Objective Evaluation of ETA Model Precipitation Forecasts over South America

Sin Chan Chou
Centro de Previsão de Tempo e Estudos Climáticos - CPTEC
Instituto Nacional de Pesquisas Espaciais – INPE, Cachoeira Paulista - SP
chou@cptec.inpe.br

Mª G. A. Justi da Silva
Departamento de Meteorologia
Universidade Federal do Rio de janeiro - UFRJ, Rio de Janeiro - RJ
justi@igeo.ufrj.br

Abstract

Equitable Threat Score (ETS) and Bias Score (BIAS) were calculated for the regional Eta model precipitation forecasts for the period from February 1997 through January 1998. The scores were also calculated separately for three regions over the continent, North, Northeast, and Centre-south, for each month, and for each season of the year, and for the different forecast ranges: T+24h, T+36h, T+48h and T+60h. During this year, North and Centre-south, regions showed higher ETS, although BIAS tended to indicate overestimate of weak and moderate rain. During autumn, the scores over Northeast region were higher compared to the other two regions. The ETS showed no significant differences for the different forecast ranges, however, BIAS were generally larger at shorter range forecasts. Forecasts with initial conditions at 12 UTC resulted in better scores than forecasts starting at 00 UTC when the number of observations are reduced over the continent.

1. Introduction

Numerical weather predictions using the Eta Model (Black, 1994), have been carried by the Centro de Previsão de Tempo e Estudos Climáticos (CPTEC), since December 1996. These forecasts are provided on a twice-a-day basis, and cover most of South America and adjacent oceans. In a higher resolution model, greater detail is aimed from these forecast fields and mainly, better quality.

Numerical weather forecasts exhibit uncertainties, which can be due to the representation of the physical processes or the computational precision used in the model. In some events the forecasts fail by misplacing the weather systems, by underestimating or overestimating their intensity; whereas in other events the forecasts capture the weather system correctly.

The knowledge of the forecast errors is important to help the forecasters issue a more reliable weather prediction. The quality of some predicted variables such as pressure, temperature and wind is superior to the precipitation, which exhibit higher space and time variability. However, predicted precipitation is the variable mostly used by users.

The objective of this work is to evaluate the precipitation forecasts from the Eta Model based on statistical score indices for the period between February 1997 and January 1998.

2. MODEL DESCRIPTION

The Eta Model is set up over a domain that covers most of South America with 40 km horizontal resolution grid, 38 layers, and model top located at 50 mb. The model is run for 60 hours, twice a day, with initial conditions at 12 UTC and 00 UTC. The step-mountain or eta vertical coordinate as defined by Mesinger (1984) is used by the model. The model topography is constructed as discrete steps, with tops located at model interfaces. The step-coordinate reduces the errors that occur when computing horizontal derivatives in the vicinity of sloping terrains. Vertical resolution is higher in the bottom layers, where the first layer is about 20 m deep, and in the layers near the tropopause.

The prognostic variables are temperature, specific humidity, horizontal wind components, surface pressure, turbulent kinetic energy and cloud water. Model variables are distributed over the semi-staggered Arakawa E grid .

A forward-backward scheme modified by Janjic (1979) is employed in the adjustment stage. The horizontal advection scheme was developed by Janjic (1984) specially for the E grid and controls the cascade of energy toward smaller scales. It is used in conjunction with a modified Euler-backward time-differencing scheme. The split-explicit approach (Gadd, 1978) is used in the model integration.

The initial conditions are taken from NCEP analyses, and lateral boundaries from 6-hourly CPTEC global model forecasts. Both conditions are input to the model in 28 vertical layers, and in the form of spectral coefficients with triangular truncation T62, which is equivalent to 1.825^o resolution in the meridional and zonal directions. Sea surface temperature is obtained from an observed weekly mean and the initial albedo from seasonal climatology.

Both grid scale (Zhao and Carr, 1997) and convective precipitation (Betts and Miller, 1986; Janjic, 1994) are produced. The turbulent exchange between model layers in the free atmosphere is based on the Mellor-Yamada Level 2.5 scheme and the exchange between the earth's surface and the lowest model layer uses the Mellor-Yamada level 2 scheme (Janjic, 1994; 1996a, 1996b). The radiation package was developed at the Geophysical Fluid Dynamics Laboratory. Short-wave radiation computation scheme is that of Lacis and Hansen (1974), and the long-wave scheme is that of Fels and Schwarztkopf (1975).

3. METHODOLOGY

The main indices used to evaluate numerical weather forecasts from regional models are (Anthes, 1983; Anthes et al., 1989): the S1 skill score, which measures model ability to forecast correctly the horizontal gradients of scalar variables; the root mean square error (RMSE) between the forecast and the observations; the threat score, which measures model ability to forecast events classified into categories; the bias score, which is used in combination with the threat score and measures model tendency to systematically overestimate or underestimate the forecast of events distributed at certain thresholds; and correlation coefficients between forecast and observations or analyses.

The threat score, TS, is defined as (e.g., Anthes et al., 1989) $TS = H / (F+O-H)$, where F is the number of points of predicted precipitation above a certain threshold, O is the number of points of observed precipitation above a threshold, and H is the number of hits. In this work the evaluation of the forecasts will be based on the Equitable Threat Score (ETS) and the Bias Score (BIAS). The ETS is defined as (e.g., Mesinger and Black, 1992):

$$ETS = \frac{H - CH}{F + O - H - CH} \quad \text{where,} \quad CH = \frac{F \times O}{N}$$

CH is the number of points of random hits, and N is the number of points in the verification domain. The ETS is equivalent to TS with a correction to remove the bias from random hits.

The BIAS is defined as the ratio between the number of points of predicted precipitation above a threshold and the number of points of observed precipitation above the threshold, BIAS = F / O. When the predicted precipitation rate is higher (lower) than the observed, this score is above (below) 1. A perfect forecast results in ETS = 1 and BIAS = 1. The reliability of these indices depend on the number of verification grid-boxes containing observations.

In this work, the model performance was evaluated according to its ability to forecast precipitation amounts above certain thresholds. The precipitation amounts used are the total precipitation accumulated in 24 hours at 12 UTC. The precipitation categories are classified according to the precipitation intensity as: rain/no-rain, light rain, moderate rain and heavy rain (Table 1). The precipitation thresholds in mm correspond to rounded numbers in inches. These numbers correspond to the same thresholds used in the NCEP precipitation forecast evaluations, and adopted here to allow comparison of the Eta model skill over South America

Table 1 – Rain classification and thresholds.

Rain intensity classification	Precipitation thresholds (mm)
Rain/no-rain	0.3
Light	2.5; 6.3
Moderate	12.7; 19.0
Heavy	25.4; 38.1; 50.8

The ETS and BIAS were calculated for the forecast time ranges: 24, 36, 48 and 60 hours. Because precipitation verification was taken at 12 UTC, the forecasts at 24 h and 36 h ranges were evaluated from model forecasts with initial conditions at 12 UTC and the forecasts at 36 h and 60 h ranges, the initial conditions were at 00 UTC. The forecasts starting at 00 UTC and 12 UTC are likely to show skill differences due to the generally reduced number of observations at 00 UTC over this continent. Precipitation observations were derived from conventional surface synoptic observations and automatic weather stations. Daily observed precipitation was interpolated into an 80-km regular grid using the Kriging method (e.g., Davis, 1973) with a maximum of 10 neighbour points for interpolation. The ETS and BIAS were estimated only in the grid-boxes which contained at least one observation.

