearly constant or actually decay. We speculate that the reason for this performance is the dominance of local forcing and the decoupling of the boundary layer at night. While errors at upper levels may grow systematically with time, the behavior at the surface is, in effect, shielded from the behavior aloft. This interpretation is supported by Cook (1998) who noted that errors at upper levels (above the boundary layer) grew monotonically, whereas errors at the surface decreased at night. We speculate that the reason for the behavior of the persistence error, which one might expect to be nearly constant with time, is that vertical mixing occurs in the afternoon when there is the strongest coupling between the surface and upper air. That is, the largest errors in persistence appear to be due to its failure (by definition) to predict regional weather changes. The strong vertical coupling during the afternoon implies that these changes are felt at the surface, but the subsequent decoupling at night means that the surface is again shielded from these changes.

We note that the use of the model is an improvement over lagged persistence during all seasons. Hence, model performance relative to persistence is not simply related to the frequency and intensity of the passage of synoptic-scale systems. For example, the performance in July 1998 was similar to that for the whole sample, in spite of the fact that this month featured a large number of clear days with weak mean flow and few passages of synoptic-scale systems. Thus, while a significant fraction of the MM5 improvement relative to persistence can be attributed to synoptic- scale changes, skill during all seasons suggests that the model is useful in predicting day-to-day changes on the mesoscale during the warm season.

3.2.2. Time averages. The RMSD statistics indicate that, on average, the model provides a reasonable representation of the diurnal cycle. However, a more stringent test of model physics is whether it can capture systematic local circulations. These are investigated in this section by comparing the horizontal variation of time averages obtained from observations and model forecasts. We note that there is no point to a comparison with lagged persistence. The bias of persistence is simply the difference between the final time and the initial time of the averaging period, divided by the length of the period. Hence, when averaging over a season, only the seasonal trend appears.

A technical issue in performing a time average is related to missing observations and forecasts. In this study, averaging was performed only for times when both a forecast and observation were available. At any given time, data from some, but not all, observing platforms may be missing. In order to not bias the spatial distribution, we required at least half of the stations to report at any given time. Otherwise, the observations and model output for that time were not included in the average.

For brevity, statistics are only presented for summer (June and July 1998); however, many of the results seen here carry over to other seasons. The exception is winter, when the presence or absence of snow cover strongly alters the local circulations seen at other times of the year. To succinctly summarize the behavior over the diurnal cycle, means for 2200 UTC (1500 MST) and 1200 UTC (0500 MST) are presented (Figs. 9, 10, respectively), corresponding roughly to the times of maximum and minimum temperature.

For the afternoon, Fig. 9 shows that the salt-breeze and the mountain-valley circulation in the model reinforce to produce northwesterly flow over much of the test range. The maximum speed is about 3 m/s. A mean divergence is apparent over the northern part of the mesonet, especially when the light mean wind at the northernmost SAMS is considered. There is a strong, systematic temperature contrast in the model between the cooler playa to the northwest and the desert to the southeast. The observations generally do not support the strength of the northwesterlies predicted by the model. This bias is consistent with that seen in Fig. 7. However, the mean difluence over the northern part of the range is still present in the observations, although it is weaker than predicted. There is also a mean confluence over the eastern part of the range which is present, though weaker, in the MM5 forecast. The observed temperatures indicate a much weaker thermal contrast between desert and playa at this time of day than predicted. To the extent that a cooler playa will result in an enhanced thermally

