Error Evaluation of NCEP and LAHM Regional Model Daily Forecasts over Southern South America

A. CELESTE SAULO, MARCELO SELUCHI,* CLAUDIA CAMPETELLA, AND LORENA FERREIRA

Departamento de Ciencias de la Atmósfera–FCEyN–UBA, Centro de Investigaciones del Mar y la Atmósfera, UBA/CONICET, Buenos Aires, Argentina

ABSTRACT

NCEP short-range operational forecast and Limited Area HIBU (Federal Hydrometeorological Institute and Belgrade University) Model (LAHM) regional model performance during a 2-month period over the southern part of South America are evaluated through the analysis of bias and rmse's. While spatial structure of errors could be only examined using gridded operational analyses as the ''ground truth,'' observed data have been used at two radiosonde stations to have an independent control of forecast and analysis quality. LAHM precipitation forecast error has been also determined using observed 24-h accumulated precipitation over a subregion of interest.

Bias and rmse are, in general, lower for Medium-Range Forecast Model (MRF) 24-h forecasts than for the regional model, though MRF errors appear to be larger than those reported by other studies carried out over the whole Southern Hemisphere, suggesting the necessity to perform regional verification analysis whenever gridded analyses and/or forecasts are being used. This recommendation particularly holds over data-void regions like South America.

While geopotential and wind biases do not exhibit a particular pattern in either forecast, there is a clear tendency to cold biases over the whole troposphere, and for the MRF in particular, growing with height.

The results obtained from LAHM evaluation suggest that continuous development is needed to keep this regional forecast system as a plausible counterpart of available global model products for fulfillment of local requirements.

1. Introduction

There is a growing tendency for the use of global numerical weather predictions (NWP) available almost in real time via the Internet. This option provides a powerful tool to forecasters all over the world, and is also very useful for the provision of initial and boundary conditions to higher-resolution regional models. In this way, regional forecasts can be more easily issued to fulfill local requirements, giving rise to a variety of local NWP efforts, first documented over the United States and Europe, but rapidly spreading over the world.

Many related topics have appeared since those pioneering works by Warner and Seaman (1990), Cotton et al. (1994) and others, devoted to real-time mesoscale prediction over the United States. Among these issues, the verification problem (e.g., Mass and Kuo 1998) is critical when running models with higher resolution than

that available in observational datasets (Davies and Carr 2000). While some of these experiences can be applied to other regions in the world, it must be kept in mind that models do not perform equally well in different regions: lack of data to correctly define initial conditions and complex terrain mainly account for this limitation. Moreover, verification should be treated cautiously, considering local data availability and/or reliability.

This issue points to a critical problem affecting South America, a region characterized by geographical singularities (the Andes, the Amazonia, the relatively narrow continental region) and a serious lack of observations (at 0000 UTC, fewer than five radiosondes are launched operationally south of 20°S). These particular conditions lead to verification restrictions even on scales larger than the mesoscale. The peculiarities of the South American region have not been considered in detail by research and operational centers when performing error evaluation: while some reports showing short- to medium-range global model forecast performance over the Southern Hemisphere [i.e., the Working Group on Numerical Experimentation of 1998), or the National Centers for Environmental Prediction (NCEP) global branch Web page] can be found, no particular studies have been carried out documenting operational NWP performance

---

over South America, at least to the authors' knowledge. However, NCEP and European Centre for Medium-Range Weather Forecasts (ECMWF) products (forecasts and analyses) are widely used, since their availability constitutes a unique tool for research and forecast activities.

NWP quality evaluation determines the confidence related to a particular forecast over a specific region, but it can also be thought of as a preliminary evaluation of simulated climate drift that can be expected from the model being tested. This idea is further illustrated by Kamga et al. (2000) who analyzed the systematic errors of the ECMWF operational model (120-h forecast) over tropical Africa. From this point of view, verification of model products can serve local forecast applications and a wider community that conducts research on a variety of timescales using global model forecasts and/or analyses. This last group of users, in the particular case of South American–related topics, has rapidly grown in recent years, as suggested by the amount of papers devoted to the description of its climate, climate variability, and related features such as the South Atlantic convergence zone, the South Pacific convergence zone, and the South America low-level jet (SALLJ), among others. Most of the above-mentioned research has been carried out using analyzed datasets, such as NCEP and ECMWF reanalysis, and some others using global models. Clearly regional-scale verification is of primary interest.

The present work is aimed at evaluating NCEP forecast and analysis system performance over the southern part of South America, during a 2-month period, as well as that of a regional model [LAHM, the Geophysical Fluid Dynamics Laboratory/Limited Area HIBU (Federal Hydrometeorological Institute and Belgrade University) Model, adapted at the Centro de Investigaciones del Mar y Atmósfera], which is experimentally being tested for operational use utilizing Medium-Range Forecast (MRF) global spectral model forecasts for initial and boundary conditions. It is believed that this work will provide some indications of quality for NCEP products that can also aid in estimating their capability to describe mean weather. In addition, it can provide guidelines for the development of regional models that should be used for higher-resolution predictions over the region of interest. It should be noted that the period is not long enough to provide statistical reliability to the results. However, it is still useful to describe main successes and failures in current forecasts over the region and to disclose future lines of work.

This work is organized as follows: section 2 briefly describes the model's main properties, the datasets, and the methodology. Section 3 presents the characteristics of the period under study, through mean analyzed and forecast fields obtained from both models and the corresponding mean errors. Section 4 focuses on errors over a particular area, linked to the Rio de la Plata basin. Section 5 shows performance against observed data,

including precipitation, and the final section summarizes the results and discusses future lines of work.

2. Data, models, and methodology

As part of a program focused on the evaluation of NWP quality over Argentina, and in the operational use of a regional model formerly used for research purposes, a regional forecast system has been experimentally set up since 1998. This system is based upon the LAHM regional model and uses daily NCEP 0000 UTC analyses and the aviation run of the MRF global spectral model for up to 3 days as initial and boundary conditions, respectively. Model characteristics (triangular truncation of T126 and 28 vertical levels) and performance evaluation of the MRF model can be found in Kalnay et al. (1998) and at NCEP Environmental Modeling Center (EMC) Web pages (http://www.emc.ncep.noaa.gov/modelinfo/index.html).

The complete set of variables, interpolated into a $1° \times 1°$ horizontal grid, and at 26 standard pressure levels, for the initial time and subsequent 12-h forecast intervals are available at around 0300 local time, and the regional model forecast is ready around 0600 local time. One strong limitation of the proposed forecast system is that only the 0000 UTC cycle is plausible because there must be enough time between the initialization procedure and the resulting regional forecast to finish the forecasts within a reasonable time period. The initialization is accordingly done with the 0000 UTC analysis, which lacks upper air data particularly over Argentina (see Fig. 1 for a reference of data used by the analysis procedure at NCEP during a day within the period of study—large boxes highlight upper air data available at the model domain).

The hydrostatic model used in this study (LAHM) is a regional model that allows both the horizontal resolution and domain to be arbitrarily chosen, with 18 fixed-sigma levels in the vertical. Its numerical formulation and parameterization of radiative and boundary layer processes are similar to those described in Orlanski and Katzfey (1987). Surface drag is formulated via a Richardson number–dependent Monin–Obukhov (1954) scheme, and vertical diffusion is treated by a first-order closure scheme with eddy diffusivity coefficients as functions of Richardson number (Nicolini and Saulo 1995). Radiation follows the Fels–Schwarzkopf (1975) algorithm and the diurnal cycle is fully considered except for variations of the sun's azimuthal angle. Moist convective parameterization is currently introduced using the cumulus cloud ensemble model developed by Arakawa and Schubert (1974). Surface hydrology is taken into account by a simple bucket scheme. Surface temperature is evaluated based on the surface energy balance equation.

For this experimental design, the selected domain encompasses an area from 65°S, 108°W to 19°S, 30°W, covering around 2.8×10^5 km^2 with a horizontal grid
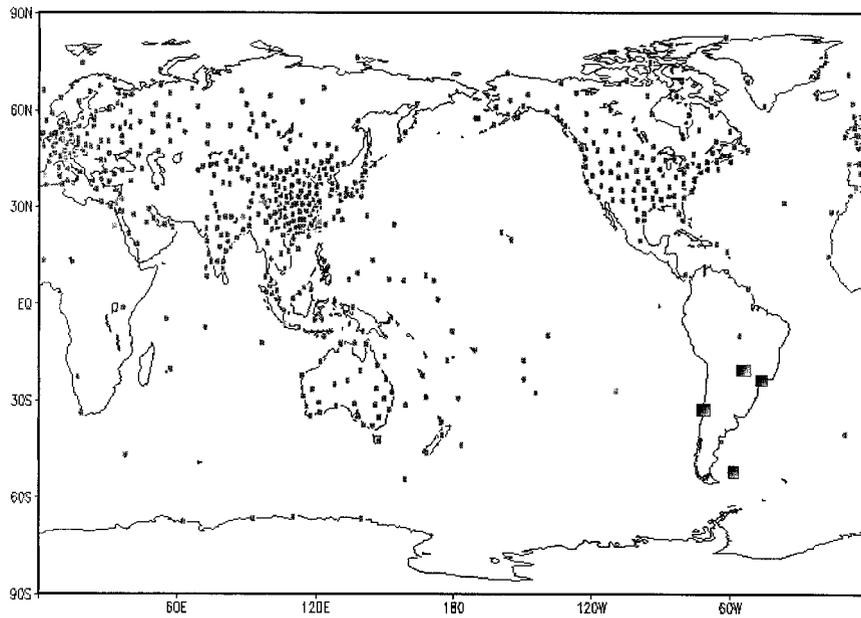
FIG. 1. Height data coverage from radiosondes at 700–300 hPa, valid for 26 Apr 1999. Radiosonde stations over the region of interest are highlighted (large boxes). (Obtained from the experimental Web page of S. Saha, NCEP/EMC.)

interval of 0.85° in both directions and 18 vertical levels (Fig. 2). These settings have been defined after several runs and are designed to achieve a reasonable resolution and areal coverage while keeping computational costs bounded. Model performance has been evaluated through a wide set of simulations, which covered many of the most frequent synoptic situations over Argentina and surrounding areas (Orlanski et al. 1991; Menendez

1994; Nicolini and Saulo 1995; Seluchi and Saulo 1998, among others). This same model has also been applied to regional wintertime climate simulations over South America (Menendez et al. 2000). These experiments allowed an improvement of model performance by modifying its physical parameterizations and tuning the coefficients.