The precipitation regime over South America exhibits large space and time variability. To distinguish the performance of the forecasts for each regime, the indices were grouped into seasons. The autumn season comprised the months of March, April and May 1997 (MAM); the winter, June, July and August 1997 (JJA); spring, September, October and November 1997 (SON); and summer, December 1997, January 1998 and February 1997 (DJF). In this way, the dry, wet and transition seasons could be characterized individually. The scores were calculated over the whole model domain (South America, SA) and also over three different regions: North sector (NO), north of 15^o S and west of 45^o W; Northeast sector (NE), north of 15^o S and east of 45^o W; and Center-South sector (CS), south of 15^o S. These sectors are shown in Figure 1.

4. RESULTS

The scores are grouped according to: the regions, seasons, forecast ranges and initial conditions. This clustering of the scores reveals the model performance according to the weather systems, to the model systematic errors and to the amount of observations used by the analyses.

During summer, the main large scale feature over the continent is the Bolivian High and the South Atlantic Convergence Zone (SACZ). The coupling of active tropical convection with equatorward penetration of frontal systems results in large amounts of precipitation. The Intertropical Convergence Zone (ITCZ) also merges with this tropical system, The meridional displacement of ITCZ is responsible for the large precipitation over coastal areas of North and Northeast regions. Mesoscale weather systems, such as mesoscale convective complexes and squall lines, can be identified embedded within these large scale convective cloud patterns. During winter, subsidence predominates over the central part of the continent resulting in generally cloud free areas; frontal systems are more intense and can reach low latitudes; ITCZ activity is located further north of the continent; and easterly disturbances bring warm-cloud precipitation to the east coast of Northeast region.

A. REGION

Figure 2 shows the ETS and BIAS scores for each region. These scores were calculated including the four forecast ranges and the whole studied period. The magnitude of the scores are comparable to those obtained with the same model over North America (Mesinger and Black, 1992; Rogers et al., 1996), and also to other regional models (Juang et al., 1997). However, the ETS over South America decrease sharply toward heavier rains. The BIAS show that the rain/no-rain category is well predicted, however, light and moderate rains are overestimated; and heavy rains are underestimated.

The model shows better precipitation forecast performance over CS sector where ETS is higher and BIAS is closer to 1. This sector is dominated by transient weather systems that travel across the continent during all year. The NO sector exhibits smaller ETS and larger BIAS indicating some deficiency in forecasting precipitation from deep tropical convective systems over the Amazon Forest. This sector suffers from low density, and inadequate distribution of observations, mostly along the rivers, which could be non-representative of the regional precipitation. The NE sector is the smallest one, but of highest surface observation density. The ETS for this sector are comparable to those obtained for the whole domain SA, however, this sector exhibited magnitude of BIAS about 1. The distribution of grid-boxes containing at least one precipitation observation on a particular day is shown in Figure 1.

B. SEASONS

Weather systems may be more frequent at different times of the year. The seasonal variation of the model performance may show its dependence on the different weather regimes. The model may show good ability to forecast certain weather systems that occur at certain times of the year, and may show deficiency to forecast other types of weather systems. The behavior of the model performance may help the identification of the physical processes which are well represented by the model and those which deserve more attention, and presumably, leading to improvements.

Figure 3 shows the curves of ETS and BIAS for each period: DJF, MAM, JJA, and SON in the SA region. MAM shows better ETS and BIAS at weak precipitation rates. For heavy precipitation rates, the forecasts are improved during JJA and SON, when ETS are higher and BIAS are closer to 1. MAM is the period when the ITCZ is positioned at its southernmost latitude and when rains are more frequent over the NE sector. On the other hand, during JJA, the central part of the continent is practically cloud free, and most of precipitation occurs to the south of 20^{o} S and north of the equator.

The whole domain, SA, shows relatively similar scores for the different seasons, this could misleadingly indicate that the model skill is independent with time of the year. The separation of the scores again into different smaller regions reveals a large seasonal variability. This can be clearly seen, for example, from the ETS and BIAS for NE sector in Figure 4. During JJA, precipitation is systematically underestimated over NE. This is not the case when the whole domain is considered because there was a compensation from overestimated precipitation over the NO sector. During JJA, most of the precipitation in the NE region is due to warm clouds that are brought into the continent by easterly waves that propagate over the Equatorial Atlantic Ocean. The clouds are not necessarily transported over the continent, a significant amount is produced locally as the strong easterlies interact with the topography along the coast. The difficulty of the model in forecasting such type of precipitation may have different reasons such as: the proximity of the region to the eastern lateral boundary of the model, the sparse or lack of observations over the ocean which prevents the model from correctly capturing the structure of the easterly disturbance, and deficiency in the representation of warm convective clouds by the convective parameterization scheme.

Whereas the lowest scores over NE sector occur in JJA, the highest occur in MAM. This shows that NE sector contributed largely to the scores of the whole domain, SA. The scores over CS sector show a seasonal performance opposite to NE, that is, during JJA over CS sector the scores are higher, and during MAM the scores are lower (Figure 5). During JJA, the precipitation over CS sector are practically due to frontal passages. Eta model forecasts for frontal systems over South America have been assessed in a study carried out by Bustamante (1998), which showed good forecasts. These scores corroborate those results.

The overestimate of precipitation forecasts over NO sector occurs systematically in all seasons of the analyzed period. The dry season, which includes JJA and SON, has smaller ETS and larger BIAS; whereas the wet season, DJF and MAM, has higher ETS and generally smaller BIAS. The precipitation regime of NO sector shows correlation with the scores (Figure 6).

C. FORECAST-PERIOD RANGES

The use of initial conditions derived from NCEP Global model analyses and lateral boundary conditions from CPTEC Global model causes a large delay to issue the short range forecasts. The Eta model is generally run for 48 hour forecasts at other weather centers. At CPTEC, the forecast range was extended to 60 hours to provide longer useful forecasts.

Figure 7 a, and b show two different forecasts, 24-h and 48-h, verifying on the same date, August 19 1997, 12 UTC. The uncertainty in the prediction of weather events such as frontal displacement poses a difficulty to the forecasters. The verifying observations suggest that the older forecast, the 48-hour, had positioned the precipitation band more correctly.

Figure 8 shows the ETS and BIAS for 4 forecast ranges: 24 h, 36 h, 48 h and 60 h. The 24-h forecasts show higher ETS in all precipitation thresholds, but some overestimate of moderate precipitation rate. This feature is more evident in the CS sector. For the whole domain, SA, the curves are approximately close to one another, as occurred to each individual region. This lack of spread of the curves indicate that there is no significant degradation with time of precipitation forecast performance produced by the Eta model at short range forecasts during this one year period. Consequently, the 60-hour forecast turned out to be equally useful for forecasting precipitation.

Although, individually chosen cases can show distinction in the quality of the different forecast ranges, the scores for ensemble of cases show no significant systematic preference for any forecast range.

D. INITIAL CONDITIONS

The uneven distribution of the conventional surface observations, which tend to lie along the coastline, the riversides and southeastern part of the South America, makes the continent notorious for the low quality/density of observations. The distribution is inadequate not only in space but also in time. Most of the remaining operating stations take the observations at 12 UTC. A sharp decrease in observation amount occurs at 00 UTC, not only upper-air but also surface observations. This affects enormously the quality of the analyses of the observations and, consequently, the quality of the initial conditions used by the regional and global models.

Since the verifying precipitation observations were taken at 12 UTC, the scores from the 4 forecast ranges were clustered into two groups: one with initial conditions at 12 UTC, using the 24-h and 48-h forecasts; and the other with initial conditions at 00 UTC, using the 36-h and 60-h forecasts.

Figure 9 show the ETS and BIAS of the forecasts for these two groups. As expected, the ETS are higher for the 12 UTC group, confirming the superior quality of the forecasts from runs which used more observations. This superiority occurs at all precipitation categories and all sectors of the domain. The 12 UTC runs also exhibit BIAS values closer to 1 at most categories, except for weak precipitation categories.

