Meteorological forecasting for the European Southern Observatory in Chile

R. Deidda([1]), M. Marrocu([1]) and A. Speranza([2])

([1]) *CRS4, Centro di Ricerca, Sviluppo e Studi Superiori in Sardegna. Via Nazario Sauro 10, I-09123 Cagliari, Italy*
([2]) *Università di Camerino, Dipartimento di Matematica e Fisica. Via Madonna delle Carceri, I-62032 Camerino, Italy*

(ricevuto ?; approvato ?)

Summary. — The potential of numerical weather prediction to supply a useful support to flexible scheduling of astronomical observations is investigated. We applied some common tools presently used in numerical meteorology at regional scale in order to evaluate the ability to forecast local meteorological conditions (cloud cover, air temperature and wind speed) at Cerro Paranal and Cerro La Silla in Chile, where telescopes of ESO (the European Southern Observatory) are sited. The first part of this paper is devoted to evaluating the accuracy of forecasts of the ECMWF (European Center for Medium-range Weather Forecast) general circulation model for ESO needs. With this aim, analysis and 24-/48-hour forecasts from ECMWF are systematically compared with observations at the ground meteorological stations of Paranal and La Silla, and with vertical profiles of radiosounding launches at Antofagasta and Quintero. The second part of this paper is aimed at improving the ECMWF forecasts at telescope sites by means of Kalman filter statistical post-processing and meteorological limited area modeling. Encouraging results are obtained concerning temperature, whereas much less satisfactory results are obtained for wind field and cloud cover. The most critical aspects of atmospheric dynamics affecting the local forecast are discussed within the limits of available information.

PACS 92.60 – Meteorology.
PACS 95.85 – Astronomical observations.

1. – Introduction

Reliable weather forecasts can significantly contribute to the optimization and to a more efficient scheduling of astronomical observations. Forecasts of cloudiness and water vapor content can provide a useful support in scheduling photometric or infrared night observations; the forecast of temperature at ground level can be used to preset air temperature within the dome so that the mirror telescope is in thermal equilibrium with the external temperature at the dome opening; the forecast of extreme wind speed at the height of the telescope are of great utility in reducing the problem of interrupting observations in course, thereby reducing loss of time and money and preserving the structure of the telescope. Another useful parameter to be known in advance is the atmospheric seeing which is related to the distribution of turbulence along the observation path [1].

The meteorological part of these topics has been the subject of a feasibility study partially financed by the European Southern Observatory (ESO) [2–4], of which in this paper we describe the main scientific results. More specifically, our research addressed to "short range" (12 to 72 hours in advance) forecasting of atmospheric conditions (in particular cloud cover, wind speed and temperature) at Cerro La Silla and Cerro Paranal in Chile (see Fig. 1).

The two sites are located at sub-tropical latitude along the western slope of the Andes, in a narrow strip characterized by arid climate. This strip, together with another North-South elongated area of high mountain climate, separates the Pacific Ocean from the pluvial-tropical region extending over the entire northern portion of South-America (again, see Fig. 1). Water vapor, vertically mixed by tropical convective activity, tends to be trapped by strong inversions over the coastal ocean and extratropical (middle-latitude) perturbations seldom penetrate as far north. As a consequence, the local climate is essentially arid. On occasion, however, particularly deep extratropical troughs can bring the middle-latitude jet stream of the southern hemisphere to the latitude of the two considered sites and, given their altitude, cause adverse atmospheric conditions. It is clear that any short range forecasting procedure for the site in question must start from an adequate forecast of the "weather perturbations" which are the primary generators of changes in the atmospheric conditions of interest for astronomical activity, but must take into account, as well, local agents (topography, heat sources, etc.) which also play an important role in the local forcing of atmospheric conditions. This can, in fact, be done in different ways and the analysis of some alternative forecasting approaches specifically finalized to ESO activities was the object of the feasibility study.

Modern short range weather forecasting is developed through a procedural chain which, almost invariably, starts from a planetary scale numerical forecast obtained by a general circulation model (GCM) in which the laws of atmospheric motion (from solar heating, to air motion, to viscous dissipation) are explicitly represented on a numerical grid: processes determining atmospheric evolution on a time-scale of the order of one day develop, in fact, on the planetary scale. Such models have been set up and are currently used with the specific purpose of forecasting the global evolution of the atmosphere and, as such, are optimized with respect to the global variance of meteorological fields (pressure, temperature, winds, etc.). As a consequence, a GCM forecast is usually not suited for local forecasting even where space-time numerical resolution seems adequate.

In order to perform local forecasts, starting from predictions of a GCM, it is usually necessary to insert additional steps in the procedural chain. A preliminary step consists in comparing the global analysis of meteorological fields, released by the same centers issuing the global forecast, with local meteorological information as provided by the local observational network. Surprising as it may seem, we do not expect the analysis to coincide with observations at specific stations, even if the observations themselves have been used in the analysis. The reason is the same mentioned above: also the analysis, as the global forecast, results from an optimization process aimed at representing the global fields and not local details. The gap between global analysis-forecast and local meteorological conditions can be filled with phenomenological-statistical techniques as well as with more locally adapted numerical models (typically the so called limited area models, LAMs).

In this paper we describe our analysis of the meteorological conditions at ESO sites in terms of the above described procedure. In section 2 the data sets analyzed and used in this study are presented. In Section 3 we discuss the comparison of direct observations at the ESO sites with the ECMWF analysis. In Section 4 we analyze ECMWF short range forecasts and compare them with local meteorological observations. In Section 5 we propose possible improvements of the ECMWF forecasts using statistical post-processing based on Kalman filtering. In section 6 we discuss the results of some numerical integrations of a limited area model nested into the GCM of ECMWF. In Section 7 we draw our conclusions.

2. – Data sets

In this study we used daily ECMWF meteorological analysis (at 00, 06, 12 and 18 GMT) and 24-/48-hours forecasts (starting at 12 GMT) from 1989 to 1993 in an area covering the entire Chile. The horizontal resolution of the grid is 1.125 degrees (about 120 km at Chilean latitudes) during the period from 01/01/89 to 16/09/91 and 0.5625 degrees (about 60 km) from 17/09/91 to 31/12/93. In Figure 2 (left) we show the whole ECMWF grid at the 0.5625 degree resolution together with the surface geopotential height in the domain of interest, while in Figure 3 (left) a cross-section at Paranal latitude is shown.

The ECMWF data set was composed of upper air fields available over 15 pressure levels together with several surface fields. Upper air fields are: geopotential height, temperature, relative humidity and components of wind velocity. Surface fields are: total cloud cover, supplementary total cloud cover, 2-meter temperature, 10-meter wind, surface geopotential height and many other fields needed for LAM integrations.

Local observations and measurements at the astronomical sites in Chile were also available through ESO in order to perform comparisons with ECMWF analysis and forecasts. The local data set consists of observations of cloud cover in Paranal, surface meteorological parameters (10-meter wind speed and direction, 2-meter air temperature) in Paranal and La Silla, and vertical profiles of temperature, relative humidity, wind speed and direction from radiosonde launches in Antofagasta and Quintero. The observation sites are shown in Figure 2, in their geographical context; some information on the sites and on their nearest ECMWF grid points can be found in Table I.

Before collecting all these data into a single and easy to access database, we performed some checks of internal consistency and eliminated some macroscopic errors from meteorological observations in La Silla and Paranal and vertical profiles from radiosoundings in Antofagasta and Quintero. The periods and the frequency of available meteorological observations are summarized in Table II.

3. – Comparison of ECMWF analysis with local observations

In order to evaluate the limits of the analysis in representing local meteorological variability, we performed, by means of the above described data sets, a systematic comparison between ECMWF analysis and meteorological observations. In particular, we compared: vertical profiles at Antofagasta and Quintero with profiles extracted from ECMWF analysis; visibility observations in Paranal with relative humidity, total cloud cover and supplementary total cloud cover ECMWF fields; 2-meter temperature and 10-meter wind speed observed at meteorological stations in La Silla and Paranal with the corresponding variables in ECMWF analysis. The period selected for all comparisons is the entire year of 1993. The ECMWF grid-resolution for this period is 0.5625 degrees, as shown in Figure 2 (left).

3˙1. *Vertical profile comparisons.* – Comparisons between vertical profiles of relative humidity, temperature and wind velocity derived from balloon launches at Antofagasta (daily at 12GMT) and at Quintero (daily at 00 and 12 GMT) were made using two methods to extract the vertical profile of a generic meteorological variable from ECMWF analysis:

- Nearest point method: the nearest point of the ECMWF grid to the local station is determined (see Table I for coordinates and distances) and the corresponding vertical profile is extracted from the analysis.

- Bilinear method: the four nearest points of the ECMWF grid are selected and the vertical profile is computed by bilinear interpolation at constant vertical levels.

Deviations between ECMWF analysis and vertical profiles from radiosoundings were evaluated by interpolating with a cubic spline every ECMWF profile (in the nearest point or the bilinearly interpolated one) at the same pressure levels at which the balloons recorded the measurements. Results of these comparisons are summarized in Table III in terms of correlations, mean errors and mean absolute errors between the vertically interpolated analysis values and the vertical profiles from radiosoundings for each launch in 1993. The table shows that the bilinear interpolation is closer to the observations than the nearest-point.

The comparisons of wind speed profiles in Quintero revealed some anomalies in the ECMWF data set. Figures 4 and 5 show plots of correlations, mean errors and mean absolute errors of winds in 1993. A different behaviour before and after March 1993 is evident. After some investigation it was clear that the Quintero balloon soundings were not used for winds from March 1993 to March 1995 due to unacceptable large differences from the first guess. These differences suddenly came back into acceptable range in April 1995, so the assumption is that some consistent errors in the measurement instruments were corrected.

3˙2. *Cloud cover comparisons.* – Total cloud cover is one of the most difficult meteorological fields to forecast and analyze. In order to estimate daily cloud cover observations (at 0 and 6 GMT) in Paranal during 1993 we used three different fields available from ECMWF analysis: total cloud cover, supplementary total cloud cover (a re-analyzed field since the first one is considered quite unreliable) and 3D upper-air field of relative humidity.

Since cloud cover is not a local field, we considered three different ways to extract information from the ECMWF data: in the first case we take only the value or the profile at the grid point nearest to Paranal, in the other two cases we consider one or two lines of neighboring points (giving a 9 or 25-point stencil). The best results were obtained by averaging the 25-point neighboring values, which correspond approximatively to a square area of side 250 km, centered in Paranal.

It is clearly difficult to extract from the above ECMWF fields a classification similar to the empirical one used in Paranal which ranges from 1 to 5 as follows: value 1 for clear, 2 for slightly cloudy, 3 for cloudy, 4 for very cloudy and 5 for overcast. For this reason, we decided to consider only two categories: clear if its value is 1 and cloudy otherwise.

In Table IV the skills, defined as the percentages of good predictions of clear and cloudy sky, of four different criteria used to estimate cloud cover in Paranal are summarized. Criterion E1 consists in guessing that the sky is always clear ("optimistic hypothesis"). Criterion E2 is based on ECMWF total cloud cover (2D field with values ranging from 0 to 1): we assumed clear sky if the average resulting from the values of the 25 neighboring points is lower than 0.9, otherwise the sky is assumed cloudy. Criterion E3 is the same as E2, but it is applied to ECMWF supplementary total cloud cover. For the last criterion E4, the ECMWF relative humidity (3D upper-air field) is used: a mean vertical profile is computed by averaging the values of the 25 neighboring points in each pressure level; then the sky is assumed cloudy if in the mean vertical profile there are at least 3 levels (above the 500 mb) with more than 70% relative humidity, otherwise the sky is considered clear.