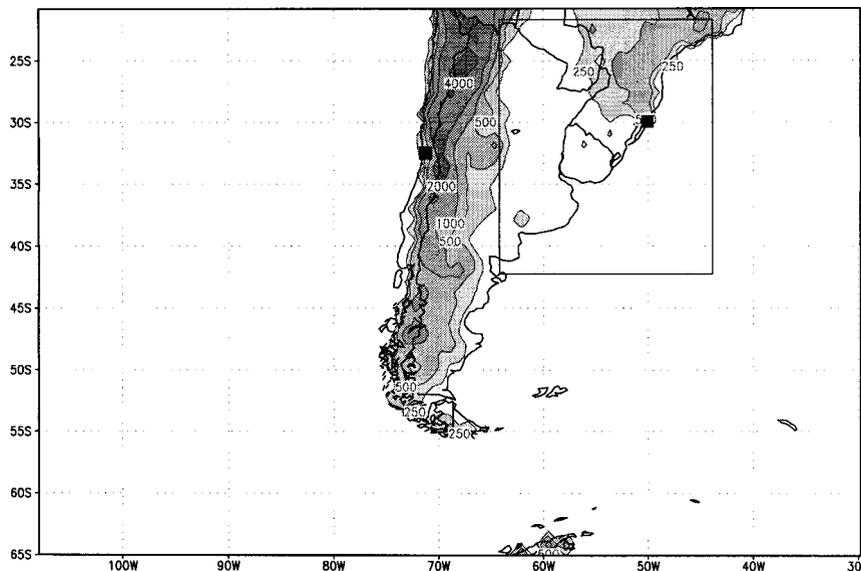As mentioned earlier, this forecast cycle has been used

FIG. 2. Model domain and topography (in m). The rectangle denotes a subdomain where mean errors where calculated. Quintero (Chile) and Porto Allegre (Brazil) stations mentioned in the text are marked with black boxes.

since August 1998. Nevertheless, there are many difficulties associated with its daily use. The most frequent problem arises from the data ingestion, which takes a long time to be completed because of the extremely slow local connection to the Internet. Also, on occasions, there have been identified errors in the transmitted MRF forecast fields, which corrupted local model runs. As a consequence, in 1999, the longest consecutive period available to perform this verification study was from 1 April to 31 May 1999.

During this period, LAHM (each 12 h) and MRF (24, 48, and 72 h) forecasts were archived and interpolated onto a common 0.85° grid, at seven standard pressure levels (1000, 850, 700, 500, 300, 200, and 150 hPa).

The 24-h accumulated precipitation forecasts during April–May 1999 were only accessible for LAHM, since daily observed precipitation is available at 1200 UTC and needs 12–36-h forecasts for comparison, which were not routinely archived for MRF products. In order to provide an idea of how the LAHM and MRF 12–36-h precipitation forecasts compare, an additional period has been included. This period encompasses three months, from October to December 2000.

Verification statistics used include root-mean-square error (rmse) and bias error (bias) that, for a given variable x, are defined as

$$\text{rmse}(x) = \left[\frac{1}{N} \sum_{n=1}^{N} (x_n^f - x_n^o)^2 \right]^{1/2} \quad \text{and}$$

$$\text{bias}(x) = \frac{1}{N} \sum_{n=1}^{N} (x_n^f - x_n^o),$$

where N is the total number of forecasts and the superscripts f and o signify forecast and observed values, respectively. The equitable threat score (ETS) for precipitation is calculated as in Mesinger (1996). Also, mean fields and their mean variability are briefly discussed, in order to highlight the regional characteristics and the main features of the period.

In order to provide a comprehensive model evaluation, this work analyizes different scores and uses alternative datasets as ''ground truth.'' Spatial structure of mean errors is focused on 850-hPa charts using gridded analysis as the truth. Since there are no upper air soundings over Argentina at 0000 UTC (suitable for validating analyses and 24-, 48-, and 72-h (forecasts) and surface data are topography dependent and do not provide a broad measure of model behavior, temporal evolution has been analyzed through the day by day changes of areal averages at 850 and 200 hPa, calculated over the rectangle shown in Fig. 2. This area is closely linked to the Rio de la Plata basin, though it is somewhat smaller (because of the regional model northern boundary) than the one strictly defined as the basin by García and Vargas (1996) who confine its northern limit around 15°S.

All the statistics previously referred to were calculated using the NCEP 0000 UTC analysis as the truth. As stated earlier, the lack of upper air observations constrains the verification over the whole domain to a check against analyzed fields. The degree of agreement that may consequently be expected between the MRF forecast and analyses may be enhanced by use of the same model in the analyses and forecasts. A similar though less marked effect may influence the regional model, since it is initialized and bound by the same model–analysis pair.

To have at least one independent way of verification, two isolated stations were used to perform model-observed data comparisons: one over Brazil (Porto Alegre, at approximately 30°S, 51°W; 3 m above sea level) and the other over Chile (Quintero, at approximately 33°S, 71°W; 8 m above sea level). This assessment further complements the information of error vertical structure obtained from areal mean values.

Finally, 24-h accumulated precipitation data from a surface synoptic network has been used to validate 12–36-h precipitation forecasts.

3. Mean fields, mean variability, and spatial error structures

In this section a brief description of the period under study is presented using information available from the monthly issued *Climate Bulletin* of the National Meteorological Service. Also, mean fields and mean variability (the standard deviation of the variable from the period mean) are calculated using both gridded analyses and forecasts (MRF and LAHM) for reference (see Fig. 3). Error statistics presented in this section are calculated using analyzed fields as the truth.

Mean 850-hPa fields (from analyses and from both 24-h forecasts) included in Fig. 3 show no particular patterns during this period, except for the trough axis evident around 40° and 50°W, which is a consequence of the high frequency of low pressure systems that developed and moved to the Atlantic Ocean between 35° and 40°S.

Intercomparison of mean analyzed and forecasted fields shows that the MRF better depicts the location of the Pacific anticyclone, while LAHM situates it somewhat west of its position. Temperatures are colder over the continent in the LAHM forecast, an aspect that will be further discussed in subsequent sections, while both forecasts agree fairly well over oceanic regions. It seems that LAHM does a better job in the representation of moisture over southeastern coastal Brazil. The strong humidity gradient over this region indicates the transition zone between tropical air masses to the north and mean latitude ones to the south. At higher levels (not shown) the agreement between forecasts is higher for all the analyzed variables.