5. DISCUSSION AND CONCLUSIONS

In this work, the Eta model forecasts over South America during the first 12 months of operation were assessed objectively using the Equitable Threat Score and the Bias score. The scores were grouped into 4 sectors of the continent: SA, NO, NE, and CS; into 4 seasons of the year: DJF, MAM, JJA, and SON; into 4 forecast ranges: 24 h, 36 h, 48 h and 60 h; and initial conditions: 00 Z and 12 Z. Better forecasts are produced in the CS sector. Inter-seasonal variation is larger in NE, where scores are the highest in MAM and lowest in JJA. BIAS larger than 1 occurs generally at weak precipitation, and smaller than 1 at heavy precipitation, leaving the moderate precipitation thresholds with BIAS closer to 1. The 60-h forecasts show comparable quality to the other forecast ranges, which suggests that an increase of the forecast range to 72 hours may still produce useful forecasts. The superior quality of the 12-UTC-run forecasts over the 00-UTC-run supports the need to increase the observation network density of the continent.

The results shown in this work refer to weather systems that predominate over the continent during the year of El Nino of 1997/1998. During El Nino years, the NO sector suffers from deficit of precipitation while the Southeastern part of the continent suffers surplus of precipitation. These conditions were observed over the continent during JJA and SON, 1997 (Bell and Halpert, 1998). The extension of these results to the following years should be done with care, mainly in a La Nina year, when the prevailing large scale atmospheric conditions are typically different. Scores estimated in the following year of this study have shown differences in the model precipitation forecast performance.

The high score at rain/no-rain category can be largely induced by the use of interpolation to create the observed precipitation grid. The interpolation smoothes the original field and spreads the precipitation nuclei; the model would not be penalized in case of small displacement of the predicted precipitation regions and this would result in higher scores.

The identification of the model deficiency and ability is crucial to help the forecasters to take decisions when a forecast report is issued. The characterization of the model errors provides more confidence on the model outputs and qualifies the model to be used as a proper forecast tool. The model products are constantly accessed by many organizations and individuals, they are used for various purposes, and considered of great usefulness. Nevertheless, continued research on identifying the model errors is an important indication of a model gaining continuous improvement and development. Cases of good and bad scores can be isolated and clustered into weather systems to identify the atmospheric structures that cause difficulties to the forecast. This can allow consideration for possible improvements to the model.

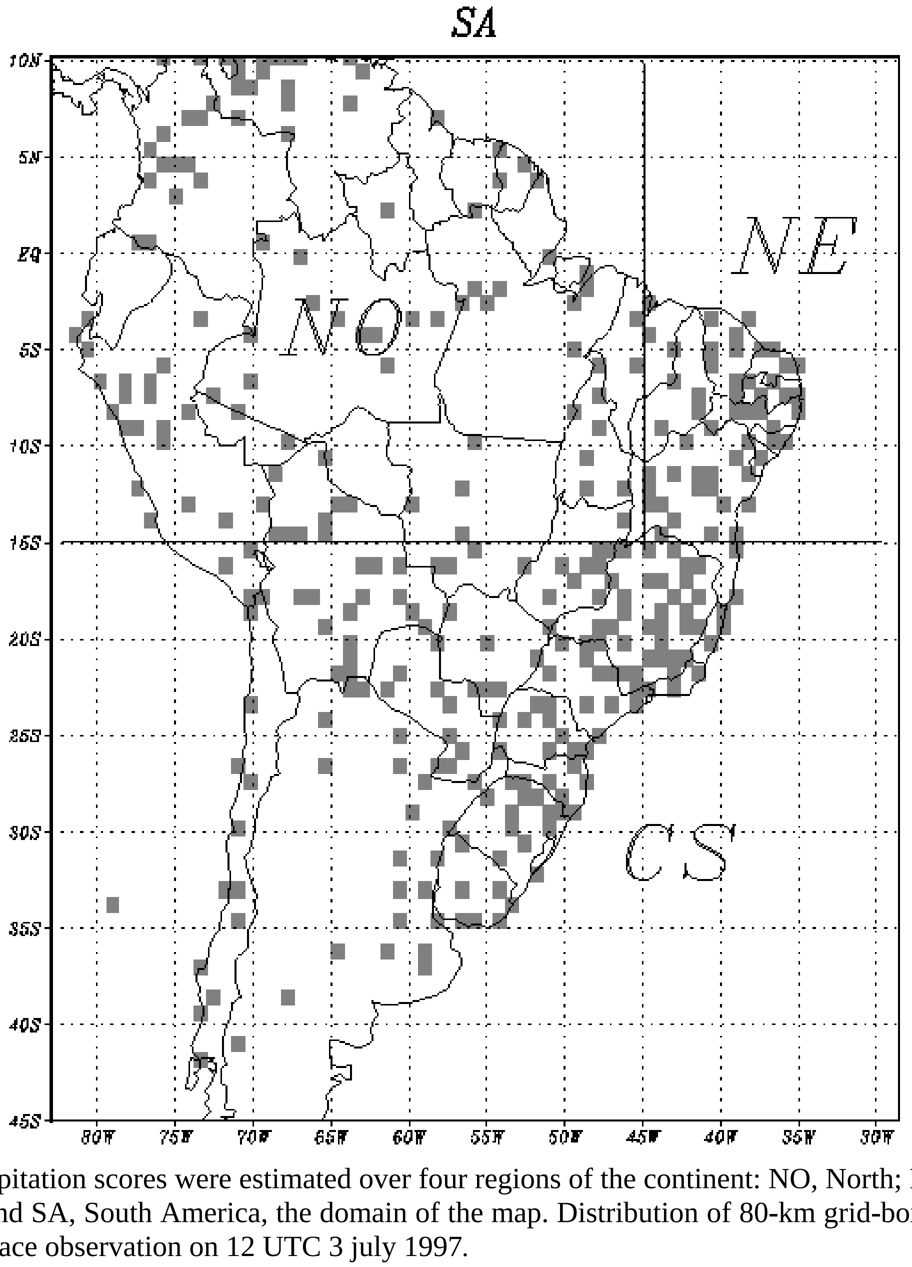

Figure 1 - Precipitation scores were estimated over four regions of the continent: NO, North; NE, Northeast; CS, Center-south; and SA, South America, the domain of the map. Distribution of 80-km grid-boxes wich contained at least one surface observation on 12 UTC 3 july 1997.