The "optimistic hypothesis" criterion (E1) was applied since Paranal is located in an arid region and most days of the year are clear: for example in 1993 80÷81% of the days were clear. By using the ECMWF fields for estimating the cloud cover with procedures E2, E3 and E4 we are able to improve the percentage of success only marginally, with respect to the "optimistic hypothesis" assumption.

3·3. *Two-meter temperature comparisons.* – Errors between ECMWF analysis and observations of 2-meter temperature (T2m) are systematically computed for the considered period. We first determine the error of the "24h persistence model", which consists in assuming that the 24-hour forecast of T2m coincides with the value of T2m observed 24 hours before. Use of such a model is justified by the fact that the time of autocorrelation is very long (typically on the order of two weeks) for the temperature field: correlations and averages of errors and of absolute errors at 00, 06, 12 and 18 GMT in 1993 are summarized in columns A of Table V for the Paranal and La Silla stations.

Next, we compared observations of 2-meter temperature with the corresponding values derived from analysis using either the 2-meter temperature of the nearest point of the ECMWF grid, or the bilinear interpolation of T2m values in the four nearest points. Results in the two cases are rather similar, even if the comparison with the bilinearly interpolated value of 2-meter temperature is generally a little better than with the nearest grid-point value. For this reason we will show only the results obtained using values from interpolation of the ECMWF field at the station site. It can be seen from columns B of Table V, by comparing mean ($\overline{\epsilon}$) and absolute ($\overline{|\epsilon|}$) deviation, that a great part of the error is a systematic error. This is an expected result, since analyzed data are referred to the altitude of the model which is quite different from the real altitude.

One way to correct such systematic error is to estimate T2m by interpolating, at the altitude of the station, the vertical profiles of upper-air temperature resulting from the analysis. Although this simple approach has still the limitation that soil effects are not properly taken into account, results shown in columns C of Table V are considerably better than the previous ones: correlation exceeds 0.9 for La Silla and mean absolute error is about 1.1 at 00 and 06 GMT. However, in all cases a sensible systematic error ($\overline{\epsilon} \neq 0$) is still present.

3·4. *Ten-meter wind velocity comparisons.* – We performed the same kind of analysis described above on observations of 10-meter wind speed (W10m). We compared, in this case, measured W10m data with values interpolated from the W10m field of ECMWF analysis, obtaining the results shown in Table VI. The agreement is, in this case, not very good: the mean absolute error is about 3 m/s, to be compared with a mean value of the field of about 6 m/s; correlations are 0.7, at most. Such a poor agreement is due to the fact that for this field the effects of soil and orography cannot be neglected. One way to account for these effects could be to re-analyse wind fields obtained from a forecast model at global scale with a "mass consistent model" which is a diagnostic model specifically designed for representing the effects of orography on wind field flows averaged over time intervals from 10 min to 1 hour (see for example [5]).

4. – Comparison of ECMWF forecasts with local observations

With the aim of verifying the reliability of ECMWF forecasts at the location of ESO observatories in Chile, we proceed essentially in the same way as in the comparisons of ECMWF analysis described in the previous section. In such a way we evaluated the worsening of forecast fields of interest with respect to the analysis.

4·1. *Vertical profile comparisons.* – We performed a systematic comparison of vertical radiosoundings of temperature and wind velocity coming from balloons released daily at 12 GMT in Antofagasta and Quintero in 1993 with the corresponding variables derived from ECMWF 24 and 48-hour forecasts. We then extended to the whole period 1989-1993 the evaluation of ECMWF forecast errors of temperature and wind.

We used the two methods (nearest point and bilinear interpolation) explained in the previous section to extract vertical profiles of meteorological variables from the ECMWF forecasts on standard pressure levels and then we applied a cubic spline to interpolate the vertical profiles at the same pressure levels at which the balloons recorded the measurements. As for the ECMWF analysis, we found that the profiles are closer to vertical observations if the bilinear interpolation is used rather than the nearest-point method.

The results of comparisons of 24 and 48-hour ECMWF forecasts with vertical profiles from radiosoundings for each launch in 1993 are summarized in Table VII in terms of average values of correlation, mean error, and mean absolute error for temperature and wind velocity profiles; for an immediate comparison, we have also reported, in the same table, results previously found with ECMWF analysis (already shown in Table III). This kind of comparison was also extended, for temperature profiles, to the whole radiosounding dataset of the 5-year period from 1989 to 1993 (see again Table VII). It can be noted that the mean error of the analysis is about 2.5 degrees, while for the 24-hour forecast it is about 3 degrees: the worsening in the global score of the forecast is only half a degree with respect to the analysis. An interpretation of this result can be that a great part of the errors in temperature forecasts can be attributed to the coarseness of the observative network and not to the ECMWF forecast model; this is plausible since the analysis is used as an initial condition for the forecast.

By comparing wind velocity profiles from radiosoundings with those obtained from ECMWF forecasts we found anomalies similar to those exhibited by the ECMWF analysis (shown in Figures 4 and 5). The causes of these anomalies were discussed in the previous section. The anomalous period during March 1993 and a different behaviour before and after that period is evident also in comparing the radiosounding wind directions with the corresponding directions interpolated from ECMWF analysis, 24-hour, and the 48-hour forecasts that are plotted for Quintero 1993 launches in Figure 6 (angles in the upper levels, between 100 and 350 mb). In the same figure the histograms of angles are also plotted.

Since conditions of very strong wind (wind speed at telescope level exceeding 15÷20 m/s) can cause the closure of the telescope dome and, consequently, the interruption of astronomical observing activities, we focused our attention on errors in the evaluation of strong winds in ECMWF vertical profiles. In particular, we computed the average on each daily vertical profile of relative wind velocity errors between radiosoundings and ECMWF for Antofagasta and Quintero launches. Relative errors of wind velocity in vertical profiles were computed using the following equation:

$$wind\_error(day) = \frac{1}{N} \sum_{N} \left| \frac{V_{ECMWF} - V_{SOUND}}{V_{SOUND}} \right| \tag{1}$$

where V_{ECMWF} is the velocity derived from ECMWF analysis or forecast, while V_{SOUND} is radiosounding velocity; N is the numbers of terms in each profiles used to compute the errors.

Relative errors of wind speed were estimated in the range of interest for this study and for the aim of ESO, by imposing the condition that in each term of the summation (1) the ECMWF velocity (V_{ECMWF}) or the radiosounding velocity (V_{SOUND}) must be greater than 10 m/s: in such a way, N is the numbers of terms in each profiles in which at least one velocity (or both) satisfy this minimum. In Table VIII are reported the scores obtained during 1993 and during the 5-year period 1989-93. The bad results for Quintero 1993 are certainly related to the anomalous period after March 1993. In the case of wind the results are worse than for temperature: the errors for the 24-hour forecast are considerably larger than for the analysis. The main cause is probably to be found in the smoothed description of orography at the spatial resolution of the ECMWF forecast model.

4·2. *Two-meter temperature comparisons.* – In Table IX we show results of comparison between observed 2-meter temperature (T2m) and the ECMWF 24-/48-hour forecast of temperature. The values compared were computed at the same altitude of the station of interest using the vertical profile bilinearly interpolated from ECMWF upper-air temperature field. For reasons already discussed in the previous section, results of direct comparisons with ECMWF 2-meter temperature fields are not presented here. It can be seen that the mean absolute error is, in general, greater than that obtained by comparing observations and analysis: in particular, for 24-hour forecasts it is about 1.3 degrees for Paranal and 1.7 degrees for La Silla. This difference is clearly related to the presence of a greater systematic error in the case of La Silla.

5. – Possible improvements of ECMWF forecasts by means of Kalman filter post-processing

Given the absence of advanced observative meteorological facilities (radar, wind profilers etc.) at the sites of the ESO observatories, we have tested two standard tools for possible improvement of the local forecast: statistical "post-processing" and the nesting of a LAM into the ECMWF global forecast model. Results of the LAM application to selected test-cases will be presented in the next section, while in this section we discuss the improvement obtained using the Kalman filtering procedure described in detail in Appendix A.

Use of Kalman filtering as a post-processing tool is aimed at filling, by means of statistical information, the gap between large scale dynamics computed by the numerical model and local evolution of the meteorological variables. This is usually done by applying some scheme of statistical regression between long series of observations of the variables of interest at the considered site and some so-called "predictors" (usually the numerically predicted values of the same variables at neighbor points in the numerical grid).

It is well-known that the output of a numerical model is affected by systematic errors which can be due, for example, to errors in the physical parameterizations and/or initialization and to unresolved sub-grid processes and local effects. For this reason, in the last years operational meteorological centers throughout the world have applied post-processing methods aimed at filtering out noise and systematic errors by using some sort of dynamical self-adaptive regression method (see, for example, [6–9]). One of the best known of these methods is the Kalman filtering procedure, which we briefly review in Appendix A in relation to the problem of post-processing the output of a meteorological model.

With the initialization described in Appendix A, we have applied the Kalman algorithm to the post-processing of ECMWF forecasts of 2-meter temperature at 12GMT, for Paranal and La Silla. The validation is made using surface observations in these two sites, during the period from 1989 to 1993, and ECMWF forecast and analysis interpolated at the station: vertically (at the altitude) and horizontally (at the position).

Results are shown in Table X: it can be seen that in comparison with the persistence, which gives an indication of the daily variability of the parameter, we obtain an improvement of the absolute mean error of about 0.5 degrees. Correlation coefficients are high and equal to those of the analysis. In the table we do not show mean absolute errors of analysis and forecast because they are affected by very large systematic errors not comparable with those of the post-processing. In particular, during daytime, 2-meter temperatures, obtained by interpolation of vertical profiles of temperature, are systematically lower than measured temperatures because of the contribution of the ground radiation to air temperature.

For this reason we have also tried to make the same comparison by using forecast and analysis of 2-meter temperature, in place of vertically interpolated values. The results are much worse than those in Table X. Essentially this is due to the fact that the parameterization of ground physics and planetary boundary layer of the ECMWF GCM is referred to properties of the ground (as orography, roughness, soil temperature, etc.) that are representative for the grid resolution of the model, so it is completely inadequate to correctly describe local dynamics at station sites. We tried also to use the value of 2-meter temperature at the ECMWF grid-point nearest to site location, in place of the horizontal interpolated value. In this case results were only little worse than those in Table X.

Another point to clarify is the role of first guess in the procedure (refer to Appendix A for symbols and notation). Values of C_0 reflect the initial knowledge of the system vector $\tilde{S}_1$. As this is never known exactly a priori, it is better to give to elements of the matrix C_0, W_1, V_1 sufficiently large values. Since all these variances are re-calculated at each time step, after a period, necessary to accumulate enough statistics, the system will "forget" the initial values. In the same way the initial value of the system vector will converge to a value that is "optimal" (in a least squares way) for the system. In order to check these statements we made a series of experiments with different first guess values of all the parameters mentioned above without obtaining appreciable differences in the results. For this reason we have chosen as first guess the one discussed in Appendix A, which has the advantage of being more intuitive and natural than many others. The only value of $\tilde{S}_t$ that varies appreciably in time is the first term which is essentially the constant term of the regression equation.

To evaluate the improvement due to the KF procedure with respect to zero knowledge (persistence), we summarize skills of both procedures in Table 11. For brevity, and because results for La Silla are very similar, only the results for Paranal are shown. Explanation of the table is in the caption but here we want to give an example of interpretation of the numbers contained in it in order to better clarify its content.