Geopotential variability is well represented in both forecasts, though they tend to exhibit greater variability at higher latitudes. In terms of forecast temperature var-
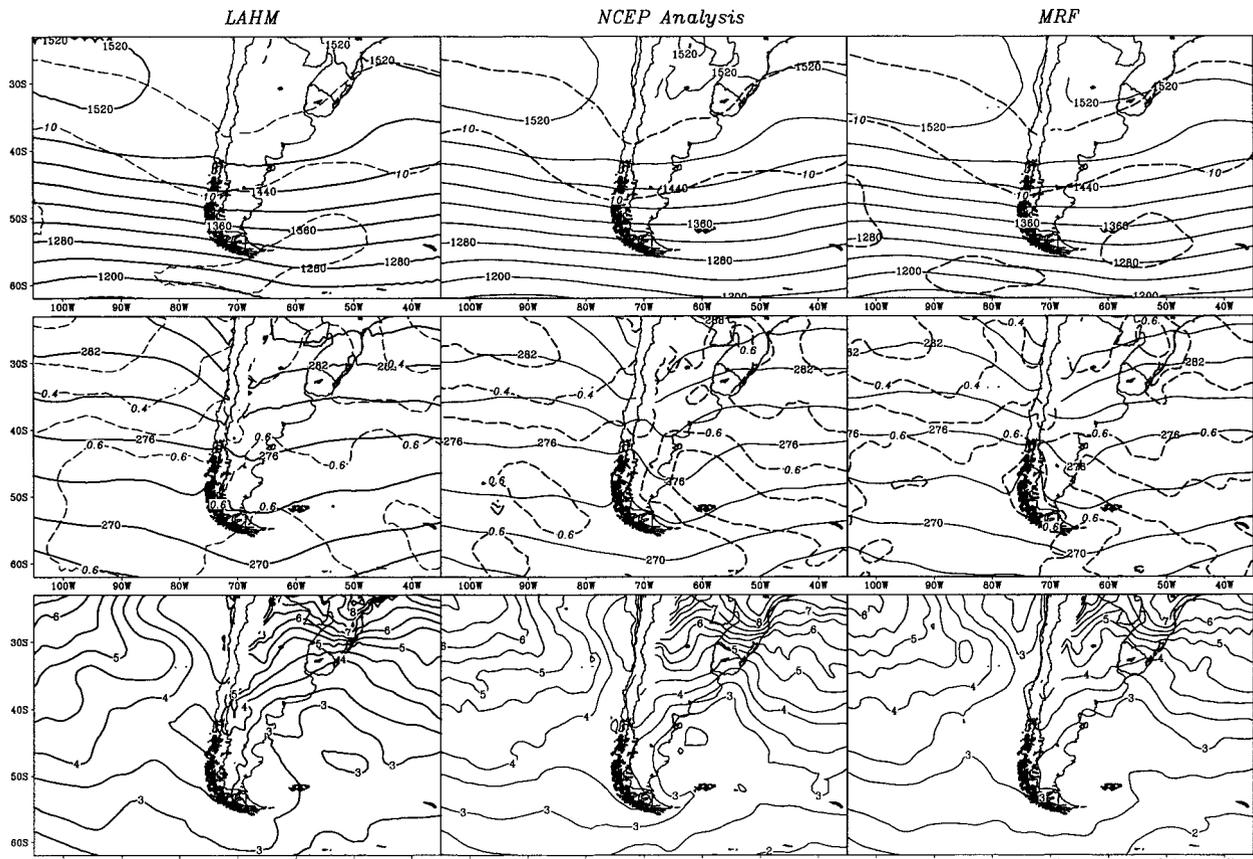
FIG. 3. (upper panels) Mean geopotential 850-hPa height (contours) and its mean variability (dashed) in m, for the 2-month period: (left) LAHM 24-h forecasts, (center) NCEP analysis, and (right) MRF 24-h forecasts. (middle panels) Same as upper panels, but for temperature in K. (lower panels) Same as above but for specific humidity in g kg^{-1}; moisture variability has not been included.

iability, there are no particular patterns to discuss except the tongue of relatively high temperature variability over northeastern Argentina, southern Brazil, and Uruguay that is well captured by both models and could be associated with the increased convective activity detected during April.

Observational data show that weather patterns exhibit different behavior in April and May in terms of variability when compared with their respective climatology. Weather in April denotes greater deviations, with mean surface temperatures between 1° and 2°C below its climatological mean north of 40°S, and slightly above it to the south. Nevertheless, during May these differences were much smaller (i.e., May was more "climatological") with a 1°C deficit over northeastern Argentina and a 1°C excess over its southern portion. Mean April precipitation was above normal over most of central Argentina, with excesses reaching 150 mm at particular areas. Most of the observed precipitation was associated with very active cold front passages and convection triggered by dynamical perturbations at upper levels. The above normal precipitation amounts during April were mainly linked to individual events that occurred between 24 and 27 April, with daily precipitation

peak values of 200 mm. Around mid-April, the first cold air irruption occurred, producing frosts over central and northern Argentina. This particular event, which led to frosts and snowfall over southern Brazil, was recently analyzed in detail by Satyamurty et al. (2001).

Rmse and bias scores for the whole domain and for both 24-h forecasts are presented in Fig. 4. In general, for all the fields there are lower rmse's for the MRF than for the LAHM. Geopotential heights show a similar spatial structure in rmse for both models with larger error over regions with higher variability, but exhibit a different pattern in bias: while LAHM tends to underestimate heights in the NW portion, mainly around the Andes, and to overestimate them in the SE part, MRF exhibits positive bias in the NE and a negative one in the SW. In both cases, this tendency increases with forecast length (not shown). LAHM systematic height errors around the Andes may be due to a different height interpolation approach over highlands, while numerical treatment of the western boundary could be responsible for error growth principally over the region where perturbations move faster to the east. The spatial distribution of bias in MRF suggests a tendency to overpredict the Atlantic anticyclone intensity and the low pres-
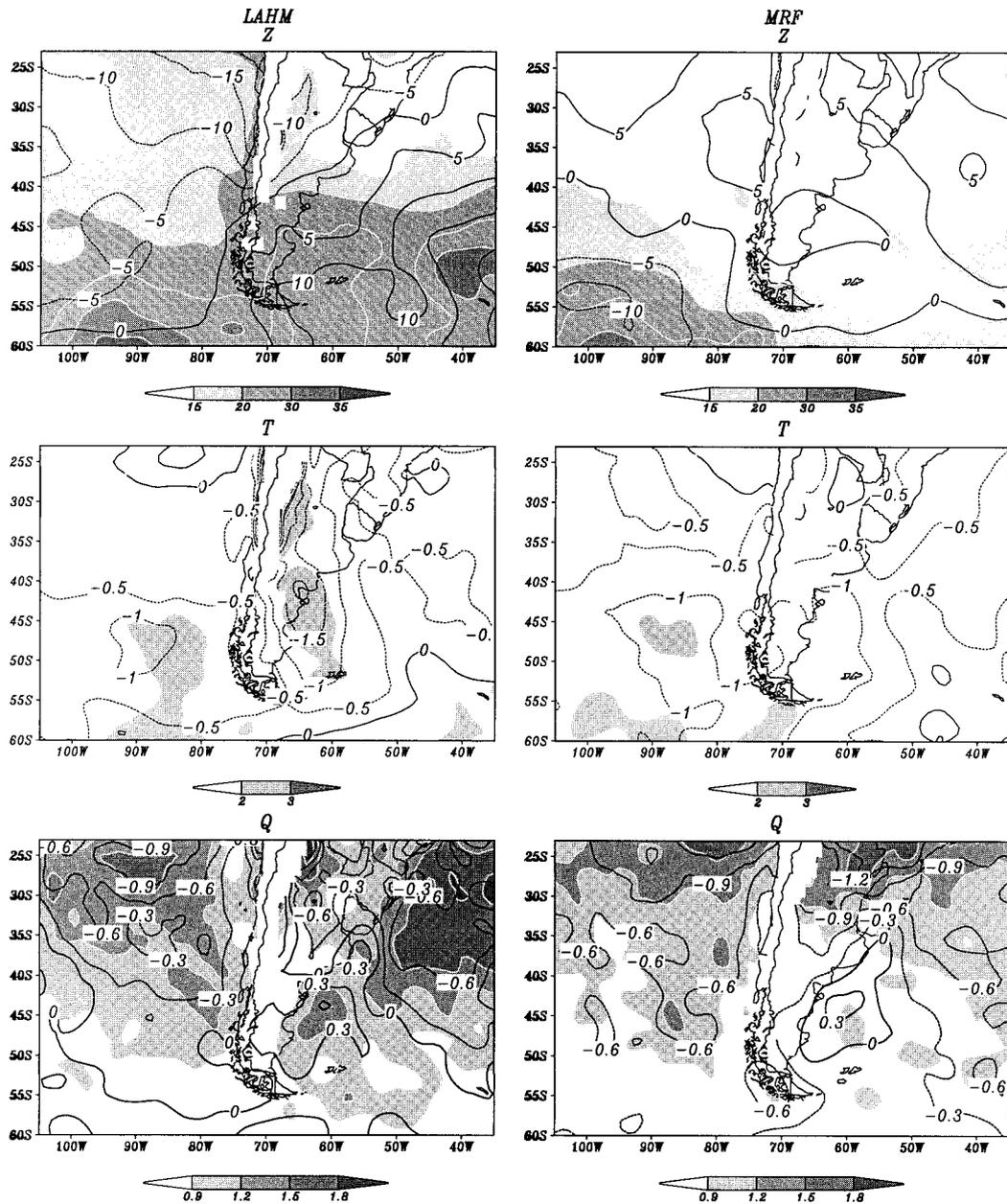
FIG. 4. (left panels) LAHM and (right panels) MRF 850-hPa 24-h forecast rmse (shaded) and bias (contoured) error for geopotencial height (m), temperature (K), and specific humidity (g kg⁻¹).

sure circumpolar trough, mainly upstream of the continent.

In temperature, a systematic error over continental regions at low levels is clearly seen in LAHM forecasts. The source of this error has been extensively discussed in Menéndez et al. (2000), and is linked to the simple bucket model used over the surface, which tends to misrepresent the sensible to latent heat flux ratio, particularly over the southern portion of South America. Errors are comparable in magnitude for both models over oceanic regions, and also display an increase with

forecast length. In the MRF, the cold bias is evident over the region south of 40°S, with peak values below −2°C in 48–72-h forecasts. Cold biases over the Southern Hemisphere lower and middle troposphere have also been identified by Kamga et al. (2000) for ECMWF 5-day forecasts over the southern portion of their domain (around 20°S), and by Moorthi (1997) for MRF January 1996 5-day forecasts. Specific humidity shows a dry bias over almost the whole domain, particularly for MRF. As it could be expected, rmse's are higher where moisture exhibits larger variability.