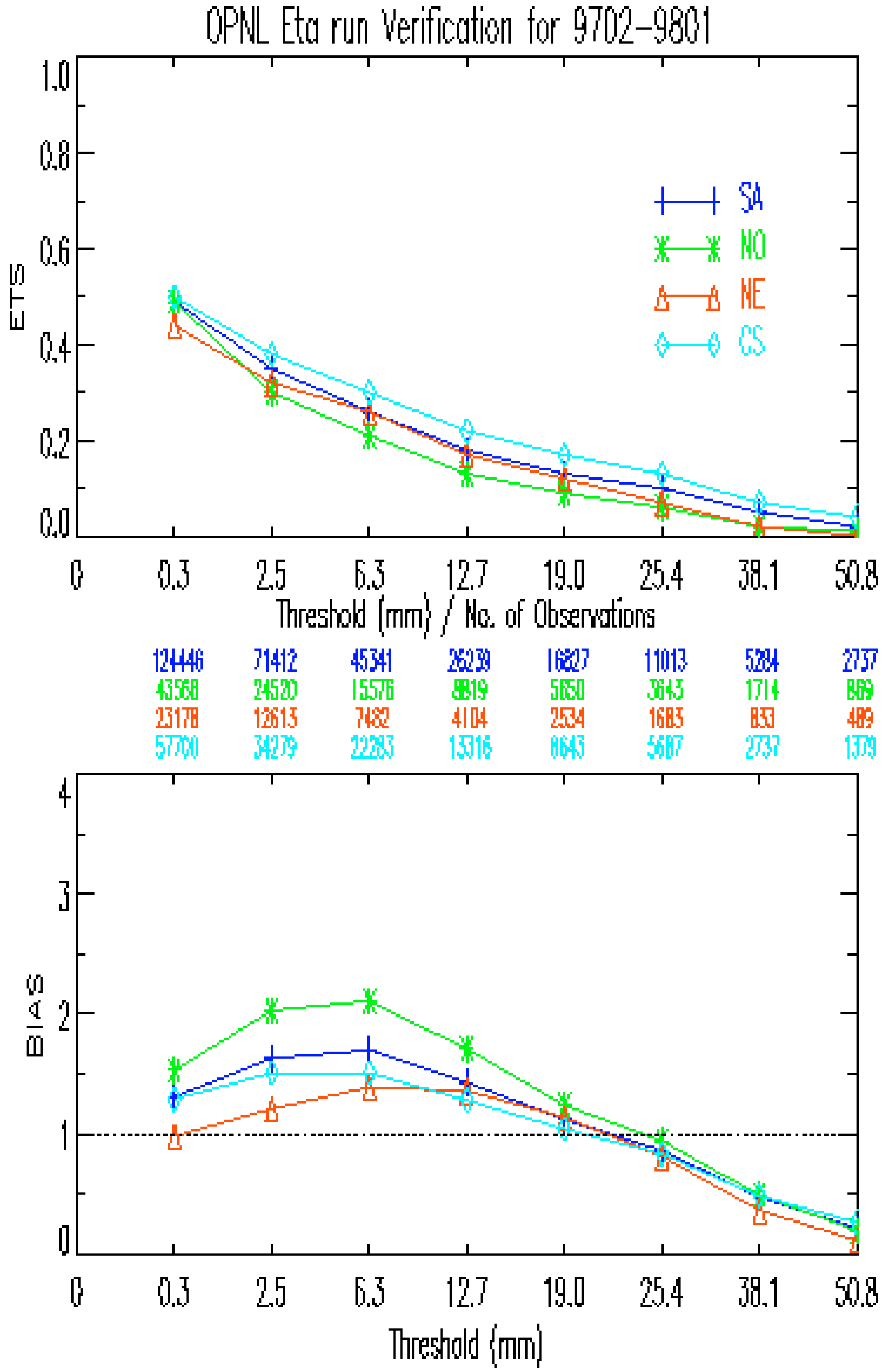

Figure 2 - Equitable threat score and bias score for the period: February 1997 through January 1998, for each region. The horizontal axis has two titles, one refers to the precipitation thresholds above, and the others, refers to the number of observations below. The 4 rows of numbers below the thresold axis give the number of gird-boxes with observation over each thresold. The first row of numbers refer to SA region; the second, NO; the third, NE; and, the fourth, CS. The higher these number, the higher the confidence on the scores.

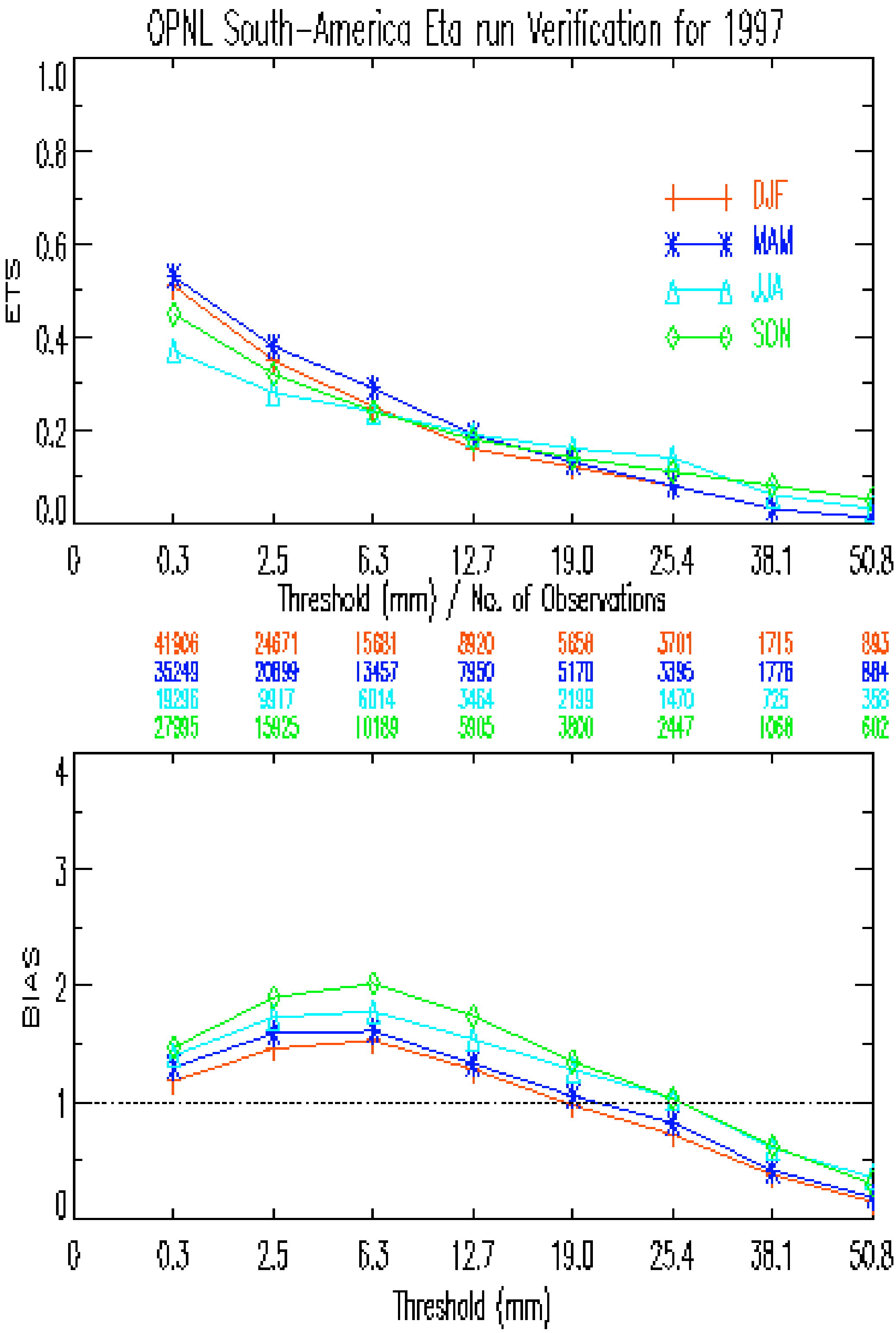

Figure 3 - As in Figure 2, except for the South America (SA) region, for each season.