Looking at the last column of Table XI, we see that number of times the error of the KF procedure in forecasting two-meter temperature within an absolute error of 1 degree is 159+158=317 (55% of the total number of data used in Table XI). At the same time the number of times the error of the persistence is within the same range is 112+117=229 (40%). The forecast of the external temperature with an error less than 1 degree is a desired condition for astronomers, as it allows to preset the air conditioning in the telescope room, assuring that the mirror is completely operational at the opening of the dome. This discussion shows then that there is a sensible added value in using the KF procedure, quantifiable as the increase in the number of operational days 317-229=88 (15%).

The same procedure was applied to 10-meter wind speed giving the results shown in Table X, but some clarification is needed about the contents of this table and about comparison of wind components in general. First of all it should be clear that any direct comparison between wind components cannot be proposed because wind at ground level is heavily influenced by orographic effects ($\sim$ 1 m) that are several orders of magnitude lower than that in the orography of the model ($\sim$ 10 km). This is reflected especially in wind direction and for this reason we have not considered it. Wind speed may be less influenced by these, in any case strong, orographic effects and so for this variable we performed the same analysis as already performed for 2-meter temperature.

As can be seen in Table X the mean absolute error of 10-meter wind speed after Kalman filtering is about 3 m/s which must be compared with the mean value of the field, 7 m/s. Correlations are about 0.6: small, but in any case, greater than correlations of forecast and analysis. It is evident, however, that analysis and forecast are very poor because correlations are small and not statistically significant: for example, in La Silla correlation of forecast is greater than correlation of analysis.

To evaluate the relevance of these results to the specific aims of the ESO astronomers, we show in Tables XII and XIII contingency between measured and Kalman filtered data for five classes of wind speed from 5 m/s to 25 m/s in Paranal and La Silla. The wind belonging to the two classes from 0 to 10 m/s, which give no problems to the astronomers, are forecast correctly 75% of the time and are present 85% of the time. The two classes between 10 and 20 m/s are forecast correctly 8% of the time and are present 15% of the time. More interesting is the analysis of the class of wind speed greater than 20 m/s on occurrence of which the telescope cannot be used. It must, first of all, be noted that the number of occurrences of this event is very low: only 1 in La Silla and 6 in Paranal. In any case the corresponding post-processed values are not in good agreement. From these results it seems that forecasts of ground wind speed from a GCM would not be useful to ESO astronomers.

As already mentioned, ECMWF analysis during the period 1992-93 shows consistently large errors, probably due to the Quintero balloon data. To check the influence of this anomalous period in our analysis of the Kalman filtering procedure, we repeated all the calculations described in this section using only data for the period 1989-1991. The results were very similar to those reported in Table X for the whole period 1989-93. Some differences were observed for correlations of analysis and forecast of temperature:

in particular we have found for La Silla in the shorter period $r_{for} = 0.87$ and $r_{ana} = 0.91$ which are greater than the corresponding values in Table X. This could be an indication that the analysis, and therefore the forecast, were better during the shorter period. We remark, in conclusion, that performances of the Kalman filtering procedure are not influenced by the apparent worsening of the analysis during the period 1992-93.

6. – Discussion of LAM test cases

In this section we discuss results of application of a LAM in a restricted region including Chile, giving particular emphasis to forecasts of 2-meter temperature and 10-meter wind at Paranal and La Silla. A general description of the equations and some detail of the numerical aspects and parameterizations of the LAM we used can be found in Appendix B and in Marrocu *et al.* [10].

The fields used as initial condition in order to solve the equations of motion for atmospheric dynamics are obtained from a GCM, provided by the ECMWF in our case. Generally one has to deal with an analysis obtained by interpolating observations from various sources (such as ships, aircrafts, balloons, satellites and meteorological stations) and "initializing" such interpolated fields in order to filter out dynamical noise and to make these fields consistent with the GCM equations of motion [11]. The ECMWF initialized analysis and the model forecasts (06 to 120 hours at 6-hour intervals, 120 to 240 hours at 12-hour intervals) are disseminated to regional meteorological centers via a dedicated telecommunication network, at four analysis times (18 GMT preceding day, 00, 06, 12 GMT current day).

A LAM solves essentially the same equations of a GCM, but inside a limited open region and, therefore, it needs a continuous updating of the physical fields at the boundary of the region of integration. The so-called time-dependent lateral boundary conditions used in an operational system are derived from the forecasts of the GCM. Since these boundary conditions must be updated at every time-step (typically a few tens of seconds) of the integration procedure of the LAM, forecasts of the GCM (typically characterized, any how, by longer integration time-steps) must be interpolated in time between the two nearest forecast hours available.

As initialization and boundary conditions are dynamically consistent with the equation of the GCM, their interpolation to the grid of the LAM (pre-processing) will produce in the first hours of integration spurious non-physical waves which are, however, rapidly damped also by using a proper relaxation procedure between LAM forecasts and lateral boundary conditions.

Since LAM forecasts are obtained at constant levels of the vertical coordinate used (pressure, potential temperature, sigma coordinate, etc.), the output of the LAM must be "post-processed" by using physical parameterizations in order to produce numerical values at given altitudes and at the ground [12]. In any case, since the orography of the model is completely different from the real one and the LAM grid square is at least 100 km^2, further statistical post-processing procedures must be applied in order to eliminate systematic geometrical errors.

We made some simulations in order to evaluate LAM performance on the region of interest. All the numerical integrations were performed using a horizontal grid with resolution of one sixth of degree (latitude step $\sim$ 17 km, longitude step $\sim$ 15 km), on 93 longitude grid-points (from 279.75E to 295.08E), 100 latitude grid-points (from 18.58S to 35.25S) and 20 vertical levels on sigma coordinates. In Figure 2 the LAM orography, obtained from an NCAR (National Center for Atmospheric Research) digital elevation map, is plotted and compared with that at 0.5625 degree resolution of ECMWF model, while in Figure 3 the corresponding W-E sections of orography at Paranal latitude are shown.

The time-step used for the integration of the dynamic equations was 15 seconds, while a 10-minute time step was used for the contributions of physical parameterizations like radiative transfer, large scale precipitation, and vertical convection.

Initial and boundary conditions needed for LAM simulations were obtained from interpolation of the ECMWF analysis at the new one sixth degree resolution on each of the 20 sigma levels. After integrations, we used a standard "post-processing" procedure to interpolate the model variables from sigma levels to standard constant pressure surfaces. In the "post-processing" phase some new "surface" fields, such as the 2-meter air temperature and 10-meter wind, were also computed.

We produced 5 different test cases, each one 72 hours long, in order to estimate LAM capability to predict local observations at the Paranal and La Silla meteorological stations:

- 3 test cases, starting at 00GMT on 04/03/93, 12/05/93 and 20/05/93, were selected during cyclogenetic perturbations. In particular, during the 2 test cases of May, the southern hemisphere jet-stream was very close to the southern boundary of the domain used for the integration. In Figure 7a,b,c the 500 mb geopotential heights after 1 hour, 24, 48, and 72 hours of integration for each simulation are plotted.

  The case starting on 04/03/93 is a case of weak penetration of a middle-latitude trough (the core of the trough appears in the LAM domain only in the last day of integration). Analysis shows interaction of the westerly jet stream with the Andes, with an initial tendency to flow splitting (particularly pronounced on March 4) and strong generation of orographic waves on successive days in which the reinforced jet-stream flows across the Andes; as already mentioned, only in the last day does the extratropical cyclone enter the LAM domain.

- 1 test case, starting at 00GMT on 17/08/93, was selected as in that period a great difference between the temperature observed in Paranal and that of ECMWF analysis and forecast was found. In Figure 7d the 500 mb geopotential height maps for this test case are plotted.

- 1 test case, starting at 00GMT on 18/03/93, was selected as representative for periods in which nothing of meteorological interest happens, as can be seen from Figure 7e where the 500 mb geopotential height maps are plotted.

Since the Paranal and La Silla stations do not coincide with a point of the grid used in the LAM simulations, the numerical variables in neighboring grid-points were used to estimate local values by means of bilinear interpolation (the coordinates and geopotential height of these points can be found in Table I).

We have compared 2-meter air temperature and 10-meter wind velocity measured at meteo-stations in Paranal and La Silla with numerical values from the LAM forecasts in the above 5 test cases. In these comparisons we have used two kinds of fields obtained from LAM variables:

- "surface" values: 2-meter temperatures and 10-meter winds were computed in the "post-processing" phase using standard parameterization of the planetary boundary layer, such as logarithmic vertical profiles; the "surface" values used in the comparisons results from bilinear interpolation of numerical values in the four nearest points to Paranal or La Silla.

- "upper air" values: we first computed 4 values, at the same geopotential height as the meteorological station, using a cubic spline interpolation of the nearest 4 profiles in the LAM grid, and then we estimated the "upper air" values by bilinear interpolation of the above 4 values.

LAM integrations were saved every hour (total of 72 hours for each test case), but we compared only values during the Chilean night from 00GMT to 10GMT (that is the time-interval of interest for astronomical activity). In Figures 8 and 9 values of 2-meter temperature and 10-meter wind speed observed in Paranal and La Silla are plotted versus "surface" or "upper air" values obtained by LAM forecast as discussed above. In Tables XIV and XV errors of 2-meter temperature in Paranal and La Silla from ECMWF analysis and from the LAM "upper air" forecast are compared. It must be pointed out that great errors in ECMWF fields are often reflected also in the LAM forecast.

We do not analyze in further detail the test case starting on 17/08/93, since we found on 18/08/93 a difference of about 8 degrees between the 2-meter temperature observed in Paranal and the values estimated by ECMWF analysis. For the remaining cases the LAM predictions of temperature seem to be in general quite good, in particular if the "upper air" values are used. Also the wind velocity forecast is better when using "upper air" values rather than "surface" values. This suggests that, for operational use of a high resolution meteorological numerical model, improvement and specialization of post-processing techniques are certainly needed.

In particular, in the case starting on 04/03/93, the LAM forecast tends to underestimate the initial jet-splitting and to overestimate orographic wave production; both features may be due to inadequate observational coverage and initialization, rather than to inadequacy of the LAM itself. The final penetration of the trough-core is well simulated, but this success is probably due to the fact that the cyclonic feature is "forced" through the southerly boundary by the boundary conditions.

Some features are in common to all the numerical experiments performed with the LAM; the most relevant are:

- the horizontal extent of the LAM domain is too small in comparison with the size of extratropical cyclones occasionally deepening equatorward up to the point of hitting the geographic area in question;

- observations are so sparse (in particular over the Pacific Ocean) that it is presently impossible both to fully exploit and adequately check the LAM performance;

- strong trains of gravity-inertia waves appear in all the LAM forecasts in the lee side of the Andes. These are probably in part real (the improved resolution allows description of gravity-inertia dynamics which is smoothed in the ECMWF analysis and forecast) and in part due to the effect of sub-grid (with respect to ECMWF grid-size) orography in the initialization procedure.

7. – Summary and concluding remarks.

We summarize here the more relevant results of our comparative study of some ECMWF fields, such as cloud cover, temperature, and wind, against the corresponding observations at the sites of astronomical facilities in Chile.

Different criteria of estimation of cloud cover in Paranal from ECMWF analysis have revealed only a marginal improvement with respect to the simple assumption, justified by the geographical location of Paranal, that the sky is always clear. Due to the poor results obtained with the ECMWF analysis, the comparisons were not extended to the 24/48-hour forecast fields of the ECMWF dataset.

By comparing vertical profiles of temperature from radiosoundings at Antofagasta and Quintero with profiles obtained from the ECMWF analysis we found a mean absolute error of about 2.5 degrees. The corresponding comparison performed with vertical profiles from 24-hour ECMWF forecasts revealed a mean absolute error of about 3 degrees. This difference of only half a degree with the error of the analysis means that a large part of the temperature errors in the vertical profiles extracted from the ECMWF 24-hour forecast is due to inadequate knowledge of the "initial condition", namely the analysis, for the ECMWF general circulation model.