4. Error time series and vertical structure of errors

This section focuses on model evaluation over a box that is around the Rio de la Plata basin (see Fig. 2). As was mentioned in the introduction, the basin is more extended to the north; still, this box is representative of a region that is of critical economical and societal interest because of its great population density, agricultural activities, and hydroelectric energy production. From the weather–climate perspective, it is a transition zone between subtropical and midlatitude weather regimes. It is characterized by significant synoptic activity to the south, such as frontal passages (Taljaard 1968; Necco 1989), a high frequency of cyclogenesis (Seluchi and Saulo 1998; Gan and Rao 1991; Sinclair 1995, among others), and its large number of mesoscale convective complexes (Velasco and Fritsch 1987). It coincides with the exit region of the SALLJ (Saulo et al. 2000; Douglas et al. 1999, among others), which provides moist and warm air into the area and may act as an organizing mechanism for convection.

The problems arising from the lack of sufficient observations over this region have been pointed out by Wang and Paegle (1996), who found that uncertainty in the analyses over a similar box to the one used in this work resulted in opposite water budgets when calculated from different analyzed datasets (NCEP and U.K. Met Office). This prior work constitutes a warning on the reliability of the analyses and consequently on model outputs, and further motivates our interest in assessing models errors over this region.

Figure 5a shows area-averaged temperature (upper curves) and geopotential heights (lower ones) for each day of the period. Besides the fair agreement between both forecasts and the analysis, it is interesting to note the variability within the period. In order to analyze biases and error growth with forecast length, Figs. 5b and 5c have been included. Geopotential forecast biases remain bound around ±30 m, except for 27 May. The 850-hPa temperature bias clearly depicts a cold bias for both models—greater for the LAHM model and increasing in time for both.

Both models tend to produce bias of the same sign along this period except for particular situations. For example, the disagreement found between forecasts in the 19–24 April period can be explained by inspection of both forecasts: a low pressure system located around 30°–35°S, 55°W, formed around 17 April and remained quasi-stationary during this particular period. This low pressure system has been overestimated by LAHM, affecting pressure values in the western boundary of the selected subdomain. Figure 6, on the other hand, shows more disagreement between model behavior at upper levels, mainly in the temperature field where LAHM shows better performance and a relative warm bias, and MRF retains its cold bias (even more than at low levels).

In order to display the vertical structure of these er-

rors, rmse and bias for the whole period, which are averaged over the subregion of interest, are presented in Fig. 7. This figure shows that in general LAHM has weaker performance, but it does have a significant portion of its failure explained by systematic errors—with a tendency to underestimate the geopotential height in the middle troposphere, and a strong cold bias near the surface. Both models show good agreement with the analyzed wind field.

These scores confirm the previously mentioned handling of surface processes by LAHM for a cold bias at lower levels, while some problem affecting the tropopause temperature representation is evident from the MRF outputs. Geopotential height and wind fields seem to be less affected (at least compared to the analyses) by systematic errors.

5. Verification against observed data

a. Radiosonde data

In this section, emphasis is placed on comparison between analysis, model forecasts, and radiosonde data at two isolated points. In order to include the analyses used in the proposed forecast system, only 0000 UTC synoptic time has been evaluated. Two radiosonde stations located at regions of interest (Porto Allegre, Brazil, and Quintero, Chile; see Fig. 2 for their locations) were selected to perform this evaluation, for which analysis and model data were horizontally interpolated to each station location. Also, to provide a complementary analysis with respect to the previous sections, rmse was calculated using an alternate expression (which will be referred to as crmse):

$$\text{crmse}(x) = \left[\frac{1}{N} \sum_{n=1}^{N} (x_n'^f - x_n'^o)^2 \right]^{1/2},$$

where the prime stands for the deviation from the corresponding time mean. Then, the performance of these forecasts can be evaluated in a way that removes the systematic bias.

This kind of comparison is, obviously, restrictive from a model evaluation point of view. On the other hand, it does not give a complete idea of performance over the region of interest, since it evaluates single points that might be affected by local features. This is indeed the case for the selected stations, since both lie near the coast. However it should be stressed that this is the only way that an independent control can be done over an area lacking upper air observations.

Scores over Porto Allegre are shown in Fig. 8. Biases denote a tendency to lower heights, at least up to 200 hPa. Surprisingly, MRF does better than the analysis over most of the troposphere, but this can be partially explained by its positive bias (compared with its own analysis) over this region, which counteracts the negative bias with respect to data at Porto Allegre. When looking at temperature error profiles, the analysis dis-