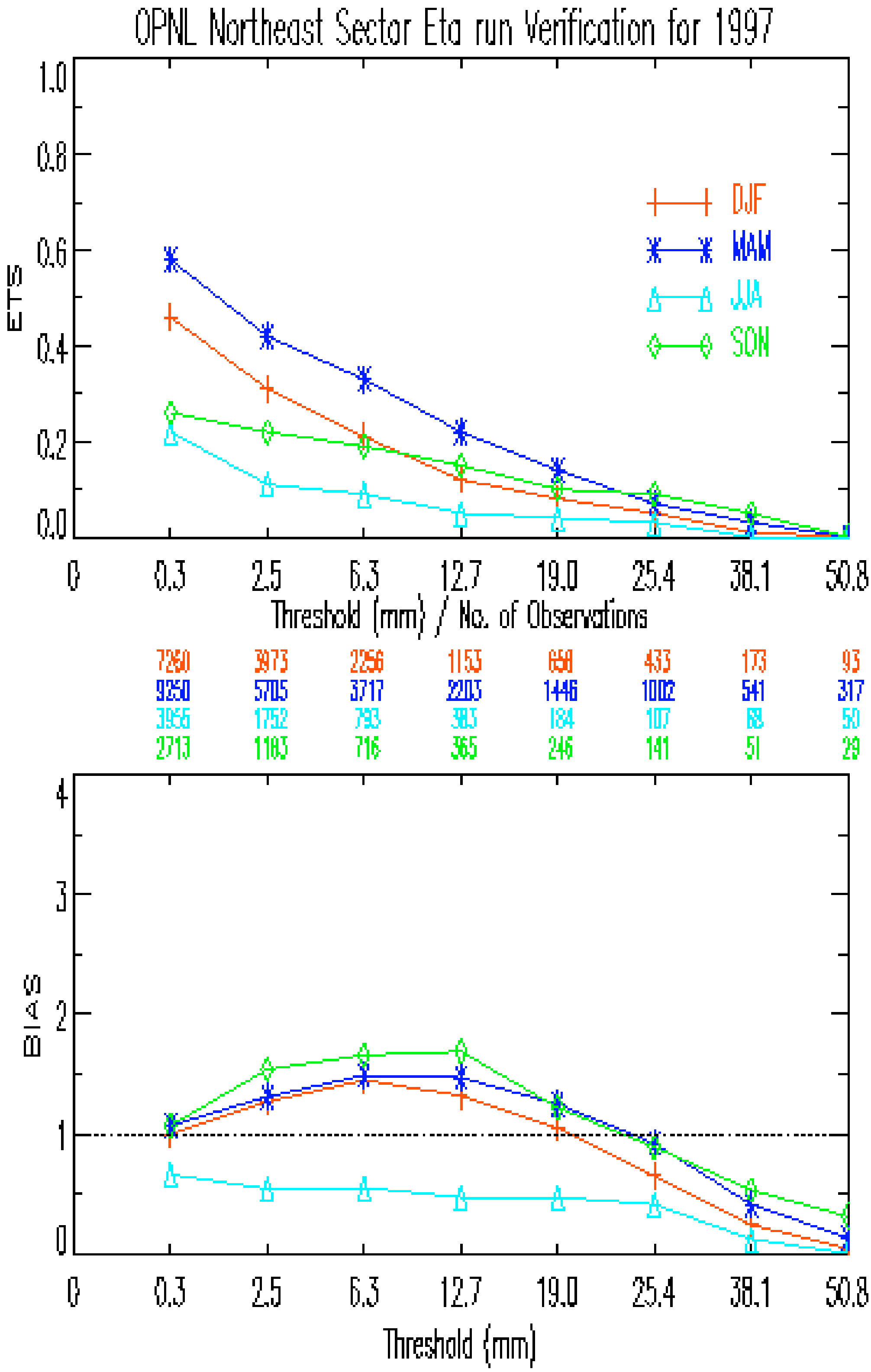

Figure 4 - As in Figure 3, except for the Northeast (NE) region, for each season.

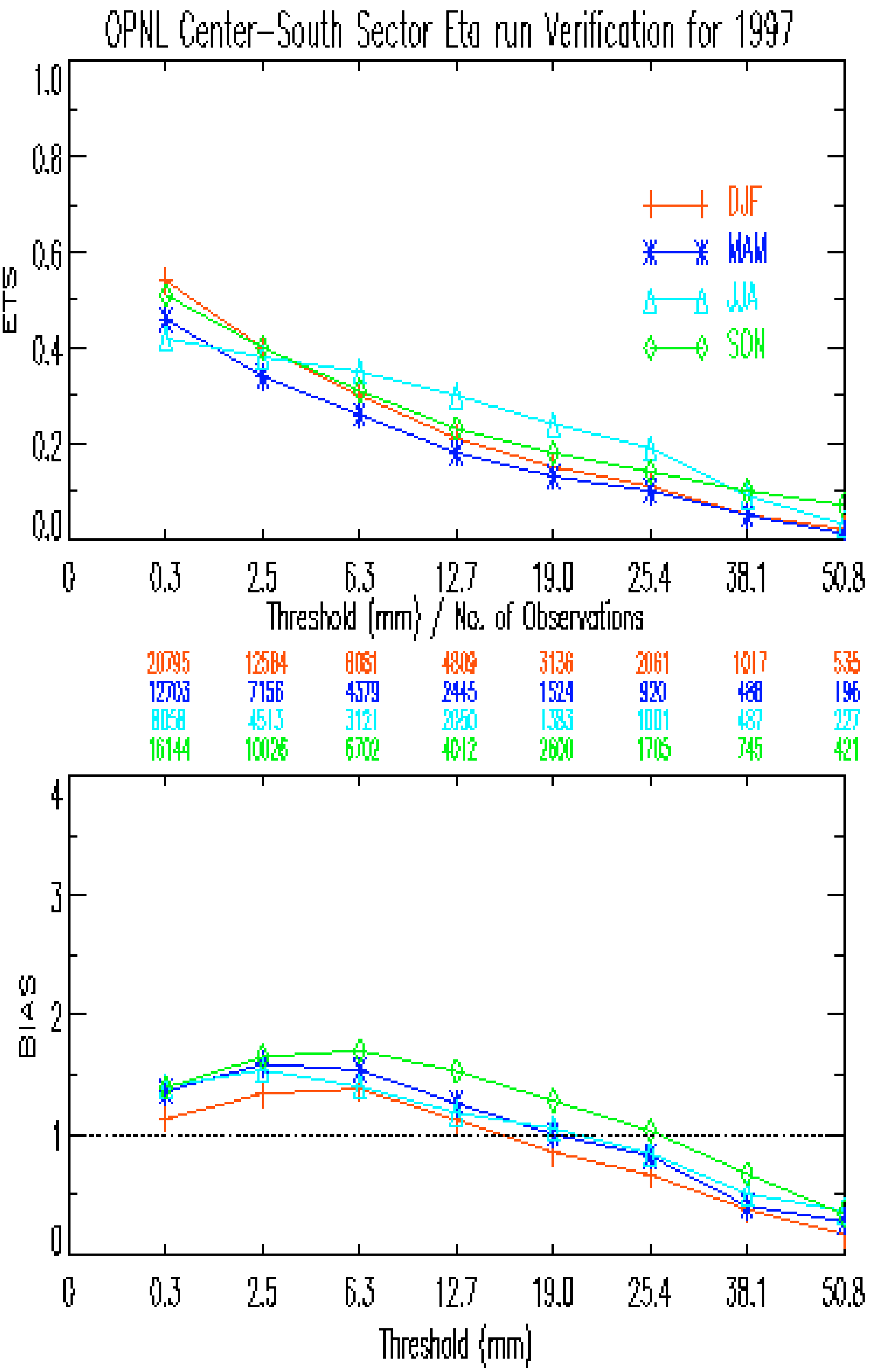

Figure 5 - As in Figure 4, except for the Center-south (CS) region, for each season.