Comparisons between vertical profiles of wind velocities from radiosoundings with the corresponding ECMWF fields were, instead, very poor, both in terms of correlations and of absolute errors. Further investigations, focused on extreme wind conditions (wind velocities higher than 10 m/s) revealed a mean absolute error of about 50% in the 24-hour ECMWF forecast profiles against an error of 25% for the corresponding analysis profiles.

Two-meter temperatures measured at the meteorological stations at Paranal and La Silla were compared with the corresponding estimates based on the the 2-meter temperature fields and upper-air 3D temperature fields from the ECMWF analysis. The comparison with the 2-meter temperature of ECMWF analysis have revealed very large errors, even larger than errors of the persistence model. The cause of these discrepancies can be essentially attributed to the smoothed representation of the orography in the ECMWF grid: the geopotential height of the ECMWF model at the location corresponding to Paranal and La Silla is very different from the real one.

Better results were instead found by interpolating the ECMWF 3D temperature field at the same geographical location and altitude of the meteorological sites: in such a way a mean absolute error of about 1.3 degrees at Paranal and 1.7 degrees at La Silla was found. Since a systematic error was still present, a statistical post-processing tool, based on the Kalman filter algorithm, was applied to the 24-hour forecast. The Kalman filter was designed to take into account not only the 24-hour forecast, but also the information coming from actual analysis and local observation. It was able to further reduce the errors in predicting 2-meter temperature to about 1.1 degree at Paranal and 1.6 degree at La Silla.

Finally, some test cases of LAM simulations in Chile were discussed with the aim to evaluate the gain, in terms of 2-meter temperature errors and 10-meter wind velocity errors, when increasing the spatial resolution from the ECMWF grid to a finer grid. Results of this investigation show that also for LAM forecasts the great part of larger errors are due to inadequacy of the initial conditions coming from the ECMWF analysis. As we have already found for the other kind of comparisons, the LAM forecast of temperature is better represented than the wind field.

We did not apply the Kalman filter post-processing to the LAM forecasts due to the small amount of data produced with the simulations, but it is clear that it can be successfully applied also to the output of an operative LAM.

Before closing, we want to remark, once again, that analysis and forecasts of all numerical variables considered here suffer from the problem of "downscaling"; in fact, the marginal statistical difference between analysis and forecast is, in itself, an indication that the problems in question already arise in the initialization phase of the forecasting procedure; in this phase the meteorological fields must be interpolated locally starting from observations collected from a very coarse observational network.

Such downscaling problems can be addressed with different approaches, not necessarily mutually exclusive:

- static (mass conserving adjustment) and statistical (Kalman filter) postprocessing techniques can prove useful in correcting part of the systematic errors, but cannot improve the local forecast in cases of dramatic changes of meteorological fields which are, however, of great importance in the management of astronomical facilities;
- limited area models have the potential for improving the local forecast, but they have to be fed with adequate initial data and extended in space to the middle-latitude westerlies in order to capture the dynamics of troughs that only on occasion penetrate the tropical circulation, but, on these same occasions, cause severe changes in the local weather at the sites of interest.

In summary, meteorology can help astronomy, as many other human activities, but at the expense of a major effort in setting up a better observative activity. The specific application discussed in this paper does not seem to justify, on its own, the big effort required in order to guarantee, in particular over the Pacific Ocean, an adequate observative cover. It seems probable however that the development of Meteorology (in particular through remote sensing) will, independently, provide in time the needed information and, as a consequence, make our methodological assessment useful.

Appendix A.

Kalman filter post-processing

A review of a Kalman filtering procedure is presented here in terms useful for application as a "post-processing" tool to the output of meteorological numerical models.

Let us suppose that the parameter to be predicted at time t is the ground temperature T and all we know are the observations of T, the analysis and the forecasts up to time t. It is then intuitive that systematic errors, if present, would be reduced if we were able to use as much as possible the available information about the system. In the traditional multiple linear regression methods, we write the so-called observation equation in which the multiregressive parameters are fixed. The Kalman filter procedure adds to the observation equations another linear equation (the system equation) in which the multiregressive parameters at time t are evaluated as a function of those at time $t-1$ and the previous error of forecast. The Kalman filter algorithm is then obtained solving a least squares problem for these two equations.

More specifically, let us suppose that the equations have the following form:

$$\left\{ \begin{array}{ll} y_t = P_t^T S_{t-1} + v_t & \text{is the observation equation} \\ S_t = S_{t-1} + w_t & \text{is the system equation} \end{array} \right. \tag{A.1}$$

where v_t is the normal error of the post-processing scheme with variance V_t, and w_t is a vector of random errors with zero mean and variance matrix $W_t = E[w_t w_t^T]$. In our specific case y_t is the temperature at ground level (the predictand), P_t is the vector whose elements are all the significant physical parameters measured or calculated up to time

t (Predictors), and S_t is a vector containing the multiregression parameters (the State vector). The matrices involved in this model can be explicitly evaluated imposing, in addition to the hypotheses already stated concerning v and w, that the mean error of S_t should be zero and that the error variance should be minimized. Assuming to know an estimate of S_0 and of its variance C_0, the Kalman filter algorithm can be stated as follows:

$$
\text{(A.2)} \qquad \begin{cases} \tilde{y}_t = P_t^T \tilde{S}_{t-1} & \text{is the forecast of } y_t^{obs} \\ e_t = y_t^{obs} - \tilde{y}_t & \text{is the forecast error} \\ R_t = C_{t-1} + W_t & \\ \tilde{Y}_t = P_t^T R_t P_t + V_t & \text{is the variance of } \tilde{y}_t \text{ given } P_t \\ G_t = R_t P_t / \tilde{Y}_t & \text{is the gain} \\ \tilde{S}_t = \tilde{S}_{t-1} + G_t e_t & \text{is the new vector of state} \\ C_t = R_t - G_t \tilde{Y}_t G_t^T & \text{is the new error variance of } \tilde{S}_t \end{cases}
$$

The algorithm is recursive and it is computationally not expensive since it involves only matrix addition and multiplication (no inversion) .

In the following we describe and discuss in more detail how to initialize the procedure and what are the elements of the vectors S_0 and P_t in order to apply such Kalman filtering procedure to post-process ECMWF forecasts of 2-meter temperature and 10-meter wind speed in La Silla and Paranal.

First of all it must be clear that before starting the procedure we do not have any historical knowledge of the system. But this is not a real disadvantage, on the contrary it is one of the most useful properties of the Kalman filter. In order to initialize the procedure we have to choose a set of significant predictors to use as elements of the vector P_t. For the 2-meter temperature it is quite natural to take as predictors: the measured value y_{t-1}^{obs}, the ECMWF analysis y_{t-1}^{ana} and the forecast value y_t^{for}. The value of y_t^{for} is the output of the ECMWF model at time $t-1$ which is the prediction for time t. The explicit form of the vector chosen in our application was: $P_t^T = [1, y_t^{for}, y_{t-1}^{ana} y_{t-1}^{obs}]$ where the first element, constant and equal to 1, is related to the constant of the multiregression equation.

Next, we need a first guess for the multiregression parameters of the state vector $\tilde{S}_0$ and its variance C_0. This choice is not critical if it is performed in a vast range of reasonable alternatives. We picked: $\tilde{S}_0^T = [0, 1, 0, 0]$ and $Diag(C_0) = [4, 1, 1, 1]$ as diagonal elements of C_0. This means that at beginning, not having any other information, we suppose that the forecast of the model is exact, $\tilde{S}_0(2) = 1$, but at the same time we presume that the standard deviations of this coefficient, as of the last two, are equal to one ($C_0(2,2) = C_0(3,3) = C_0(4,4) = 1$).

$C_0(1,1) = 4$, related to $\tilde{S}_t(1)$, means that the additive constant in the multiregression formula, at time $t = 1$, is 0, but with a standard deviation of $\sqrt{4} = 2$ degrees.

W_t and V_t must be known at the beginning, but can be updated after initialization, using a short past series of about 10-14 days of observations and forecasts. As initial values we have chosen $V_1 = 4$ and $W_1 = C_0$.

Appendix B.

Limited area model description

In this Appendix we give a more technical description of the LAM used in this work. LAM is a meteorological model which integrates in time the full three-dimensional set of partial differential equations that constitute the so-called primitive equations, using centered finite differences. The σ coordinate, defined as $\sigma = p/p_s$, is used as vertical coordinate, where p is the pressure in the upper air and p_s is corresponding pressure at ground. The vertical discretization is of the Lorenz type, with the vertical velocity defined at intermediate levels with respect to those where the prognostic variables are defined. The horizontal discretization of the model equations is done using standard centered finite differences on the staggered Arakawa C-grid, using latitude (ϕ) and longitude (λ) as independent horizontal spatial variables. The prognostic variables are the horizontal wind components (the zonal wind u and the meridional wind v), the potential temperature (θ), the specific humidity (q) and the surface pressure (p_s). The equations for the evolution of these five variables consist of the two components of the horizontal momentum equation:

$$\frac{\partial u}{\partial t} = (f+\xi)v - \dot{\sigma}\frac{\partial u}{\partial \sigma} - \frac{R_d T_v}{a h_\lambda}\frac{\partial ln p_s}{\partial \lambda} - \frac{1}{a h_\lambda}\frac{\partial}{\partial \lambda}(\Phi + E) + K_u + F_u, \tag{B.1}$$

$$\frac{\partial v}{\partial t} = (f+\xi)u - \dot{\sigma}\frac{\partial v}{\partial \sigma} - \frac{R_d T_v}{a h_\phi}\frac{\partial ln p_s}{\partial \phi} - \frac{1}{a h_\phi}\frac{\partial}{\partial \phi}(\Phi + E) + K_v + F_v, \tag{B.2}$$

the continuity equation integrated over the vertical coordinate, yielding an equation for the surface pressure p_s:

$$\frac{\partial p_s}{\partial t} = -\int_0^1 \nabla\cdot(\mathbf{V}_h \frac{\partial p}{\partial \sigma}) d\sigma, \tag{B.3}$$

the thermodynamic equation:

$$\frac{\partial \theta}{\partial t} = -\frac{u}{a h_\lambda}\frac{\partial \theta}{\partial \lambda} - \frac{v}{a h_\phi}\frac{\partial \theta}{\partial \phi} - \dot{\sigma}\frac{\partial \theta}{\partial \sigma} + K_\theta + F_\theta, \tag{B.4}$$

and the equation for the conservation of water:

$$\frac{\partial q}{\partial t} = -\frac{u}{a h_\lambda}\frac{\partial q}{\partial \lambda} - \frac{v}{a h_\phi}\frac{\partial q}{\partial \phi} - \dot{\sigma}\frac{\partial q}{\partial \sigma} + K_q + F_q. \tag{B.5}$$

The diagnostic equations are the hydrostatic equation:

$$\frac{\partial \Phi}{\partial \sigma} = -\frac{R_d T_v}{p}\frac{\partial p}{\partial \sigma} \tag{B.6}$$

and the continuity equation which is solved at each vertical level for the vertical velocity ($\dot{\sigma}$):

$$\dot{\sigma}\frac{\partial p}{\partial \sigma} = (1 - \frac{\partial p}{\partial p_s})\frac{\partial p_s}{\partial t} + \int_\sigma^1 \nabla\cdot(\mathbf{V}_h \frac{\partial p}{\partial \sigma}) d\sigma. \tag{B.7}$$

The system of equations is closed with the ideal law of gases and the standard definitions of the potential (θ) and virtual (T_v) temperatures. In the above equations,

(h_λ, h_ϕ) are scale factors ($(h_\lambda, h_\phi) = (cos\phi, 1)$ in spherical coordinates), R_d is the "gas constant" for dry air, a is the mean radius of the earth, $f = 2\Omega sin\phi$ is the Coriolis parameter, ξ is the vertical component of the relative vorticity, Φ is the geopotential, E is the kinetic energy of the horizontal motion, $(F_u, F_v, F_\theta, F_q)$ are the contributions to the tendencies coming from the parameterization of the physical processes (such as solar and terrestrial radiation, convection, cloud processes, surface friction and heat flux, etc), and $(K_u, K_v, K_\theta, K_q)$ represent the contribution from the horizontal and vertical diffusion. In particular, a second order horizontal diffusion operator is applied to the divergence of the horizontal velocity field, to reduce the growth of gravity waves and the development of unbalanced components giving rise to "noise" in the meteorological fields. The divergence damping is obtained by adding to the right hand side of the two horizontal components of the momentum equation a term proportional respectively to $\partial/\partial\lambda$ and $\partial/\partial\phi$ of the divergence itself. The horizontal diffusion consists in a fourth order operator applied to all prognostic variables except p_s to limit energy concentration on the grid points.