a. 24 hr forecast

b. 850 hPa height forecast bias

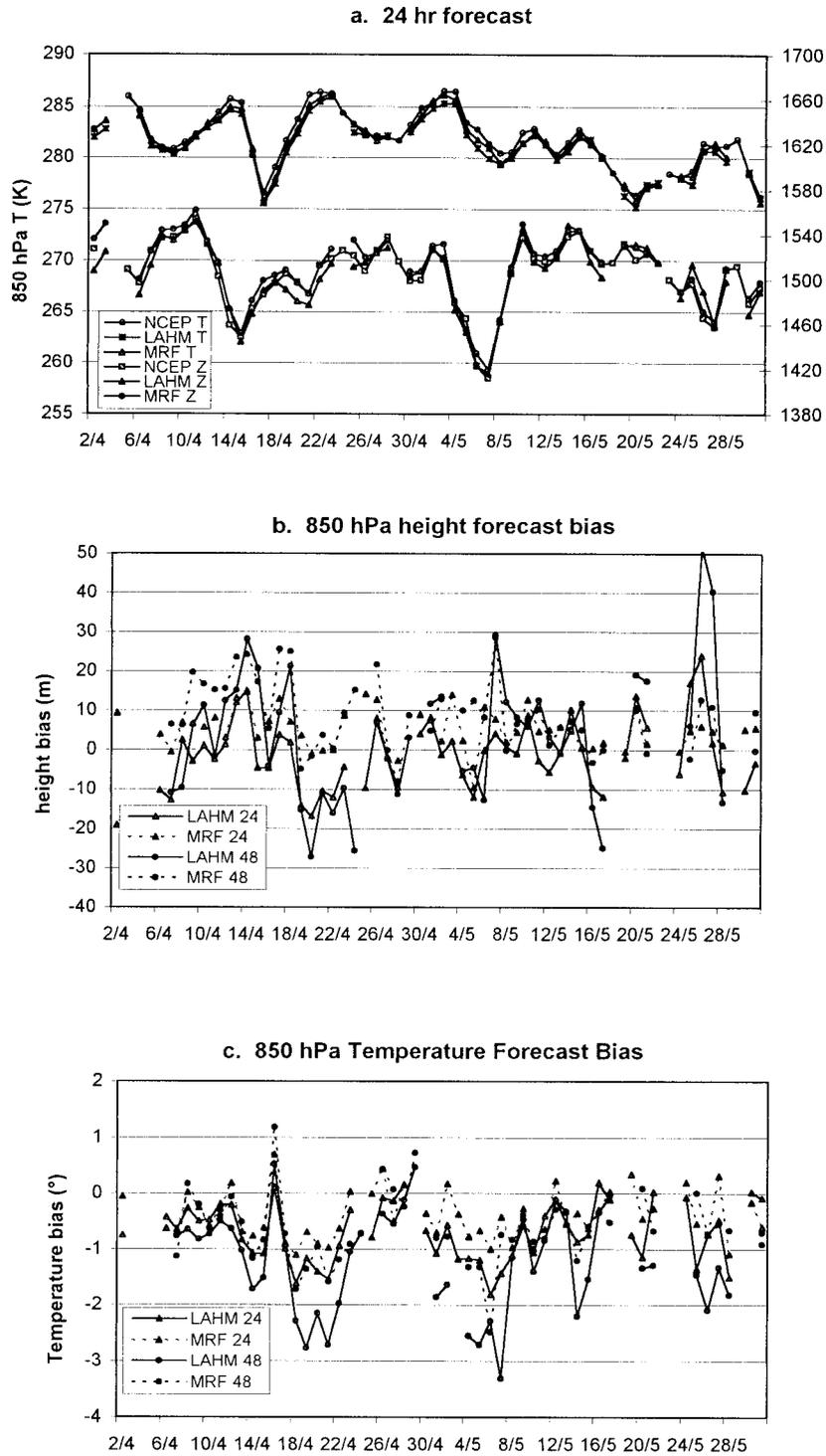## c. 850 hPa Temperature Forecast Bias

FIG. 5. Temporal variation of area-averaged values (over the rectangle shown in Fig. 2): (a) MRF and LAHM 24-h forecasts and NCEP analysis of temperature (K) and geopotencial height (m) (upper and lower curves, respectively), (b) height forecasts bias (m) for MRF and LAHM 24- and 48-h forecasts, and (c) temperature forecasts bias (K) for MRF and LAHM 24- and 48-h forecasts.
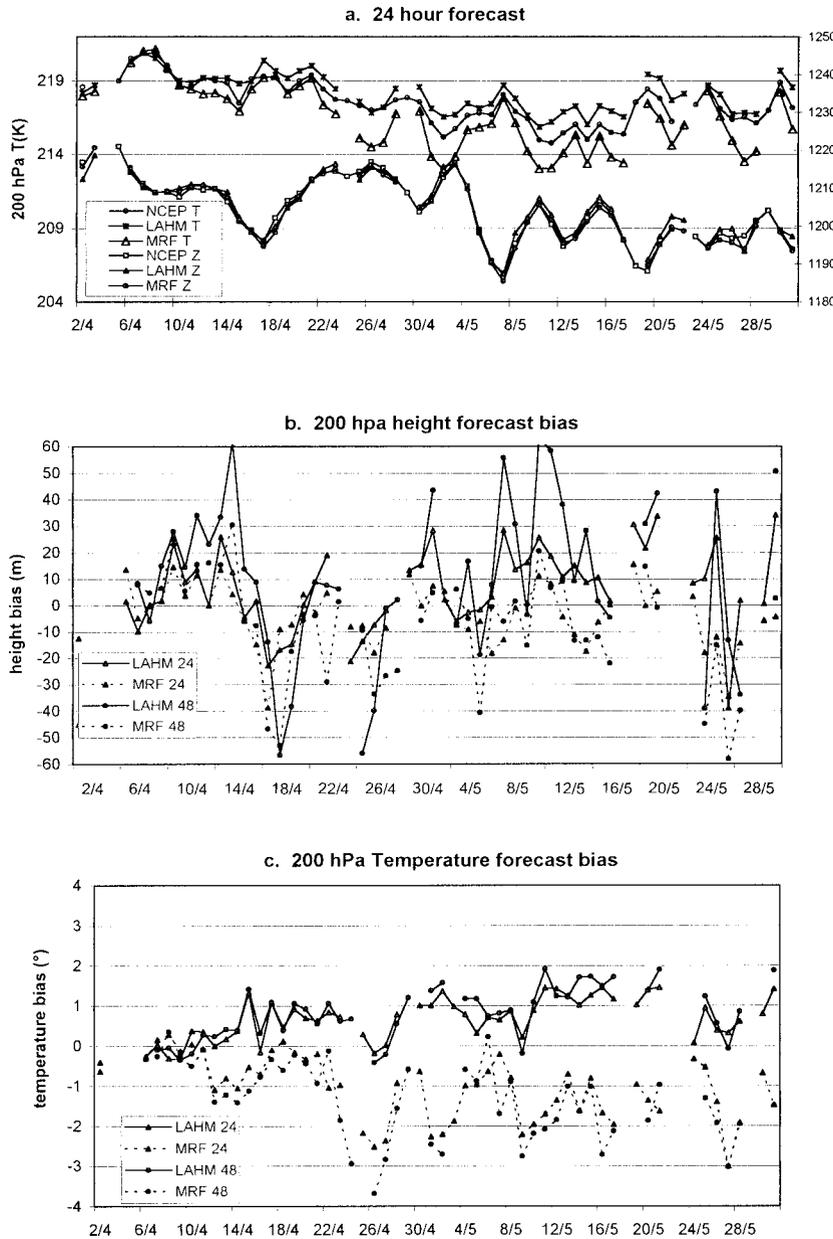
a. 24 hour forecast

b. 200 hpa height forecast bias

c. 200 hPa Temperature forecast bias

FIG. 6. Same as in Fig. 5 but for 200 hPa.

plays a cold bias of around −1° to −2°, and a crmse of around 2°, throughout the troposphere. The analysis error profile is closely followed by errors of both models (except for LAHM near the surface). In contrast to what has been observed in geopotential height errors, which tend to be less in MRF that in the analysis, temperature forecast errors are greater since the biases remain in the same sense (the model is colder than the analysis, which is also colder than the observation).

Figure 9 shows a similar analysis performed over Quintero. At this point it is worthwhile to point out the particular geographical characteristics of this station, which lies upwind of the Andes, 8 m above sea level

but very near the highest mountain peaks. Its vertical temperature profile is characterized by an inversion around 850 hPa. While both models have relatively high resolution, it is still too coarse to consider that surface singularities like the ones mentioned above will be correctly represented by model fields. However it is interesting to evaluate a station with characteristics different than those of Porto Allegre.

Above 850 hPa, there is similar behavior of geopotential errors in the analysis and forecasts compared to that found in Porto Allegre. Temperature errors do not behave similarly: the cold bias is higher at lower levels, less above 500 hPa and gradually warmer at higher lev-

Root Mean Square Error
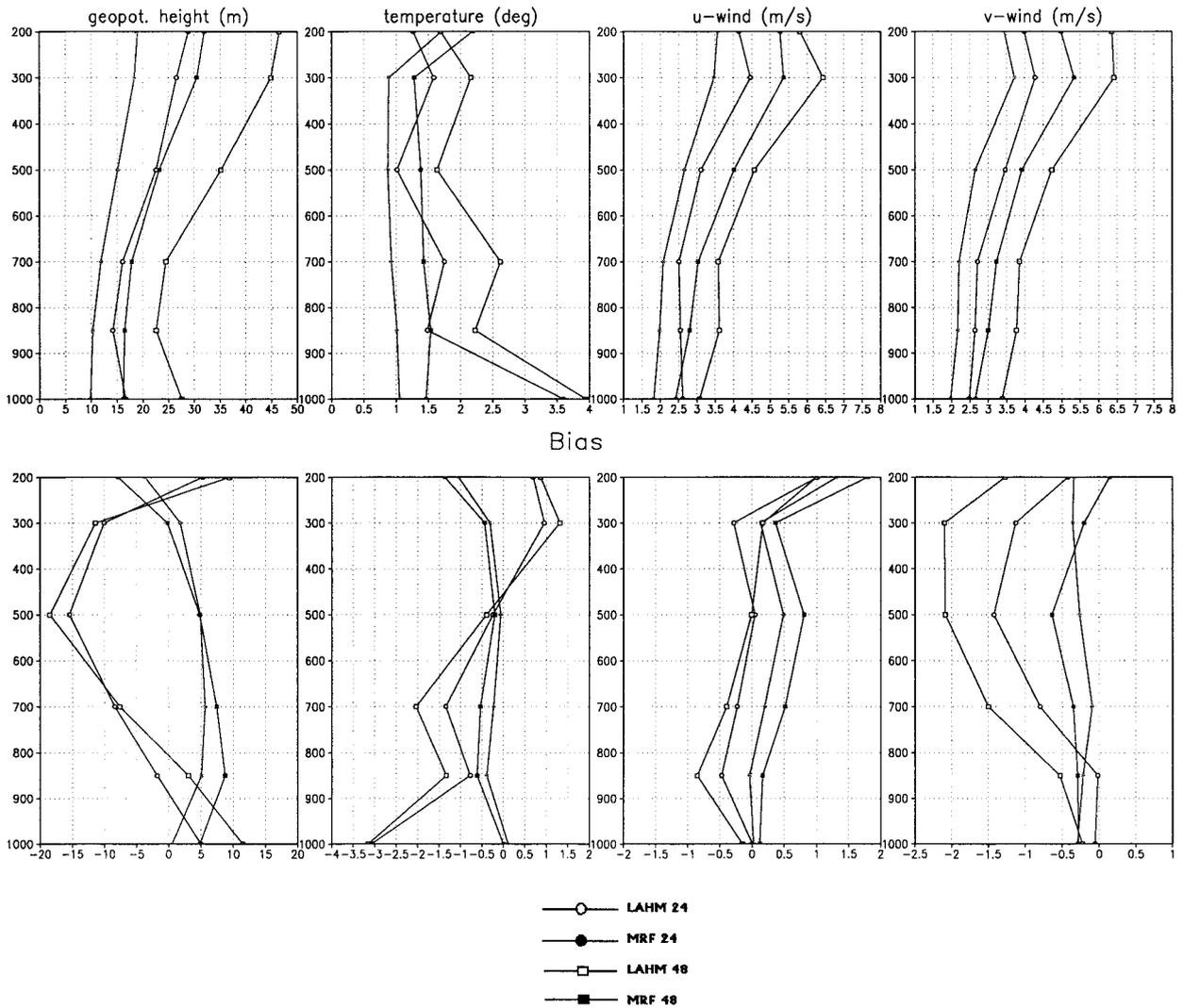

Bias

LAHM 24
MRF 24
LAHM 48
MRF 48

Fɪɢ. 7. Error variation with height of LAHM and MRF 24- and 48-h forecasts.

els, while crmse's are larger, particularily for MRF. Between the surface and 850 hPa it is hard to determine whether temperature errors (clearly larger than at Porto Allegre) could be produced by a misrepresentation of the Andes effect at this particular site.