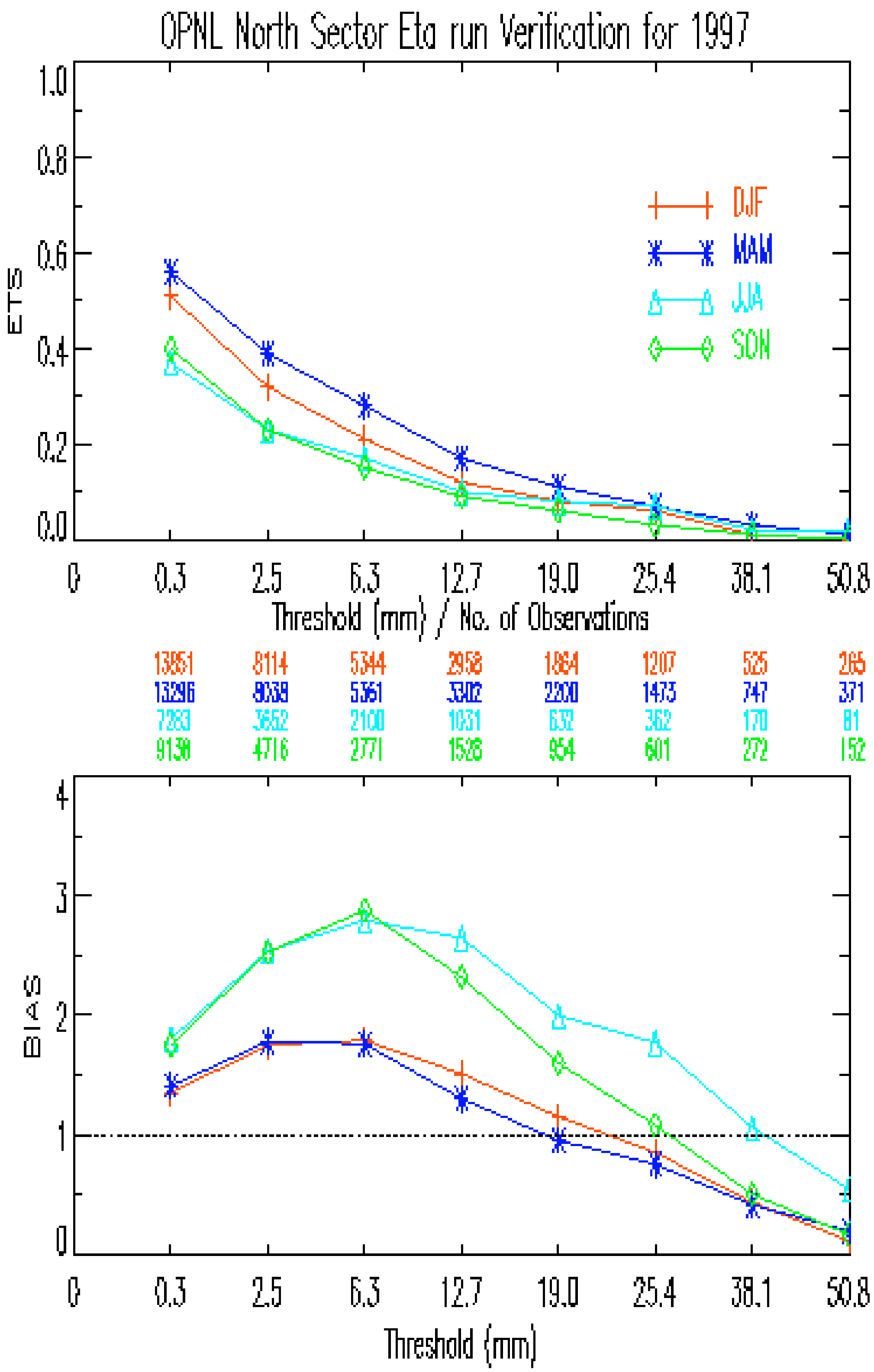

Figure 6 - As in Figure 5, except for the North (NO) region, for each season.

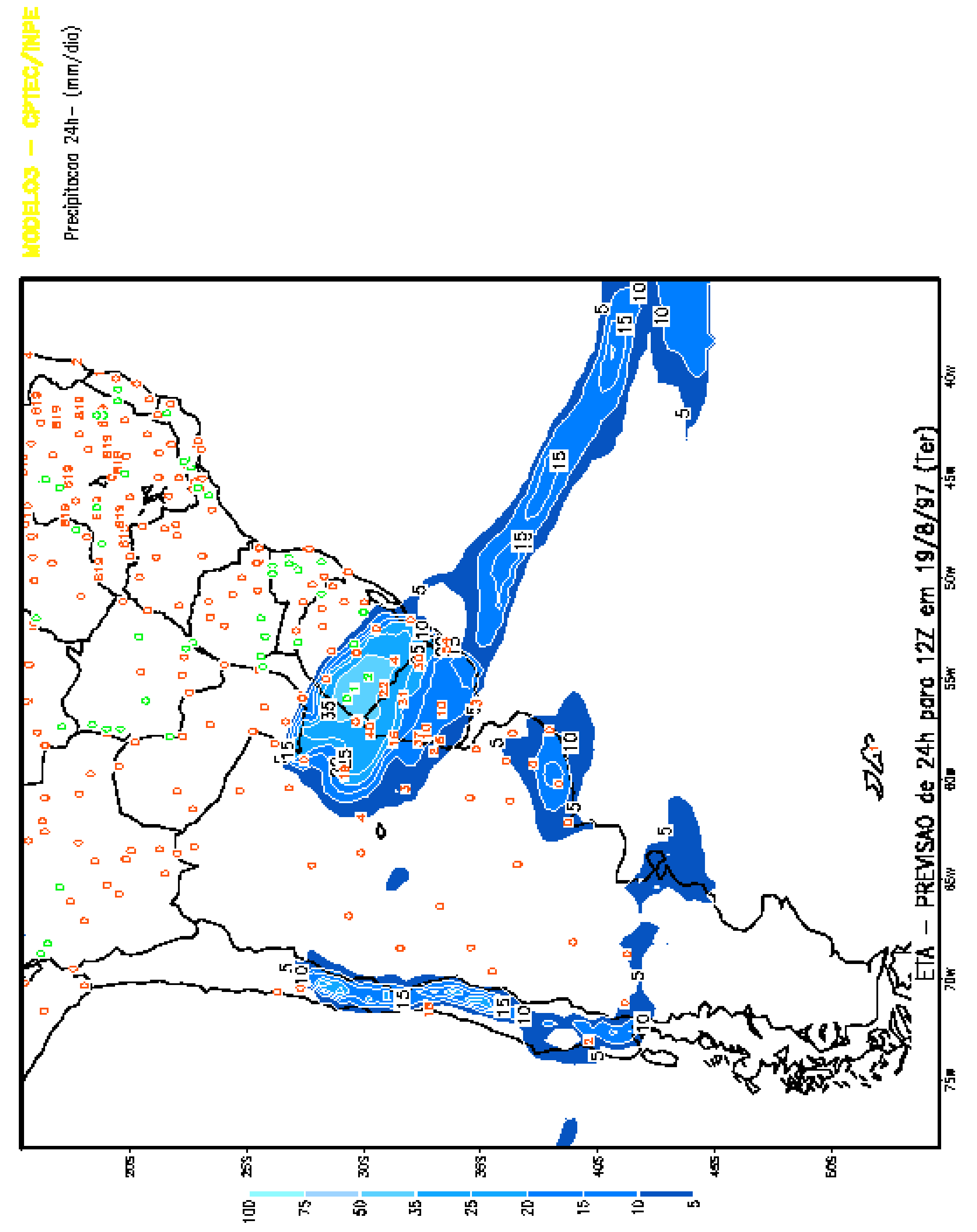

Figure 7(a) - 24-hour Eta Model forecast of total precipitation verifying on 12 UTC August 19, 1997.