The use of the potential temperature θ as one of the prognostic variables avoids the need of a spatially dependent diffusion coefficient, especially when the σ surfaces become rough over complex terrain, and it also simplifies the thermodynamic equation (the "$\omega - \alpha$ term" does not appear).

The time integration scheme is, in the present version, a leapfrog scheme, which reads for the generic variable $H(x, t)$:

$$H(x, t + dt) = H(x, t - dt) + 2dt\frac{\partial H(x, t)}{\partial t}, \tag{B.8}$$

where $\partial H/\partial t$ is the tendency of H calculated at position x and at time t; an Asseling time filtering is applied to suppress the computational mode due to the three time-level scheme, which reads again for the variable $H(x, t)$:

$$H(x, t) = H(x, t) + \alpha(H(x - dt) - 2H(x, t) + H(x, t + dt)), \tag{B.9}$$

where α is about 0.05.

The lateral boundary conditions are implemented with the relaxation scheme of [13] and [14] applied to all prognostic variables.

The physical package describes the contribution to the evolution of the prognostic variables due to those processes characterized by typical spatial scales smaller than the model resolution. These processes include vertical diffusion, dry adiabatic adjustment, soil water and energy balance, large scale precipitation and evaporation, moist convection and radiation. For more details, the interested reader can refer directly to the references [15–20].

The time stepping procedure consists in a first integration in the interior points of the domain, i.e. excluding the boundary lines, that is performed to calculate the tendencies due to the dynamics, the vertical diffusion, the radiation and surface processes. At following time-steps the prognostic variables are advanced at time $t+dt$ with the leapfrog scheme, then the variables at time t are filtered using their values at times $t - dt$ and $t + dt$. The changes in the prognostic variables due to the dry and moist convective adjustment and large scale rain are calculated and added with an Euler scheme. At this point, variables are first shifted in time (from t to $t - dt$ and from $t + dt$ to t) and then the contribution from the horizontal diffusion and divergence damping is added using an Euler scheme. Finally, the prognostic fields are relaxed to the updated boundary values, completing the time step.

* * *

The authors acknowledge the Sardinian Regional Authorities for additional support. We are grateful to Prof. Roberto Benzi for early guidance in the project, and to Marc Sarazin for interesting and helpful discussions.

REFERENCES

[1] Sarazin M., *Predicting observing conditions at ESO observatories - reality and perspectives.* Technical Report 89, ESO (1997).

[2] Benzi R., Deidda R., Marrocu R. and Speranza A., *Feasibility Study of a meteorological Prediction Model for ESO Observatories in Chile (phases A1, A2).* Technical Report CRS4-TECH-REP-96/51, CRS4, Cagliari, Italy (1996).

[3] Deidda R., Marrocu R. and Speranza A., *Feasibility Study of a meteorological Prediction Model for ESO Observatories in Chile (phases A3, A4).* Technical Report CRS4-TECH-REP-97/38, CRS4, Cagliari, Italy (1997).

[4] Deidda R., Marrocu R. and Speranza A., *Proposal for a system of operational forecast finalized to Meteorological Prediction for ESO Observatories in Chile (phase A5).* Technical Report CRS4-TECH-REP-97/56, CRS4, Cagliari, Italy (1997).

[5] Ratto C.F., Festa R., Romeo C., Frumento O.A. and Galluzzi M., *Environ. Software*, **9** (1994) 247.

[6] Simonsen C., *Self adaptive model output statistics based on Kalman filtering.* Technical Report, Danish Meteorological Institute (1991).

[7] Cacciamani C. and De Simone C., *Minimum temperature forecasts at the regional meteorological service of the Emilia Romagna region (North Italy) by the application of the Kalman filter technique. ECMWF Newsletter*, **60** (1992) .

[8] Persson A., *Kalman filtering. A new approach to adaptive statistical interpretation of numerical meteorological forecasts. ECMWF Newsletter*, **46** (1989) .

[9] Cattani D., *Application d'un filtre de Kalman pour adapter les temperatures a 2 metres fournies par le modele ECMWF aux station meteorologiques de la Suisse.* Technical Report No. 175, ISM (1994).

[10] Marrocu M., Scardovelli R. and Malguzzi P., *Parallel Computing*, **24** (1998) 911.

[11] Daley R., *Rew. Geoph. Space Phys.*, **19** (1981) 450.

[12] Geleyn J.-F., *Tellus*, **40A** (1988) 347.

[13] Davies H. C., *Quart. J. Roy. Meteor. Soc.*, **102** (1976) 405.

[14] Kållberg P., *Test of lateral boundary relaxation scheme in a barotropic model.* Technical Report No. 3, ECMWF (1977).

[15] Buzzi A., Fantini M., Malguzzi P. and Nerozzi F., *Meteorol. Atmos. Phys.*, **53** (1994) 137.

[16] Charnock H., *Quart. J. Roy. Meteor. Soc.*, **81** (1955) 639.

[17] Emanuel K. A., *J. Atmos.Sci.*, **48** (1991) 2313.

[18] Geleyn J.-F. and Hollingsworth A., *Contrib. Atmos. Phys.*, **52** (1979) 1.

[19] Louis J.-F., Tiedtke M. and Geleyn J.-F., *A short history of the PBL parameterization at ECMWF*, in *Proceedings from the ECMWF Workshop on Planetary Boundary Layer Parameterization (25-27 November 1981).* 1982, pp. 59-79.

[20] Ritter B. and Geleyn J.-F., *Mon. Wea. Rev.*, **120** (1992) 303.

TABLE I. – *Geographical information about La Silla and Paranal, where surface meteorological stations are located, about Antofagasta and Quintero, where balloons for radiosounding are released, and about nearest points (NP) in the ECMWF grid.*

	meteo stations		radiosoundings	
	La Silla	Paranal	Antofagasta	Quintero
Altitude	2400 m	2636 m		
Longitude	70W42	70W24	70W26	71W32
Latitude	29S16	24S37	23S26	32S47
Distance from coast	40 km	12 km		
NP altitude	2680 m	1835 m	1294 m	1182 m
NP longitude	70W52.5	70W18.75	70W18.75	71W26.25
NP Latitude	29S15	24S45	23S37.5	32S37.5
Distance from NP	17 km	17 km	25 km	20 km

TABLE II. – *The ESO local data base: cloud cover in Paranal; 2-meter air temperature (T2m) and 10-meter wind (W10m) from meteorological stations in Paranal and La Silla; vertical profiles of relative humidity (RH), temperature (T) and wind (W) from radiosounding at Antofagasta and Quintero.*

Variables	observed period	site & frequency
cloud cover	from 2/6/92 to 31/12/93	Paranal (every 2 hours: at 0,2,4,6,8,10)
Meteo stations: T2m and W10m	from 1/10/89 to 30/09/94	La Silla (every 20 min) Paranal (every 20 min)
Radio soundings: RH, T and W	from 1/1/85 to 30/06/94	Antofagasta (daily at 12) Quintero (daily at 0,12)

TABLE III. – *Average of daily correlations (r), mean errors ($\overline{\epsilon}$), and mean absolute errors ($\overline{|\epsilon|}$) of comparisons between vertical profiles from radiosoundings in 1993 and the corresponding ECMWF analysis profiles on the nearest point (NP profile) or those interpolated in the sites (INTERP profile). RH is relative humidity, T is temperature and W is wind velocity.*

		NP profile			INTERP profile		
site & hour	field	r	$\overline{\epsilon}$	$\overline{\vert\epsilon\vert}$	r	$\overline{\epsilon}$	$\overline{\vert\epsilon\vert}$
Antofagasta	RH (%)	0.426	0.48	17.12	0.501	1.46	16.55
at	T (C)	0.985	1.50	2.66	0.986	1.39	2.54
12 GMT	W (m/s)	0.913	-0.21	2.17	0.918	-0.36	2.11
Quintero	RH (%)	0.664	0.29	13.82	0.693	0.91	13.46
at	T (C)	0.994	0.84	2.24	0.994	0.73	2.15
00 GMT	W (m/s)	0.862	0.34	3.97	0.862	0.46	3.98
Quintero	RH (%)	0.650	-1.72	13.94	0.694	-0.85	13.42
at	T (C)	0.990	1.29	2.59	0.991	1.15	2.48
12 GMT	W (m/s)	0.851	0.96	4.27	0.854	1.04	4.26

TABLE IV. – *Percentages of success for cloud cover estimations at 0 and 6 GMT for different estimation criteria. E1 is the "optimistic hypothesis", that consists in guessing that the sky is always clear. E2 is the skill of the total cloud cover of the ECMWF analysis, E3 is that obtained by the ECMWF supplementary total cloud cover and E4 is obtained using humidity upper-air field.*

	E1	E2	E3	E4
0 GMT	80%	80%	82%	81%
6 GMT	81%	82%	83%	85%

TABLE V. – *Correlations (r), mean errors ($\overline{\epsilon}$), and mean absolute errors ($\overline{|\epsilon|}$) between 2-meter temperature observations at La Silla and Paranal in 1993 and different estimates from ECMWF analysis at synoptic hours (H). Columns A: results of the "24h persistence model". Columns B: results using the 2-meter temperature field of ECMWF analysis; estimates are derived from bilinear interpolation of the the 4 neighboring points in ECMWF grid. Columns C: results using the upper-air temperature profiles of ECMWF analysis; values are interpolated horizontally (at station position) and vertically (at station altitude).*

		A		B			C		
Station	H	$\overline{\|\epsilon\|}$	r	$\overline{\epsilon}$	$\overline{\|\epsilon\|}$	r	$\overline{\epsilon}$	$\overline{\|\epsilon\|}$	r
	0	1.4	0.76	-1.2	1.8	0.67	0.8	1.3	0.81
Paranal	6	1.5	0.77	2.5	3.1	0.49	0.8	1.3	0.86
	12	1.6	0.77	2.1	2.8	0.53	0.2	1.0	0.90
	18	1.4	0.71	-5.8	5.8	0.73	-0.9	1.4	0.80
	0	1.9	0.79	4.3	4.3	0.85	-0.3	1.1	0.93
La Silla	6	1.9	0.80	8.8	8.9	0.58	-0.2	1.1	0.94
	12	1.7	0.79	8.1	8.1	0.59	-1.5	1.8	0.93
	18	1.8	0.81	-0.3	1.6	0.85	-1.4	1.7	0.93

TABLE VI. – *Correlations (r), mean errors ($\overline{\epsilon}$), and mean absolute errors ($\overline{|\epsilon|}$) of comparison between observations of 10-meter wind speed (W10m) at La Silla and Paranal in 1993 and bilinearly interpolated values from 10-meter wind speed of ECMWF analysis at synoptic hours (H). MV is the mean value.*