Figure 10 shows NCEP analyses and MRF 24–48-h forecast bias and rmse for geopotential height (mean over the Southern Hemisphere and calculated using rawinsonde data as ground truth). This figure has been adapted from the information that is displayed operationally in an experimental Web site from the EMC's Global Modeling Branch. Figure 10 has been included to further illustrate that our particular quality evaluation at individual radiosonde stations is below the "hemispheric mean" performance, as can be inferred from comparison with Figs. 8 and 9. We have also calculated standard rmse for geopotential height and temperature

(not shown), in order to compare with similar figures obtained at EMC, and found that our scores are lower.

b. Precipitation

One of the key motivations to develop a forecast system is to improve quantitative precipitation forecast at the regional scale. It is widely known that higher-resolution models with suitable parameterization of convection exhibit better performance in representing processes leading to precipitation. In order to evaluate what is our real capability to provide useful precipitation forecasts and to design guidelines for future developments, LAHM 12–36-h forecasts of daily accumulated precipitation were evaluated over the box of interest, where observed data from the surface synoptic network was

Geopotential height bias
Pto. Allegre 00 UTC

Geopotential height CRMSE
Pto Allegre 00 UTC

Temperature bias
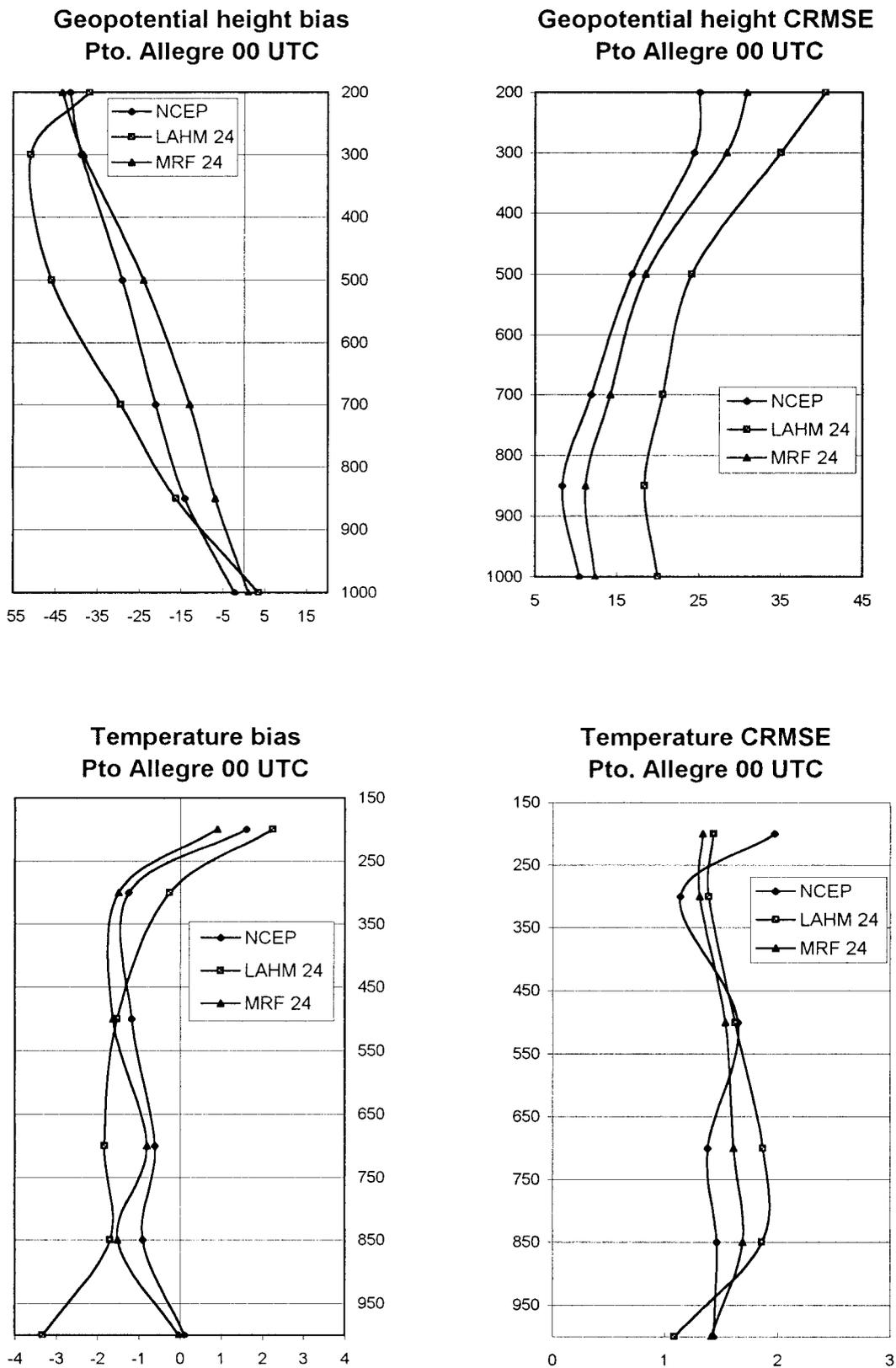## Pto Allegre 00 UTC

Temperature CRMSE
Pto. Allegre 00 UTC

FIG. 8. Crmse and bias with height at Porto Allegre (Brazil) for NCEP analysis, and MRF and LAHM 24-h forecasts. (upper panels) Geopotential height (m); (lower panels) temperature (K).

Geopotential height bias
Quintero 00 UTC

Geopotential height CRMSE
Quintero 00 UTC

Temperature bias
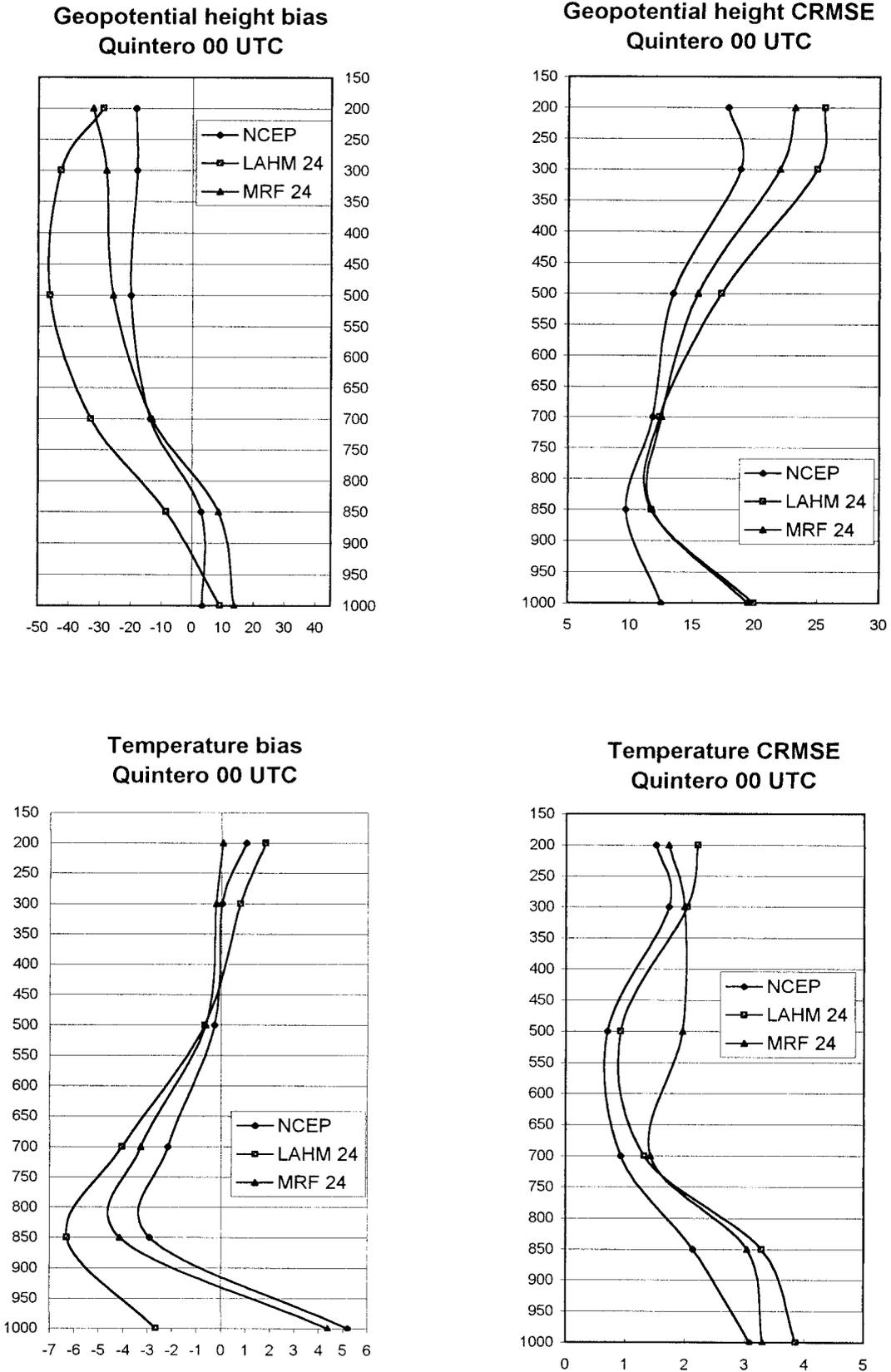## Quintero 00 UTC

Temperature CRMSE
Quintero 00 UTC

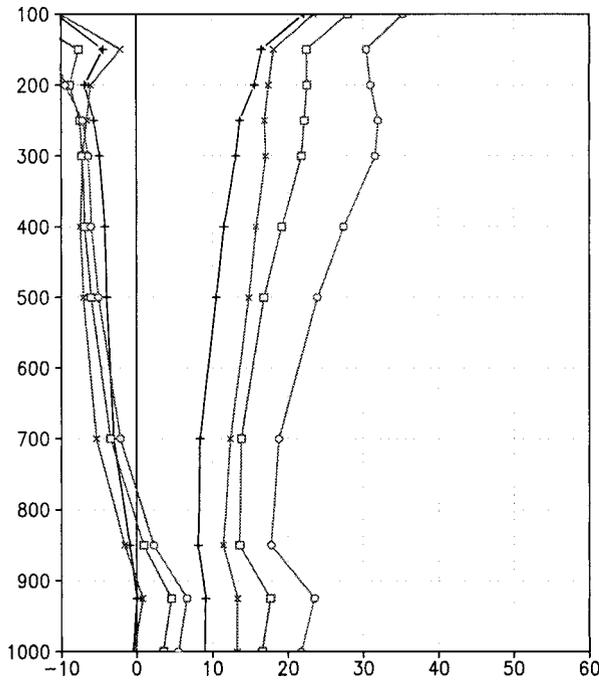FIG. 9. Same as in Fig. 8 but at Quintero (Chile).