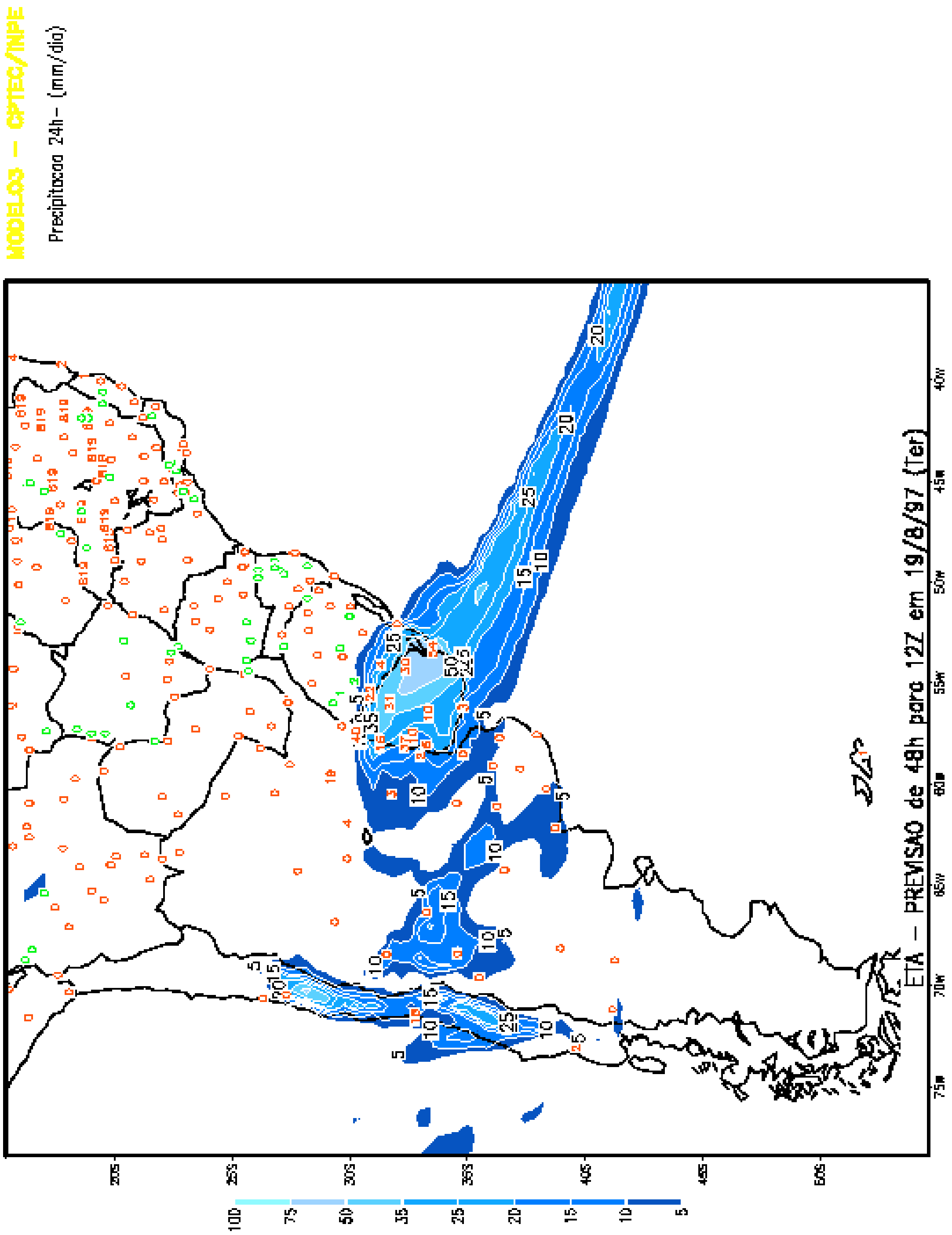

Figure 7 (b) - 48-hours Eta Model forecast verifying on the same day.

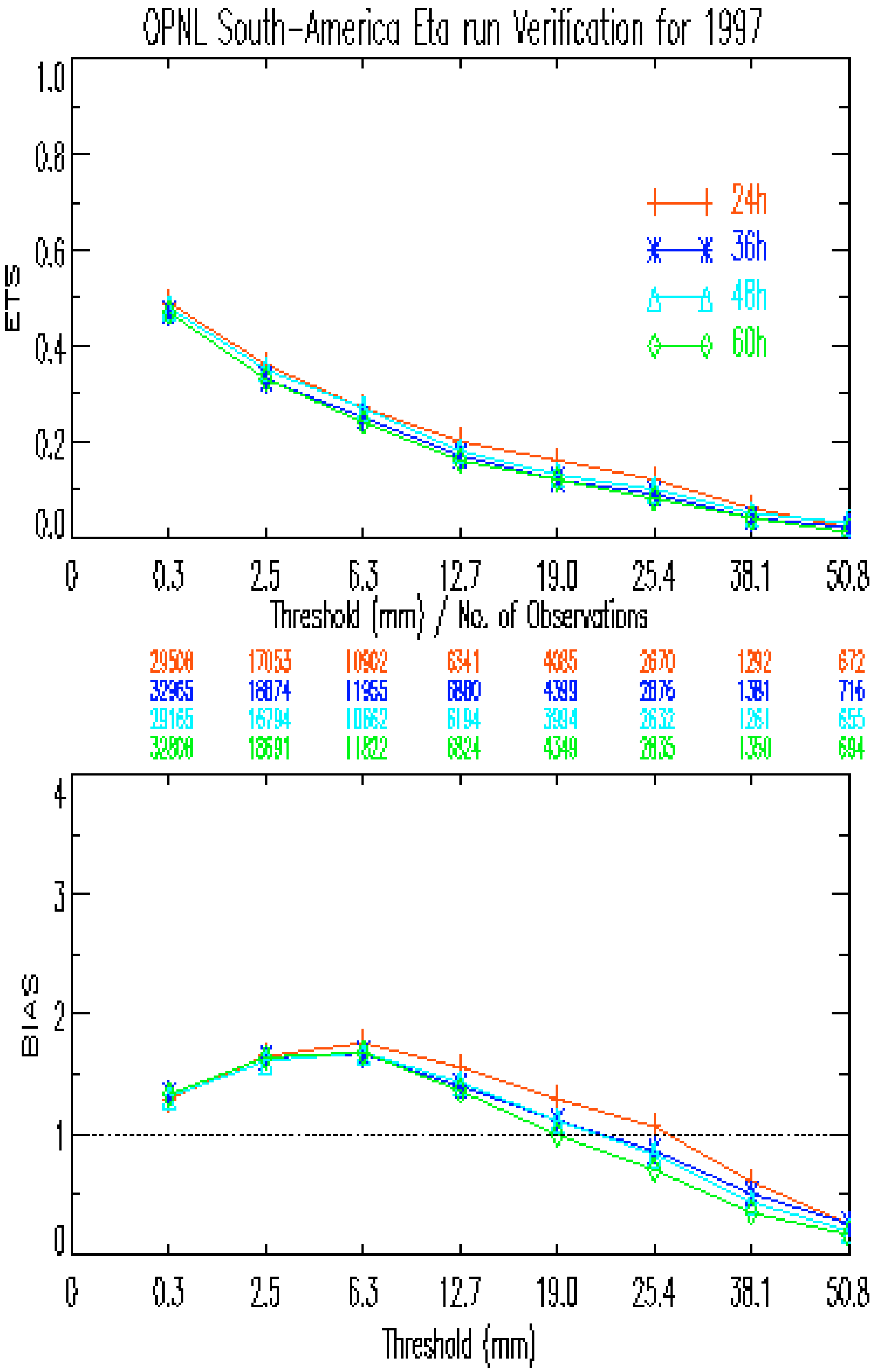

Figure 8 - As in Figure 2, except for different forecast ranges.