Station	H	MV	$\overline{\epsilon}$	$\overline{\|\epsilon\|}$	r
	0	6.5	-2.7	3.3	0.48
Paranal	6	6.8	-2.5	3.3	0.57
	12	6.7	-2.2	3.1	0.68
	18	8.6	-4.5	4.8	0.50
	0	4.4	-0.6	2.1	0.67
La Silla	6	5.6	-1.8	2.7	0.65
	12	5.6	-2.0	3.0	0.60
	18	5.6	-1.6	2.3	0.59

TABLE VII. – *Average of daily correlations (r), mean errors ($\overline{\epsilon}$), and mean absolute errors ($\overline{|\epsilon|}$) between vertical profiles from radiosoundings in 1993 and the corresponding profiles obtained by bilinear interpolation of grid values from ECMWF analysis and 24/48-hour forecast. The averages of absolute errors on temperature profiles are also shown for the whole 89-93 period.*

		1993			89-93
radiosoundings	ECMWF profile	r	$\overline{\epsilon}$	$\overline{\|\epsilon\|}$	$\overline{\|\epsilon\|}$
Antofagasta	Analysis	0.986	1.39	2.54	2.50
12GMT	24-h Forecast	0.985	1.41	3.00	2.91
Temperature (C)	48-h Forecast	0.984	1.32	3.24	3.11
Quintero	Analysis	0.991	1.15	2.48	2.52
12GMT	24-h Forecast	0.989	1.35	3.39	3.25
Temperature (C)	48-h Forecast	0.988	1.37	3.96	3.71
Antofagasta	Analysis	0.918	-0.36	2.11	
12GMT	24-h Forecast	0.813	-3.36	4.31	
Wind (m/s)	48-h Forecast	0.781	-0.24	5.09	
Quintero	Analysis	0.854	1.04	4.26	
12GMT	24-h Forecast	0.785	0.95	6.13	
Wind (m/s)	48-h Forecast	0.748	1.05	7.24	

TABLE VIII. – *Extreme wind events (velocities bigger than 10 m/s): averages of daily relative errors of radiosounding wind velocity against ECMWF analysis and 24/48-hour forecast during 1993 period and during the 1989-93 period.*

	Antofagasta		Quintero	
	1993	89-93	1993	89-93
Analysis	20%	21%	52%	31%
24h For.	48%	46%	98%	62%
48h For.	63%	58%	118%	80%

TABLE IX. – *Correlations (r), mean errors ($\overline{\epsilon}$), and mean absolute errors ($\overline{|\epsilon|}$) between 2-meter temperature observations at La Silla and Paranal during the 1989-93 period and values interpolated from 24/48-hour ECMWF forecast of upper-air temperature.*

Station	model	$\overline{\epsilon}$	$\overline{\|\epsilon\|}$	r
Paranal	24-h ECMWF forecast	-0.1	1.3	0.83
at 12 GMT	48-h ECMWF forecast	-0.2	1.4	0.80
La Silla	24-h ECMWF forecast	-1.2	1.7	0.90
at 12 GMT	48-h ECMWF forecast	-1.2	1.7	0.88

TABLE X. – *Absolute mean errors ($\overline{|\epsilon_\circ|}$) and correlations ($r_\circ$) between observed 2-meter temperatures (T2m) or 10-meter wind (W10m) and estimations by Kalman filter post-processing ($\circ = Kal$), persistence ($\circ = per$), ECMWF 24-hour forecast ($\circ = for$), and ECMWF analysis ($\circ = ana$) at 12 GMT. N is number of data used for comparison during the period from 1989 to 1993. The small differences with respect to Tables V, VI and IX are due to the different number of values used.*

field	site	N	$\overline{\lvert\epsilon_{Kal}\rvert}$	r_{Kal}	$\overline{\lvert\epsilon_{per}\rvert}$	r_{per}	r_{for}	r_{ana}
T2m	Paranal	572	1.1	0.88	1.5	0.80	0.79	0.87
(C)	La Silla	1114	1.6	0.86	2.1	0.77	0.82	0.86
W10m	Paranal	608	3.0	0.59	3.4	0.56	0.50	0.64
(m/s)	La Silla	1242	2.5	0.64	2.9	0.59	0.57	0.43

TABLE XI. – *Contingency table between errors of Kalman filtering post-processing procedure (KFPP) and T_{2m} daily variation at Paranal during the period 89-93 at 12GMT. Sum of the counts along rows (shown in the last column of the table) is the number of times T_{2m} has been forecast within an error indicated in the first column of the table itself. Sum of the numbers along a column (shown in the last row), is the number of times T_{2m} has had a diurnal variation within the value shown in the first row. The generic count within a cell of the table gives number of times T_{2m} has been forecast within an error indicated in the corresponding cell of the first columns when the daily variation of T_{2m} was that indicated in the corresponding cell of the last row.*

		Persistence errors (K)													
		-6	-5	-4	-3	-2	-1	1	2	3	4	5	6		
	-6	0	0	0	1	1	0	0	0	0	0	0	0	2	
	-5	0	0	0	0	0	0	0	0	0	0	0	0	0	
	-4	0	1	1	4	2	1	0	0	0	0	0	1	10	
	-3	0	1	3	8	9	3	3	2	1	1	0	0	31	
KFPP	-2	1	2	11	14	17	22	9	3	0	0	0	0	79	Global
errors	-1	1	5	3	14	38	38	34	15	9	2	0	0	159	scores
(K)	1	2	2	2	4	15	34	47	28	16	5	3	0	158	of KFPP
	2	0	0	2	1	5	9	21	15	18	5	5	2	83	
	3	0	0	0	0	4	5	2	6	10	1	2	3	33	
	4	0	0	0	0	0	0	1	3	4	2	1	1	12	
	5	0	0	0	0	0	0	0	1	0	1	1	0	3	
	6	0	0	0	0	0	0	0	0	1	0	0	1	2	
		4	11	22	46	91	112	117	73	59	17	12	8		
		Global scores of persistence													

TABLE XII. – *Contingency counts between measured 10-meter wind speed and Kalman filter post-processing, for la Silla. Total number of data points is 1242 for the period from 1989 to 1993.*

		Kalman Filter estimates				
	V (m/s)	0-5	5-10	10-15	15-20	20-25
	0-5	392	137	17	3	0
	5-10	172	250	99	13	1
measured	10-15	4	50	59	21	2
	15-20	2	2	10	5	2
	20-25	0	0	0	0	1

TABLE XIII. – *Contingency counts between measured 10-meter wind speed and Kalman filter post-processing, for Paranal. Total number of data points is 608 for the period from 1989 to 1993.*

		Kalman Filter estimates				
	V (m/s)	0-5	5-10	10-15	15-20	20-25
	0-5	135	61	10	1	0
	5-10	103	118	61	14	1
measured	10-15	1.5	26	25	18	0
	15-20	0	1	10	6	3
	20-25	0	2	1	3	0

TABLE XIV. – *2-meter temperature at Paranal: LAM forecast errors ($\epsilon_{lam} = T_{lam} - T_{observed}$) and ECMWF analysis errors ($\epsilon_{ana} = T_{analysis} - T_{observed}$) for the 5 test cases discussed in the text.*

		Forecasting time			
		24	30	48	54
	ϵ_{lam}	2.66	2.10	.55	1.07
04/03/93	ϵ_{ana}	2.43	2.54	-.34	.75
	ϵ_{lam}	2.20	1.74	.12	.85
12/05/93	ϵ_{ana}	2.34	2.28	.23	2.07
	ϵ_{lam}	1.34	2.09	.03	****
20/05/93	ϵ_{ana}	.58	1.93	-1.47	****
	ϵ_{lam}	2.78	5.00	8.31	****
17/08/93	ϵ_{ana}	3.05	5.93	8.47	****
	ϵ_{lam}	.39	-.44	1.15	-.56
18/03/93	ϵ_{ana}	.36	.13	.32	-.67

TABLE XV. – *2-meter temperature at La Silla: LAM forecast errors ($\epsilon_{lam} = T_{lam} - T_{observed}$) and ECMWF analysis errors ($\epsilon_{ana} = T_{analysis} - T_{observed}$) for the 5 test cases discussed in the text.*

		Forecasting time			
		24	30	48	54
04/03/93	ϵ_{lam}	1.53	.49	2.03	2.33
	ϵ_{ana}	.36	.55	.74	1.30
12/05/93	ϵ_{lam}	2.28	1.61	-2.00	-2.59
	ϵ_{ana}	.49	-.31	-1.91	-.16
20/05/93	ϵ_{lam}	.32	-1.01	.63	3.39
	ϵ_{ana}	-.82	-2.73	-.08	1.73
17/08/93	ϵ_{lam}	****	****	****	****
	ϵ_{ana}	****	****	****	****
18/03/93	ϵ_{lam}	-.16	-3.80	.24	.52
	ϵ_{ana}	-1.30	-2.75	-.73	-.09

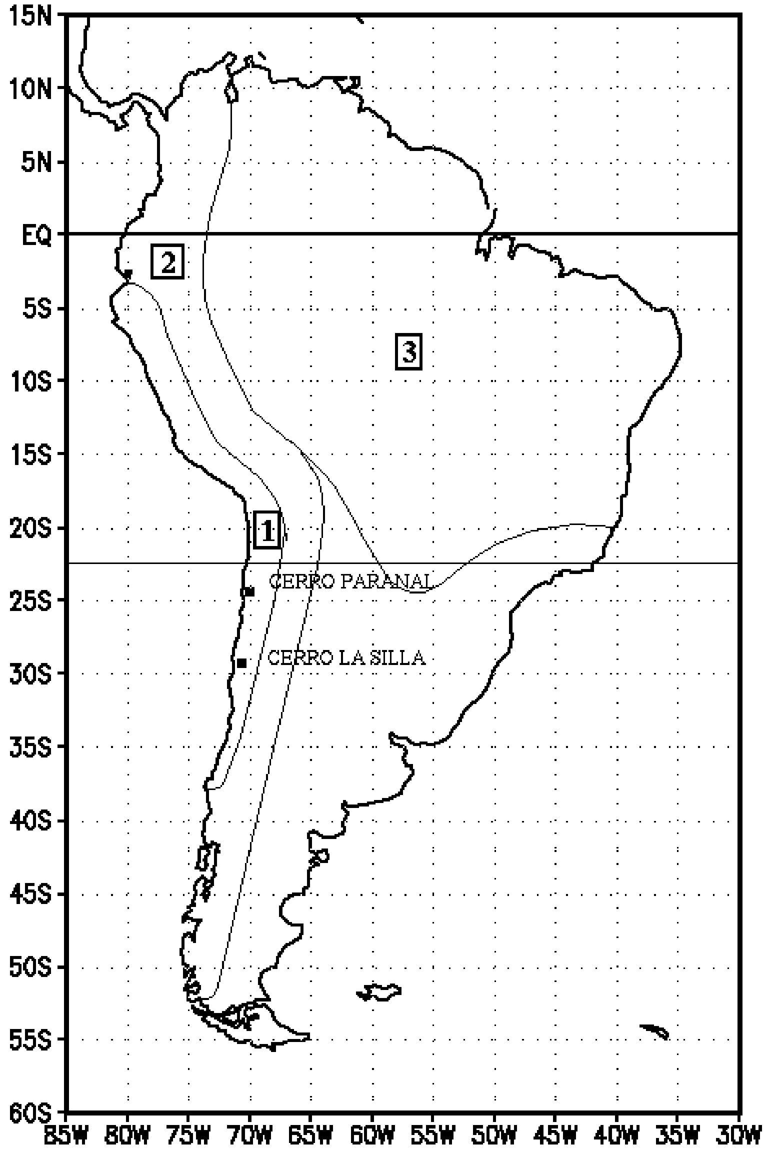

Fig. 1. – The three different climate areas in the southern tropical belt of South America. Region 1 is characterized by an arid climate, region 2 by a mountain climate and region 3 by a tropical climate. The position of the two ESO sites (Cerro Paranal and Cerro La Silla) is also shown.