FIG. 10. (right curves) Geopotential height rmse and (left curves) bias (in m) over the Southern Hemisphere: -+-, MRF analysis; -×-, MRF first guess; -□-, MRF 24-h forecast; -○-, MRF 48-h forecast. (Obtained from the experimental Web page of S. Saha, NCEP/EMC.)

available (provided by Centro de Previsão de Tempo e Estudos Climáticos of Brazil).

Figure 11 shows accumulated precipitation during April and May 1999, observed and obtained from

LAHM 12–36-h forecasts. Precipitation is underestimated in northern Argentina and around the northern boundary (which also coincides with the northern boundary of the integration domain). The 200-mm maximum near 32°S, 60°W is relatively well captured but the model fails to reproduce the areal coverage of the 150-mm isoplet that surrounds two maxima: that already mentioned over central Argentina and one over southern Brazil. The latter one is overestimated by the model, which at this area exhibits "gridpoint instability," associated with orographically forced ascent over a relatively high topography leading to spurious precipitation (see Fig. 2 to locate this local maximum in terrain height). A daily inspection of accumulated precipitation shows that LAHM did a good job in reproducing the event that explained most of the precipitation in April (already mentioned in section 3), responsible for the 200-mm maximum around 32°S.

Since it is relevant to compare model behavior with respect to precipitation, another period during which MRF 12–36-h precipitation forecasts were archived is included. This period (Oct–Dec 2000) is completely different from the one previously analyzed, not only because it corresponds to the warm season, but also because of the high frequency of convective systems observed over Argentina, which are certainly more difficult to capture for models with low resolution such as those employed in this study.

When accumulated precipitation is compared (Fig. 12), a spurious maximum is again detected at 28°S, 50°W in both models (larger for LAHM), whose origin has already been discussed. There are more similarities
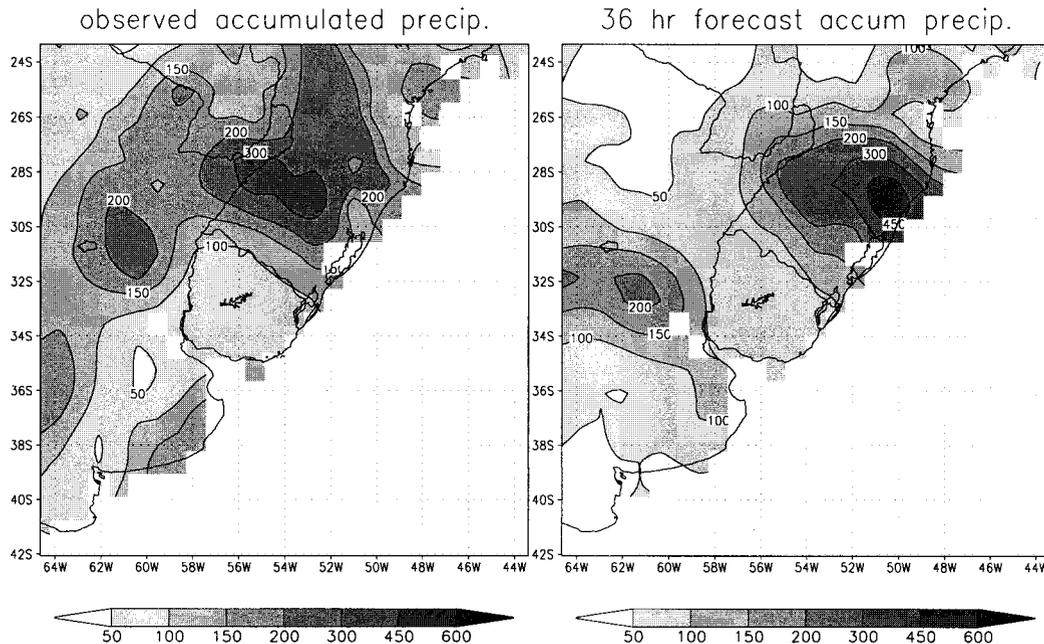

FIG. 11. Accumulated precipitation (mm) during Apr and May 1999: (left) Observed: (right) LAHM 12–36-h forecast.
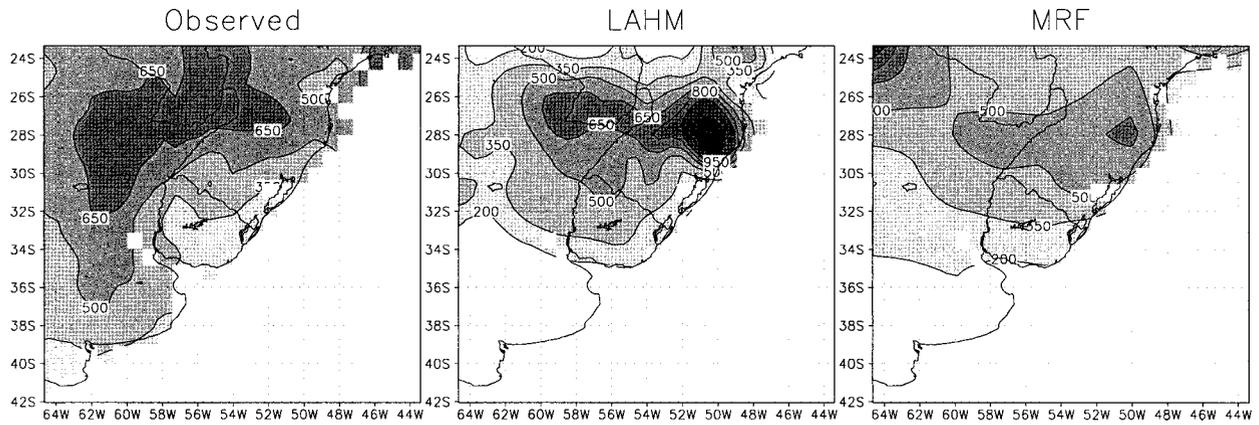
FIG. 12. Accumulated precipitation (mm) from Oct to Dec 2000: (left) observed, (center) LAHM 12–36-h forecast, and (right) MRF 12–36-h forecast.

between the forecast fields than between the observation and any forecast. Both models tend to underestimate the area covered by the above 500-mm isopleths, with the MRF being somewhat worse. Over the northwestern boundary, MRF displays another maximum that is absent in the observed precipitation field. Both models fail to represent the precipitation area south of 34°S.

FIG. 13. The ETS and bias scores for LAHM and MRF 12–36-h forecast precipitation from Oct to Dec 2000.

From the skill scores perspective (ETS and bias in Figs. 13a and 13b, respectively), MRF displays better ETS for the lower thresholds, while this behavior reverses at the higher ones; on the other hand, LAHM biases are closer to one at all thresholds. These results may be interpreted as a tendency of the regional model to better represent smaller-scale precipitation patterns, while MRF is more accurate in predicting lighter rain events. However, it is not clear whether the ETS is a complete, meaningful measure of precipitation forecast skill over this area, since the values obtained for both models, but particularly for MRF, are well above those reported in other regions in the world. For example, in Mesinger (1996) and McBride and Ebert (2000), the ETS of the MRF model is below 0.3 for all categories. Perhaps the stronger limitation of this measure of skill when applied to this region is that the amount of observations available is scarce compared to the referred studies, leading in this way to less representative results.
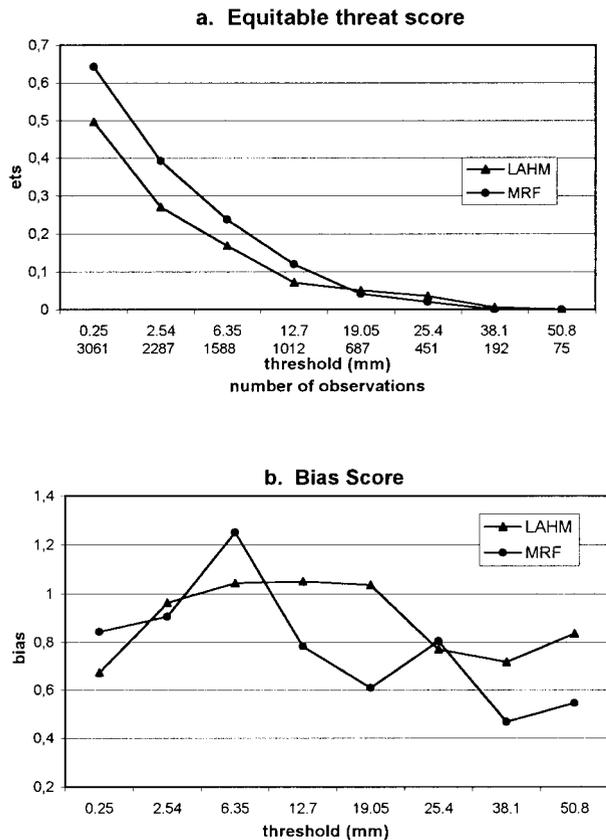
6. Discussion of results

When compared to gridded analyses, it is seen that MRF and LAHM provide useful short-range forecasts over southern South America, with bias and rmse errors bound within reasonable limits. Time and vertical structure of errors indicate their growth with longer-time forecast periods. LAHM shows a strong cold bias near the surface, which does not propagate to higher levels, related to the simple bucket model that solves the continental surface energy budget. To a lesser extent, MRF is also cold biased with respect to the analyzed temperature field. In general, LAHM has a significant portion of error explained by systematic errors, suggesting that further adjustments have to be done in order to improve forecast quality, mainly in regard to surface processes.