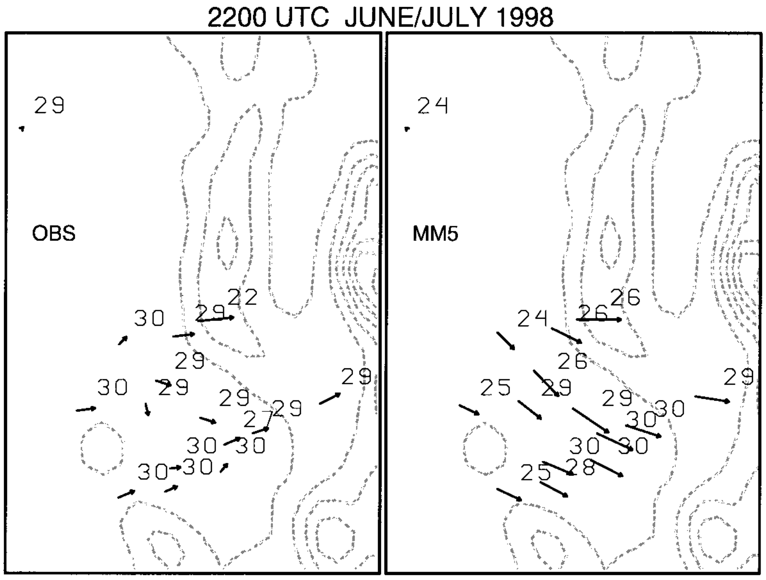

Fig. 9. Mean wind and temperature (°C) for summer (June and July 1998) for 2200 UTC (1500 MST), a 10-h forecast for MM5.

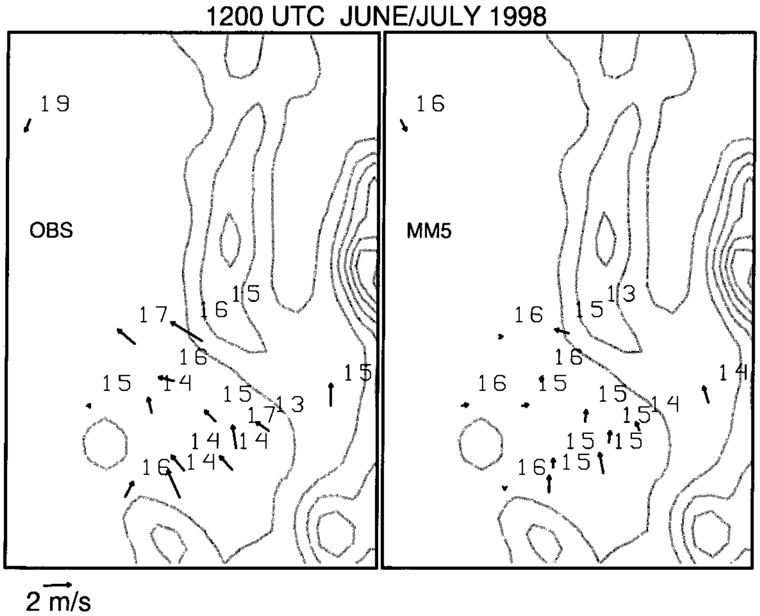

Fig. 10. As in Fig. 9, but for 1200 UTC (0500 MST), a 24-h forecast for MM5.

direct mean circulation, the northwesterly flow bias in the model is consistent with its accentuated cooling (or a lack of warming) over the playa.

In the early morning (Fig. 10), MM5 predicts weak southerlies over much of the range. The temperature variation between playa and desert has reversed, with the playa now being about 2°C warmer than the desert. Observations generally support this trend of a reversal in the flow over the range and in the temperature contrast between desert and playa. However, there is more of an easterly flow component in the observations and there is a stronger temperature contrast than MM5 predicts. At all times of day, MM5 exhibits a cold bias over the playa, and this bias contributes to the range-average cold bias noted in Fig. 7 in addition to the bias forced by the Eta model first guess. Note that the stronger southeasterly flow in the observations indicates a thermally direct circulation consistent with the larger observed temperature gradient directed toward the playa.

From these results, we note that MM5 predicts the general sense of both the mountain-valley circulation and the reversal in temperature gradient between desert and playa that occur over the course of the diurnal cycle. However, it is apparent that some of the parameters which govern the surface heat balance over the playa are in error, or the surface-energy-budget physics or PBL parameterization contain significant error. The albedo, moisture availability and thermal inertia are all possible candidates to be examined, and upon which further testing should be performed.

3.2.3. Anomaly correlations. To motivate the discussion in this section we return briefly to the salt-breeze event of 9 September 1997. A time series from SAMS-9 (location in Fig. 2) and from the model output interpolated to the location of SAMS-9 indicates that the model forecast of winds during the day and early evening captured the major wind shift (Fig. 11). However, during the night, an even stronger wind shift occurred at SAMS-9 which was not forecast by the model. In this case the model predicted southerlies across the range, but in reality the winds were quite variable. This kind of variability, especially at night, is a major concern to forecasters. It raises the question of how well this variability might be predicted by MM5.