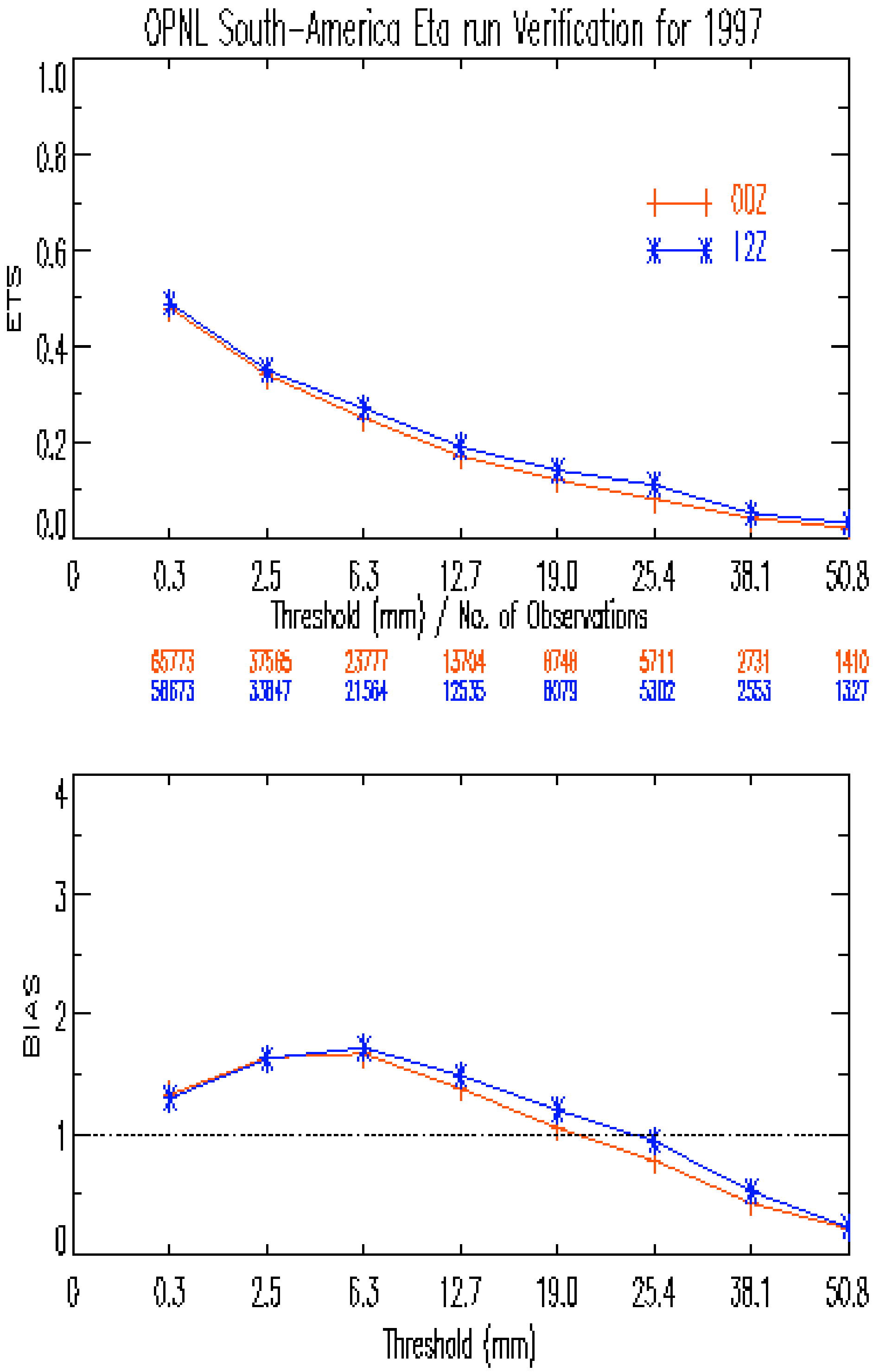

Figure 9 - As in Figure 2, except for different initial conditions.

6. References

Anthes, R. A., 1983: Regional models of the atmosphere in middle latitudes. *Mon. Wea. Rev.*, **111**, 1306-1335.

Anthes, R. A., Y.-H. Kuo, E.-Y. Hsie, S. Low-Nam, and T. W. Bettge, 1989: Estimation of skill and uncertainty in regional numerical models. *Quart. J. Roy. Meteor. Soc.*, **115**, 763-806.

Bell, G. D., and M. S. Halpert, 1998: Climate assessment for 1997. *Bulletin of the American Meteorological Society*. **79**, S1-S50.

Betts, A. K., and M. J. Miller, 1986: A new convective adjustment scheme. Part II: Single column model tests using GATE wave, BOMEX and artic air-mass data sets. *Quart. J. Roy. Meteor. Soc.*, **112**, 693-709.

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

Bustamante, J. F., 1998: Regional numerical weather prediction models: comparison between FSU and Eta model performance over South America (in portuguese). MSc dissertation. Instituto Nacional de Pesquisas Espaciais. São José dos Campos, SP. Brazil.

Davis, J. C., 1973: *Statistics and data analysis in geology*. J. Wiley. 550pp.

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.

Gadd, J., 1978: A split-explicit integration scheme for numerical weather prediction. *Quart. J. Roy. Meteor. Soc.*, **104**, 569-582.

Janjic, Z. I., 1979: Forward-backward scheme modified to prevent two-grid-interval noise and its application in sigma coordinate models. *Contributions to Atmospheric Physics*, **52**, 69-84.

Janjic, Z. I., 1984: Non-linear advection schemes and energy cascade on semi-staggered grids. *Monthly Weather Review*, **112**, 1234-1245.

Janjic, Z. I., 1994: The step-mountain eta coordinate model: further developments of the convection, viscous sublayer and turbulence closure schemes. *Mon. Wea. Rev.*, **122**, 927-945.

Janjic, Z. I., 1996a: The Mellor-Yamada level 2.5 scheme in the NCEP Eta model. Preprints, *Eleventh Conference on Numerical Weather Prediction*, Norfolk, VA, American Meteorological Society, Boston, MA, 333-334.

Janjic, Z. I., 1996b: The surface layer in the NCEP Eta model. Preprints, *Eleventh Conference on Numerical Weather Prediction*, Norfolk, VA, American Meteorological Society, Boston, MA, 354-355.

Juang, H. H.; S. Hong, and M. Kanamitsu, 1997: The NCEP regional spectral model: an update. *Bull. Amer. Met. Soc*., **78**,. 2125-2143.

Lacis, A. A., and J. E. Hansen, 1974: A parameterization of the absorption of solar radiation in the earth's atmosphere. *J. Atmos. Sci*., **31**, 118-133.

Mesinger, F., 1984: A blocking technique for representation of mountains in atmospheric models. *Rivista di Meteorologia Aeronautica*, **44**, 195-202.

Mesinger, F., and T. L. Black,. 1992: On the impact on forecast accuracy of the step-mountain (eta) vs. sigma coordinate. *Meteor. Atmos. Phys*, **50**, 47-60.

Rogers, E., T. L. Black, D. G. Deaven, G. J. DiMego, Q. Zhao, M. E. Baldwin, N. W. Junker, and Y. Lin, 1996: Changes to the operational "early" eta analysis/forecast system at the National Centers for Environmental Prediction. *Wea. Forecasting*, **11**,. 391-413.

Zhao, Q., and F. H. Carr, 1997: A prognostic cloud scheme for operational NWP Models. *Mon. Wea. Rev*., **125**, 1931-1953.