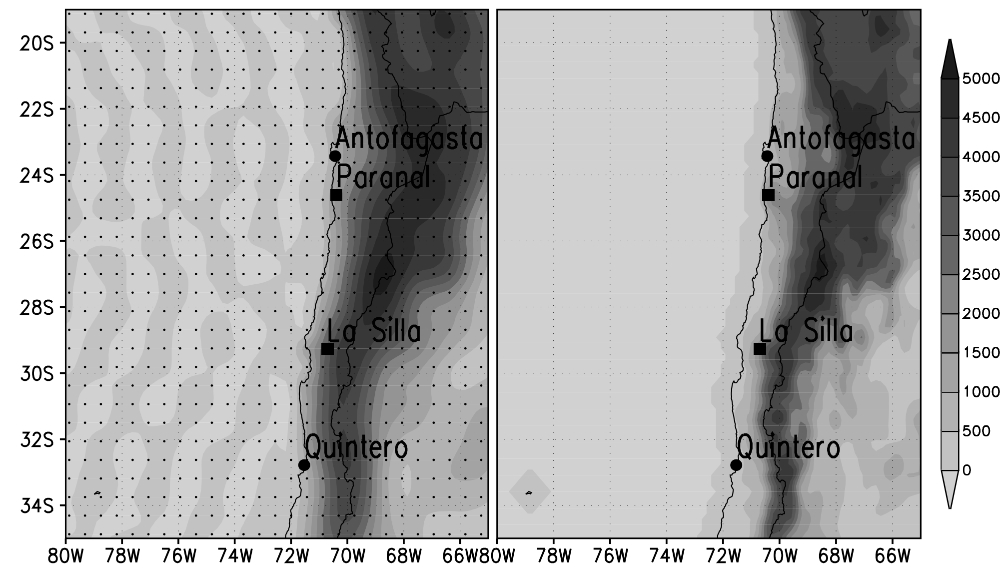

Fig. 2. – **left)** ECMWF orography (meters) at the horizontal resolution of 0.5625 degree and the corresponding grid-point mask; **right)** orography used for LAM simulations derived from NCAR surface geopotential height (meters) at one sixth degree of resolution. In both maps locations of Antofagasta, Paranal, La Silla, and Quintero are also shown.

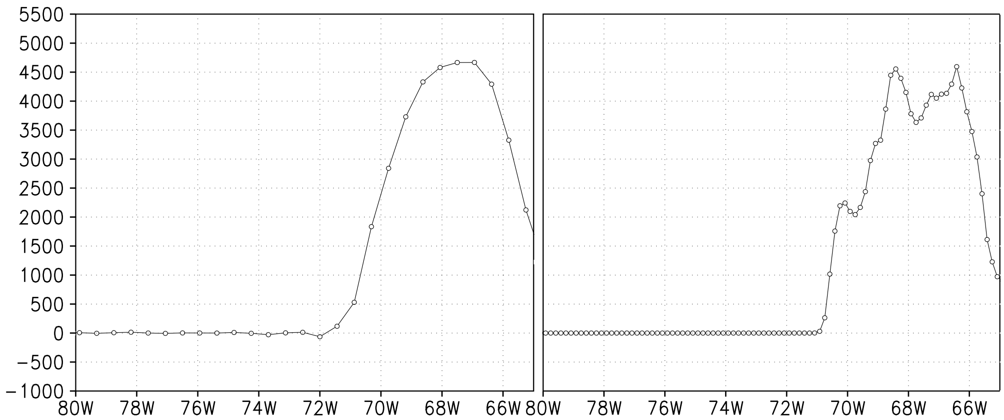

Fig. 3. – Comparisons between W-E orography sections (meters) at Paranal latitude: **left)** section of the orography of the ECMWF model, **right)** section of the orography used for LAM simulations.

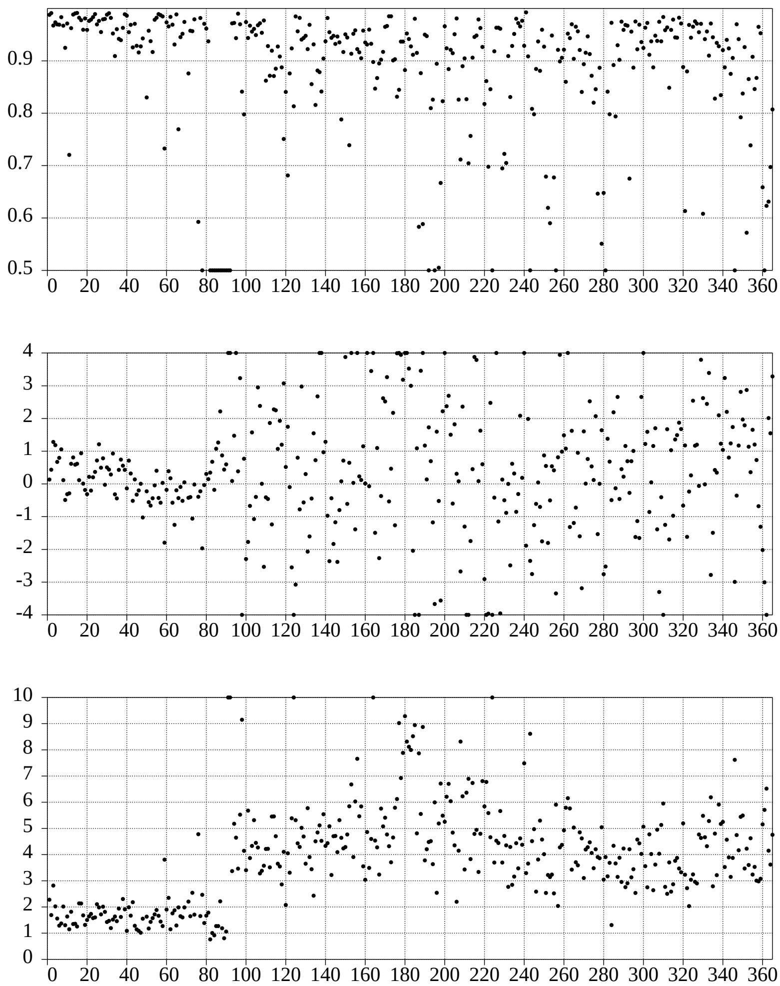

Fig. 4. – Daily (at 00 GMT) correlations (top), mean errors (middle), and mean absolute errors (bottom) between vertical profiles of wind velocity (m/s) from radiosoundings at Quintero in 1993 and the corresponding ECMWF analysis profiles interpolated at the site position.

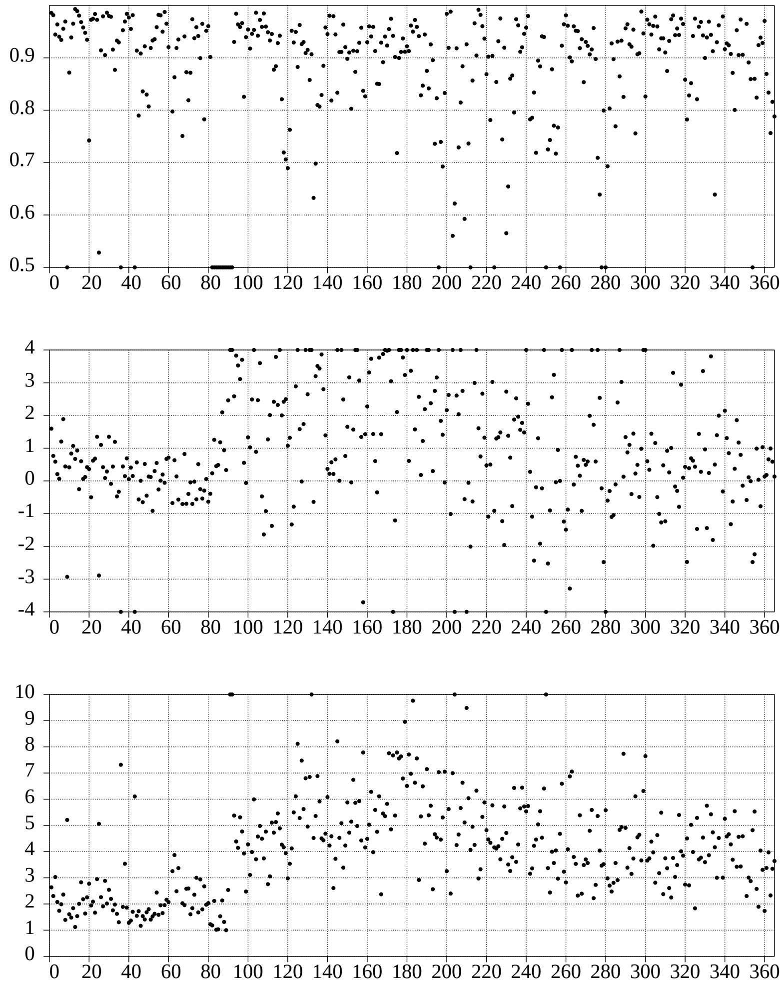

Fig. 5. – Daily (at 12 GMT) correlations (top), mean errors (middle), and mean absolute errors (bottom) between vertical profiles of wind velocity (m/s) from radiosoundings at Quintero in 1993 and the corresponding ECMWF analysis profiles interpolated at the site position.

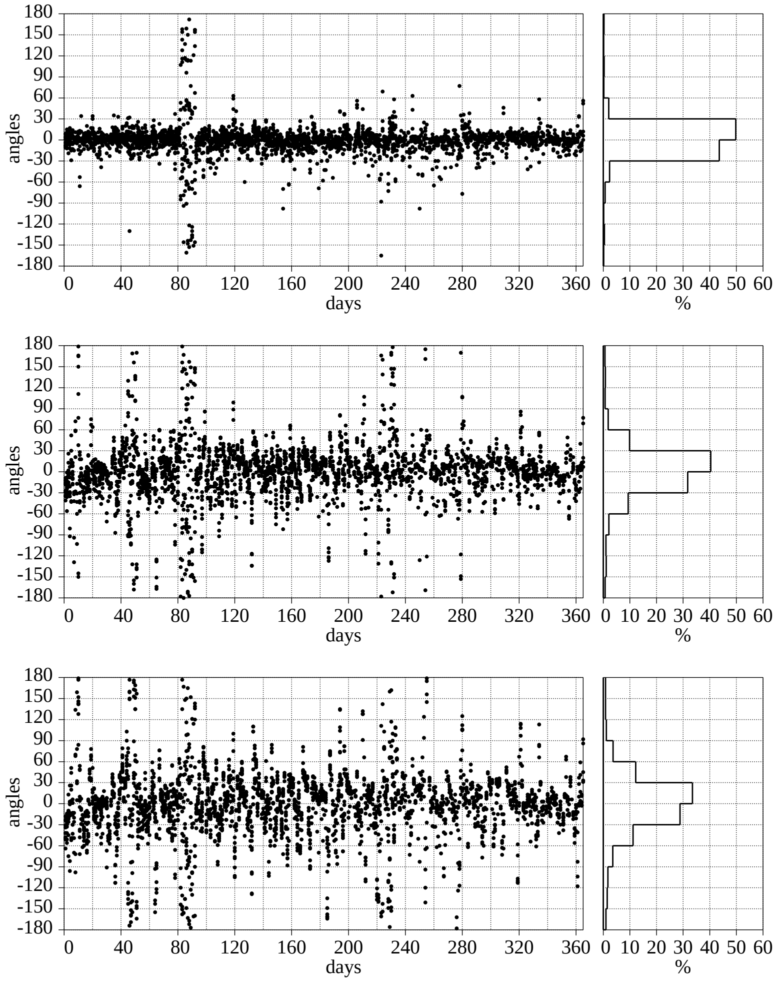

Fig. 6. – Angles between wind directions of vertical profiles from Quintero 12GMT radiosoundings and wind directions of ECMWF profiles from analysis (top), 24-hour forecast (middle), and 48-hour forecast (bottom): for each day in 1993 all angles measured from 100 to 350 mb are plotted.