A simple tool to examine the predictability of nonpersistent, range-scale circulations is an anomaly correlation. To calculate an anomaly at a given time for a given station, we first compute the deviation of the observed variable from its time mean for a given hour of the day. This effectively removes the systematic range scale circulation. Then the range-average of this temporal deviation is computed and the anomaly is defined as the departure from the range average. The same is done with MM5 forecasts. In this case, the time and range averages will differ between the observations and MM5, owing to the biases in the model discussed previously. We obtain anomalies for

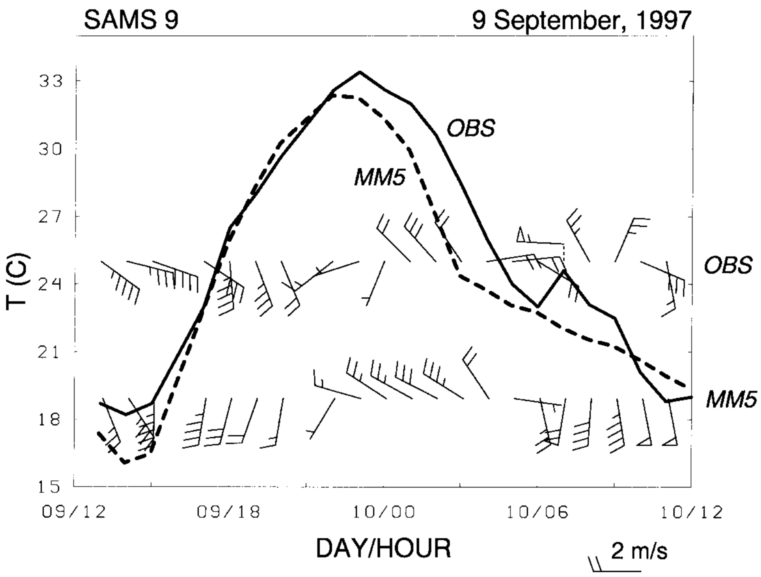

Fig. 11. Time series of wind (long barb = 1 m/s) and temperature (°C) for 9 September 1997.

each variable at each station and at each time. A simple correlation is computed between forecasted and observed anomalies, and the correlation coefficient for the entire sample of forecasts is then calculated.

Table 2 displays the correlation information. Nominally, correlations of 0.6 or higher are deemed significant or "useful". It is apparent that model and persistence forecasts each perform similarly, and poorly. Also of interest in Table 2 are the RMS values for the anomalies themselves. First, at least for winds, the magnitude of the observed RMS values is smaller, but not negligible, incomparison to the RMS values for the range mean (i.e., the mesoscale and synoptic-scale day-to-day variability). Hence, the local, nonpersistent fluctuations constitute a nontrivial portion of the total wind variability. However, the day-to-day fluctuations in range-mean temperature are much larger than the non-recurring anomalies of temperature. Second, the model RMS values are consistently 1.5 to 2 times smaller than those for the observations. While not apparent in the time averages, it is clear that at any given instant, MM5 fails to produce the proper variance among the mesonet stations.

Table 2. *Columns from left to right; anomaly correlation for 24-h persistence; anomaly correlation for MM5 forecasts; standard deviation of anomalies from observations; standard deviation of anomalies for MM5; units for standard deviations are in °C for temperature (T) and m/s for wind (U and V, the zonal and meridional wind components, respectively)*

	An. cor. persist	An. cor. MM5	SD OBS	SD MM5
Fall 97				
T	0.307	0.340	1.690	0.887
U	0.018	0.178	1.507	0.927
V	0.064	0.147	1.745	1.118
Winter 97				
T	0.386	0.321	1.910	1.089
U	0.034	0.106	1.352	0.772
V	0.176	−0.038	1.859	1.113
Spring 98				
T	0.273	0.125	1.345	0.923
U	0.029	0.065	1.676	1.001
V	0.073	0.114	2.011	1.137
Summer 98				
T	0.200	0.106	1.330	1.076
U	0.032	0.097	1.700	0.907
V	0.062	0.062	1.887	1.053

We can further investigate the skill of persistence forecasts in predicting the range-scale fluctuations discussed above by considering forecasts of different time lags (Fig. 12). If we adopt a "usefulness" threshold of 0.6, persistence is a useful forecast for not more than 45 min to 1 h, depending on the variable and season. The exponential decay of persistence forecast skill with time shown in Fig. 12 is also evident in Anthes and Baumhefner (1984), who focused on synoptic-scale forecast errors. If, based on synoptic-scale results, we postulate that a dynamical model can extend predictability by a factor of 2–3 over persistence, we find that there is little hope of predicting range-scale, non-recurring circulations with a lead time of more than about 2 h.