As has been found by other studies carried out over different regions (i.e., McBride and Ebert 2000; White

et al. 1999), MRF skill is better than that of the regional model even when the latter is run at a somewhat higher resolution. White et al. (1999), for example, who did a short-term forecast validation, comparing research and operational models over the western United States, found that on shorter timescales (up to 24 h) the Meso-Eta Model did the best, while at 36 h MRF provided the best forecasts. Their explanation for this result is that, ''the value added from mesoscale information in high resolution models is lost due to phase and amplitude errors in the individual mesoscale structures.'' Also, as stated in the introduction, the validation of high-resolution models is a problem still to be solved, mainly when it is done against coarser-resolution observational datasets. When analyzing the spatial structure of errors, it is evident that the MRF benefits from the internal consistency between the analysis and the corresponding forecast.

Nevertheless a problem identified by the present study is a cold bias with respect to two isolated points in the analysis and, consistently in both models, that affects the troposphere. This assessment clearly needs further analysis in order to determine whether this problem is of systematic origin and remains bound to southern South America or not. At the Argentine National Meteorological Weather Service, this cold bias has also been detected when NCEP analyses are compared to Ezeiza (35°S, 58°W) 1200 UTC radiosondes (H. Ciappesoni 2000, personal communication). Besides the obvious impact that this bias has on short-range forecasts, through an imprecise description of the initial thermal structure over this region, it also certainly distorts mean temperature fields derived from this dataset (i.e., the ''temperature climatology'' becomes locally altered). It should be noted that this study also documents a cold bias of the MRF global model (and of LAHM) with respect to the analysis. As is also mentioned in Wang and Paegle (1996), this cold bias could be enhanced by four-dimensional data assimilation generated analysis, which in void-data regions is of questionable value in filling the gaps between observations and is even more model dependent.

From this preliminary result it should be stressed that a quality control of how analysis and forecasts represent local characteristics should be carried out at each specific area of interest.

The evaluation of precipitation forecast quality using standard skill scores like ETS and bias leads to the conclusion that the MRF forecast is better for lower thresholds and LAHM for higher ones. However these scores should be interpreted cautiously, since the MRF ETS's seem to be better at this region than at any other region evaluated in previous studies, a result that is not in agreement with the overall quality of the model, which shows less skill over the Southern Hemisphere than over the Northern Hemisphere. It is believed that an accumulated precipitation field gives a better idea of the model precipitation forecast quality than the pre-

vious scores. In this sense, LAHM provides better geographical detail of precipitation patterns, while it tends to overestimate precipitation at relatively high terrain points.

Though the overall evaluation of both models shows that the MRF performs better, it is evident that LAHM's quality is also acceptable. This result encourages the development of this regional forecast system. However, to get clear advantages of higher-resolution models compared to global ones, emphasis should be directed to shorter-time range prediction (i.e., up to 36 h), in an attempt to get detailed structures, mainly those associated with heavy precipitation, that might not be represented by coarser grids. Currently, and as a first action resulting from this investigation, the LAHM domain and resolution have been modified to get more utility from this forecast system. Current horizontal grid spacing is 0.65° with an areal coverage extended to the north, up to 4°S. Nevertheless, from the perspective of regional higher-resolution modeling efforts, it is believed that substantial benefits will be accessed only when the initial state becomes more precisely represented. As soon as this requirement becomes fulfilled, better skill in mesoscale short-range prediction over South America could be expected with the use of regional models.

REFERENCES

Arakawa, A., and W. H. Schubert, 1974: Interaction of a cumulus cloud ensemble with large-scale environment, Part I. *J. Atmos. Sci.,* **31,** 674–701.

Cotton, W. R., G. Thompson, and P. W. Mielke Jr., 1994: Real-time mesoscale prediction on workstations. *Bull. Amer. Meteor. Soc.,* **75,** 349–362.

Davies, C., and F. Carr, 2000: Summary of the 1998 Workshop on Mesoscale Model Verification. *Bull. Amer. Meteor. Soc.,* **81,** 809–819.

Douglas, M., M. Nicolini, and C. Saulo, 1999: The low-level jet at Santa Cruz, Bolivia during January–March 1998 pilot balloon observations and model comparisons. Preprints, *10th Symp. on Global Change Studies,* Dallas, TX, Amer. Meteor. Soc., 223–226.

Fels, S. B., and M. D. Schwarzkopf, 1975: The simplified exchange approximation: A new method for radiative transfer calculations. *J. Atmos. Sci.,* **32,** 1475–1488.

Gan, M. A., and V. B. Rao, 1991: Surface cyclogenesis over South America. *Mon. Wea. Rev.,* **119,** 1293–1302.

García, N. O., and W. M. Vargas, 1996: The spatial variability of

runoff and precipitation in the Río de la Plata basin. *J. Sci. Hydrol.,* **41,** 279–299.

Kalnay, E., S. J. Lord, and R. D. McPherson, 1998: Maturity of operational numerical weather prediction: Medium range. *Bull. Amer. Meteor. Soc.,* **79,** 2753–2769.

Kamga, A., S. Fongang, and A. Viltard, 2000: Systematic errors of the ECMWF operational model over tropical Africa. *Mon. Wea. Rev.,* **128,** 1949–1959.

Mass, C., and Y.-H. Kuo, 1998: Regional real-time numerical weather prediction: Current status and future potential. *Bull. Amer. Meteor. Soc.,* **79,** 253–263.

McBride, J. L., and E. E. Ebert, 2000: Verification of quantitative precipitation forecasts from operational numerical weather prediction models over Australia. *Wea. Forecasting,* **15,** 103–121.

Menéndez, C. G., 1994: Análisis de un ciclón subantártico. *Meteorologica,* **19,** 33–42.

——, A. C. Saulo, and L. Li, 2000: Simulation of South American wintertime climate with a nested modelling system. *Climate Dyn.,* **17,** 219–231.

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

Monin, A. S., and A. M. Obukhov, 1954: Basic laws of turbulent mixing in the ground layer of the atmosphere. *Akad. Nauk SSSR Geofiz. Inst. Tr.,* **151,** 163–187.

Moorthi, S., 1997: NWP experiments with a gridpoint semi-Lagrangian semi-implicit global model at NCEP. *Mon. Wea. Rev.,* **125,** 74–98.

Necco, G. V., 1989: Extratropical weather in South America. Preprints, *Third Conf. on Southern Hemisphere Meteorology and Oceanography,* Buenos Aires, Argentina, Amer. Meteor. Soc., 149–155.

Nicolini, M., and A. C. Saulo, 1995: Experiments using the LAHM/CIMA model over Argentina in convective situations: Prelimi-nary results of precipitation fields. Programme Weather Prediction Research Rep. No. 7, WMO/TD, No. 699, World Meteorological Organization, 5 pp.

Orlanski, I., and J. J. Katzfey, 1987: Sensitivity of model simulations for a coastal cyclone. *Mon. Wea. Rev.,* **115,** 2792–2821.

——, ——, C. Menendez, and M. Marino, 1991: Simulation of an extratropical cyclone in the Southern Hemisphere: Model sensitivity. *J. Atmos. Sci.,* **48,** 2293–2311.

Satyamurty, S., J. F. Bustamante, M. J. Bottino, M. E. Seluchi, M. C. Lourenço, and L. G. Gonçalvez, 2001: An early freeze southern Brazil in April 1999 and its CPTEC NWP guidance. *Meteor. Appl.,* **8,** 1–16.

Saulo, A. C., M. Nicolini, and S. C. Chou, 2000: Model characterization of the South American low-level flow during the 1997–1998 spring–summer season. *Climate Dyn.,* **16,** 867–881.

Seluchi, M., and A. C. Saulo, 1998. Possible mechanisms yielding an explosive coastal cyclogenesis over South America: Experiments using a limited area model. *Aust. Meteor. Mag.,* **47,** 309–320.

Sinclair, M. R., 1995: A climatology of cyclogenesis for the Southern Hemisphere. *Mon. Wea. Rev.,* **123,** 1601–1619.

Taljaard, J. J., 1968: Climatic frontal zones of the S. H. *Notos,* **17,** 23–24.

Velasco, I., and J. M. Fritsch, 1987: Mesoscale convective complexes over the Americas. *J. Geophy. Res.,* **92,** 9591–9613.

Wang, M., and J. Paegle, 1996: Impact of analysis uncertainty upon regional atmospheric moisture flux. *J. Geophys. Res.,* **101,** 7291–7303.

Warner, T. W., and N. L. Seaman, 1990: A real time mesoscale numerical weather prediction system used for research, teaching, and public service at The Pennsylvania State University. *Bull. Amer. Meteor. Soc.,* **71,** 792–805.

White, B. G., J. Paegle, W. J. Steenburgh, J. D. Horel, R. T. Swanson, L. K. Cook, D. J. Onton, and J. G. Miles, 1999: Short-term forecast validation of six models. *Wea. Forecasting,* **14,** 84–108.