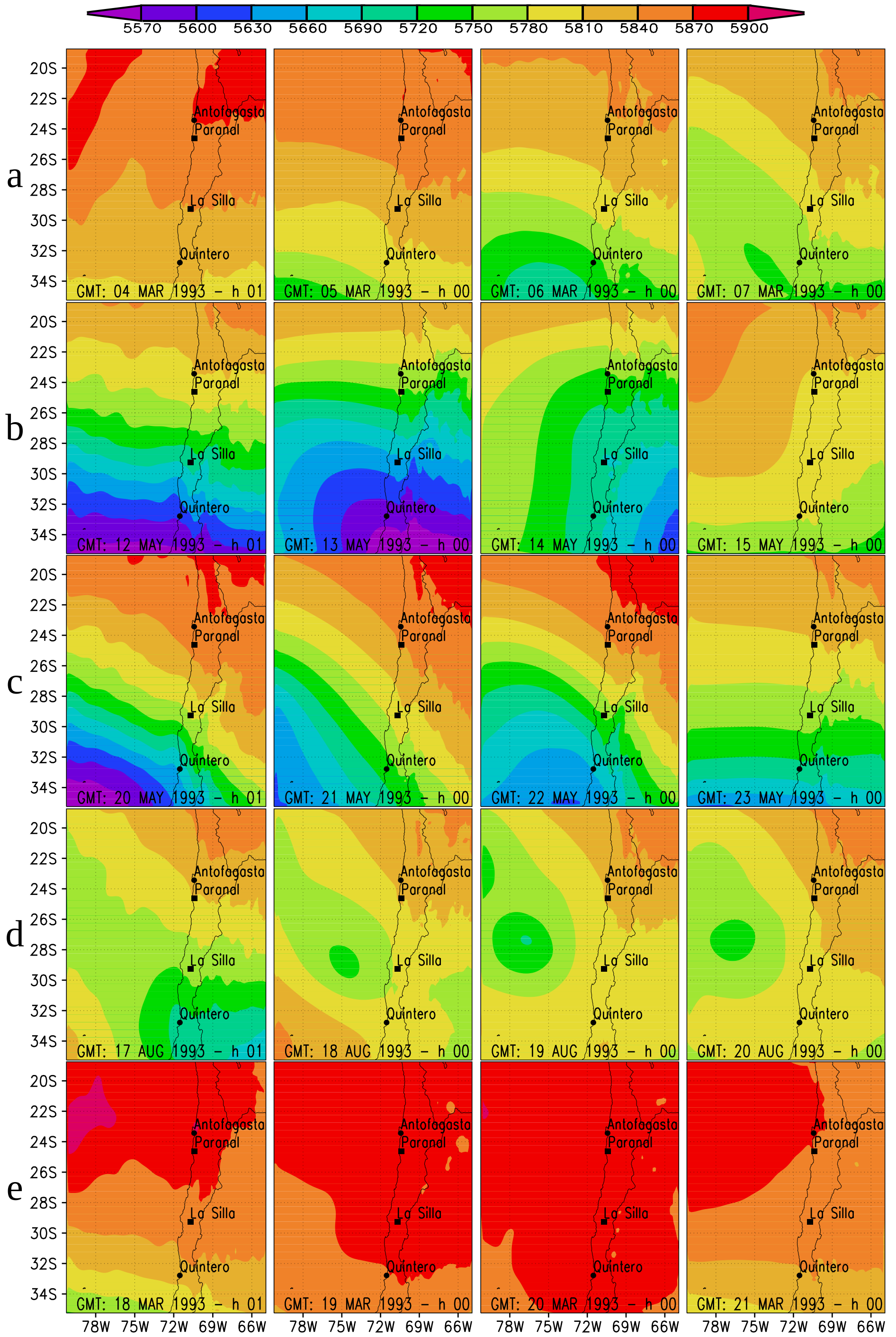

Fig. 7. – Forecasts of geopotential height (meters) at 500 mb after 1 hour, 24, 48, and 72 hours by LAM integration with a one sixth degree orography (see Figure 2, right) started at 00GMT on: a) 4 March 1993, b) 12 May 1993, c) 20 May 1993, d) 17 August 1993, e) 18 March 1993.

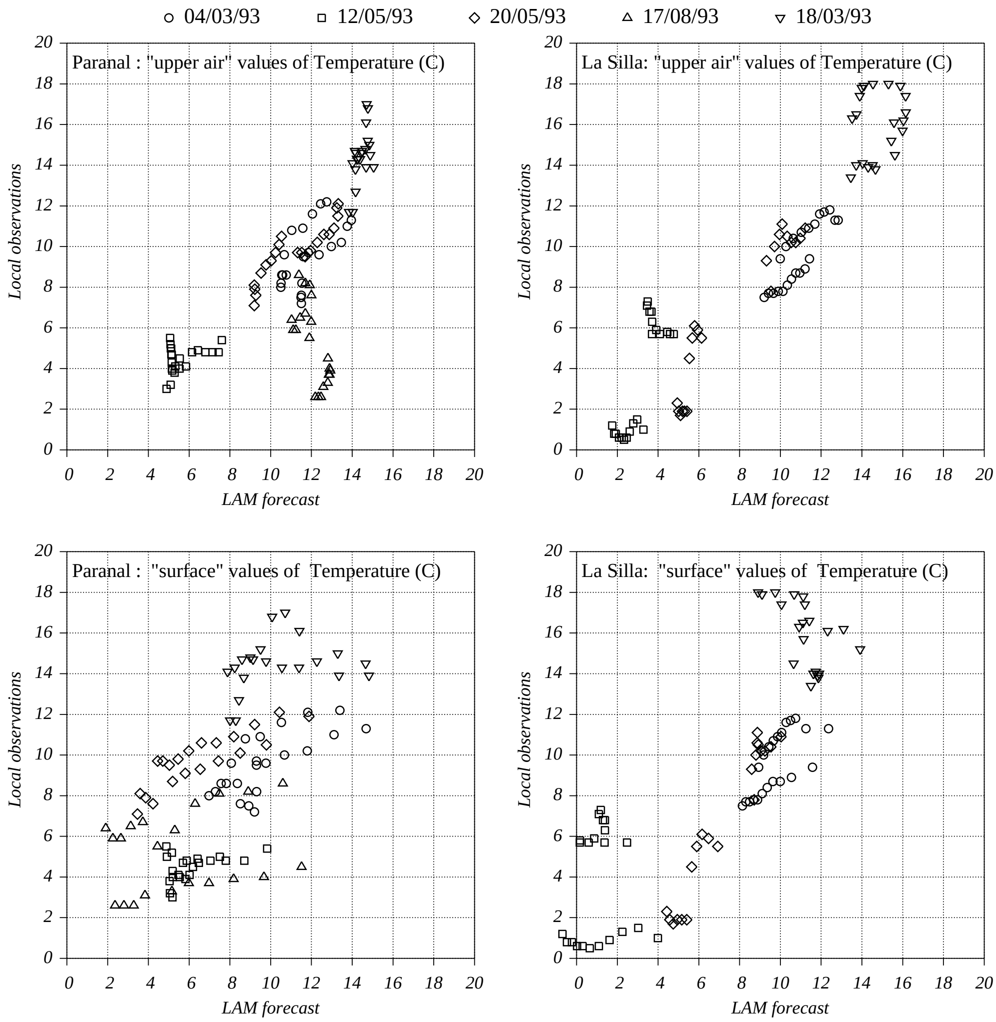

Fig. 8. – Scatter-plots of 2-meter air temperature observed at Paranal and La Silla versus estimates based on LAM "upper air" or "surface" values for the five test cases.

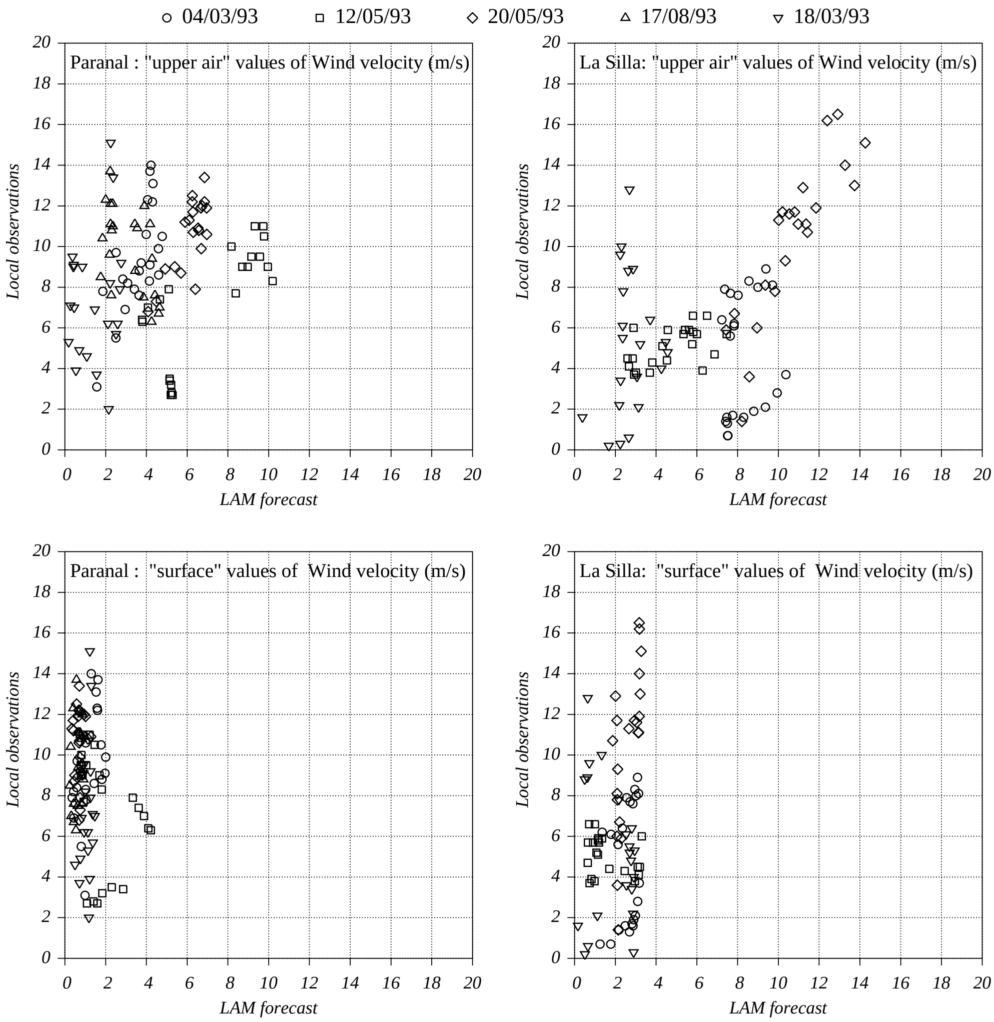

Fig. 9. – Scatter-plots of 10-meter wind velocity observed at Paranal and La Silla versus estimates based on LAM "upper air" or "surface" values for the five test cases.