If we assume that the fluctuations dominating the statistics have length scales comparable to about half the scale of the test range (about 30 km), we find ourselves in a predictability/spatial- scale portion of parameter space broadly consistent with that discussed by Lorenz (1969) in which the predictability of scales of a few tens of km is of order an hour or two (see his Fig. 2). This represents a clear separation between the recurring, locally forced circulations, predictable on synoptic time scales, and the nonrecurring fluctuations, predictable only on short time scales. Because the amplitude of the non-recurring circulations is at least as large as that for the systematic circulations over the test range, the notion of greatly enhanced predictability near complex terrain is not accurate.

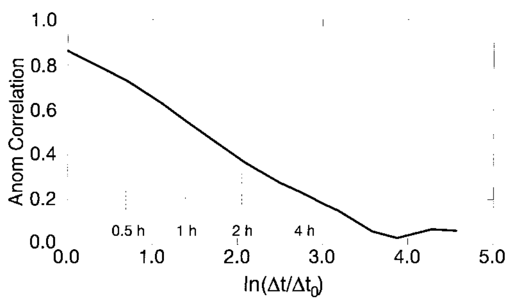

Fig. 12. Anomaly correlation (defined in text) as a function of the natural log of the normalized forecast length for time-lagged persistence. Here, $\Delta t_0 = 0.25$ h.

4. Conclusions

In the present paper, we have shown results from an operational forecast system using the MM5 model with horizontal grid increments as fine as 1.1 km. The particular application of the model was for producing forecasts over Dugway Proving Ground in west-central Utah, a desert region with complex terrain. The primary concerns for this paper were the model's ability to represent near-surface circulations in this environment and the extent to which the dominance of local forcing enhanced predictability on the meso-gamma scale (a few km to a few tens of km).

After examining more than 170 forecasts over 4 seasons and utilizing high-resolution mesonet data for model verification, we can make the following statements about model performance.

- MM5 was able to capture some important aspects of the recurring range-scale circulations in forecasts out to 24 h. Systematic errors in the model appear related to the treatment of the surface energy balance over the salt flats (playa).
- MM5 performed better than a forecast based on a 24-h-lagged persistence. The persistence forecast of this type typically performs well near complex terrain, especially in the absence of strong synoptic-scale systems. MM5 showed improvement over persistence during all 4 seasons. MM5 errors for temperature were smaller than those of the Eta model. The two models exhibited comparable wind errors.
- MM5 was unable to predict the non-recurring range-scale circulations. These have inherently short time scales and are predictable (anomaly correlation greater than 0.6) by a persistence forecast of not more than 1 h.

A noteworthy facet of the behavior of MM5 errors is that they did not grow monotonically with foreast time. Errors peaked during the late afternoon and evening and then decreased during the night. We speculate that this was a result of boundary-layer decoupling, which allowed the local, repeatable forcing to dominate, and effectively shielded the near-surface layer from errors aloft, which presumably were growing with time. During the daytime, when errors were growing with time, the time scale for error doubling was at least 24–36 h, a characteristic synoptic time scale.

We estimate that the nonrecurring range-scale circulations are characterized by a predictability of not more than about 2 h. Because the model was initialized with coarse resolution data, the error associated with these nonrecurring features was essentially saturated by the time the model spun up its fine-scale detail. Hence, the growth of errors associated with these features does not appear in the RMSD statistics.

In the near term, similar forecast/verification exercises will be conducted at other TECOM test ranges. The particular forecast challenges at each range differ, but the focus on the interaction of synoptic-scale and mesoscale transients with local forcing will be a common element. We plan to consider dynamic initialization in order to begin predicting some of the high-frequency, range- scale circulations. This will require a different forecast strategy with an analysis/forecast cycle of comparably high frequency. We anticipate that, with better depiction of the non-recurring local- scale circulations, MM5 can be used successfully as a driver for other models, such as to predict the path of debris resulting from a test or to predict the propagation of sound or the dispersion of a smoke plume.

5. Acknowledgments

This research was funded by the US Army Test and Evaluation Command through an Interagency Agreement with the National Science Foundation.

REFERENCES

Anthes, R. A. and Baumhefner, D. P. 1984. A diagram depicting forecast skill and predictability. *Bull Amer. Meteor. Soc.* **65**, 701–703.

Black, T. 1994. The new NMC mesoscale Eta model: description and forecast examples. *Wea. Forecasting* **9**, 265–278.

Colle, B. A. and Mass, C. F. 1996. An observations and modeling study of the interaction of low-level southwesterly flow with the Olympic Mountains during COAST IOP4. *Mon. Wea. Rev.* **124**, 2152–2175.

Cook, L. K. 1998. *An evaluation of mesoscale model performance over the Intermountain Region of the US.* MS Thesis, University of Utah, Salt Lake City, UT. 78 pp.

Dudhia, J. 1989. Numerical study of convection observed during the winter monsoon experiment using a mesoscale two-dimensional model. *J. Atmos. Sci.* **46**, 3077–3107.

Grell, G. A, Dudhia, J. and Stauffer, D. R. 1994. *A description of the 5th-generation Penn State/NCAR mesoscale model (MM5)*. NCAR Technical note TN-398+STR. 121 pp. Available from corresponding author at NCAR, PO Box 3000, Boulder, Colorado, 80307.

Grell, G. 1993. Prognostic evaluation of assumptions used by cumulus parameterizations. *Mon. Wea. Rev.* **121**, 764–787.

Hong, S.-Y. and Pan, H.-L. 1996. Nocturnal boundary layer vertical diffusion in a medium-range forecast model. *Mon. Wea. Rev.* **124**, 2322–2339.

Klemp, J. B. and Lilly, D. K. 1978. Numerical simulation of hydrostatic mountain waves. *J. Atmos. Sci.* **35**, 78–107.

Lorenz, E. N. 1969. The predictability of a flow which posesses many scales of motion. *Tellus* **21**, 278–307.

Loveland, T. R., Merchant, J. W., Brown, J. F., Ohlen, D. O., Reed, B. C., Olson, P. and Hutchinson, J. 1995. Seasonal land-cover regions of the United States. *Annals of the Association of American Geographers* **85**, 339–355.

Manning, K. W. and Davis, C. A. 1997. Verification and sensitivity experiments for the WISP94 MM5 forecasts. *Wea. Forecasting* **12**, 719–735.

NCAR staff 1996. Technical approach to developing 4D weather prototype for TECOM. *Final report: NCAR evaluation of TECOM meteorological support needs.* UCAR, Boulder, CO. Available from corresponding author at NCAR, PO Box 3000, Boulder, Colorado, 80307.

Orlanski, I. 1975. A rational subdivision of scales for atmospheric processes. *Bull. Amer. Meteor. Soc.* **56**, 527–530.

Physick, W. L. and Tapper, N. J. 1990. A numerical study of circulations induced by a dry salt lake. *Mon. Wea. Rev.* **118**, 1029–1042.

Shafer, M. A. and Hughes, T. W. 1996. Automated quality assurance of data from the Oklahoma mesonetwork. *Preprints, 12th International Conference on Interactive information and processing systems for meteorology, oceanography and hydrology*. Atlanta, GA. Available from the American Meteorological Society, 45 Beacon Street, Boston Massachusetts 02108.

Ulaby, F. T., Dellwig, L. F. and Schmugge, T. 1975. Satellite microwave observations of the Utah Great Salt Lake Desert. *Radio Science*, 947–963.

Vislocky. R. L. and Fritsch, J. M. 1997. An automated, observations-based system for short-term prediction of ceiling and visibility. *Wea. Forecasting* **12**, 31–43.

Weingarten, D., Feckter, J., Clough, C., Villarreal, A. and Jameson, T. 1996. Range meteorological problems and issues. *Enhancing US Army Range meteorological capabilities* (ed. Shantz, A. A.). UCAR, Boulder, CO. Available from corresponding author at NCAR, PO Box 3000, Boulder, Colorado, 80307.