Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

MNRAS **449,** 1664–1678 (2015) doi:10.1093/mnras/stv332

Forecast of surface layer meteorological parameters at Cerro Paranal with a mesoscale atmospherical model

F. Lascaux,★ E. Masciadri and L. Fini

INAF Osservatorio Astrofisico di Arcetri, Largo Enrico Fermi 5, I-501 25 Florence, Italy

Accepted 2015 February 13. Received 2015 January 13; in original form 2014 October 28

ABSTRACT

In this paper, we aim to prove the feasibility of the forecast of all the most relevant classical atmospherical parameters for astronomical applications (wind speed, wind direction, temperature) above the European Southern Observatory ground-based site of Cerro Paranal using a mesoscale atmospherical model called MESO-NH. In an earlier paper, we preliminarily treated the model performances obtained in reconstructing some key atmospherical parameters in the surface layer 0–30 m, studying the bias and the root-mean-square error (RMSE) on a statistical sample of 20 nights. The results were very encouraging and therefore it appeared mandatory to confirm such good results on a much richer statistical sample. In this paper, the study has been extended to a total sample of 129 nights between 2007 and 2011, distributed in different parts of the solar year. This large sample made our analysis more robust and definitive in terms of the model performances and permitted us to confirm the excellent performances of the model. Besides, we present an independent analysis of the model performances using the method of the contingency tables. Such a method permitted us to provide complementary key information with respect to the bias and the RMSE, which is particularly useful for an operational implementation of a forecast system.

Key words: turbulence – atmospheric effects – methods: data analysis – methods: numerical – site testing.

1 INTRODUCTION

This paper is part of a general study about the feasibility of the forecast of meteorological parameters and optical turbulence (OT) at the European Southern Observatory (ESO) sites of Cerro Paranal and Cerro Armazones in the framework of the Modelling ESO Sites (MOSE) project. The MOSE project is presented extensively in a previous paper (Masciadri, Lascaux & Fini 2013). We only recall here that the MOSE project aims at proving the feasibility of the forecast of all the most relevant classical atmospherical parameters for astronomical applications, such as wind speed intensity, wind direction, temperature and relative humidity (RH), and the OT (C_N^2 profiles) with the integrated astroclimatic parameters derived from C_N^2: that is, the seeing (ε), the isoplanatic angle (θ_0) and the wavefront coherence time (τ_0) above the two ESO sites of Cerro Paranal, which is the site of the Very Large Telescope (VLT), and Cerro Armazones, the site selected for the European Extremely Large Telescope (E-ELT).

The final outcome of the project is to investigate the opportunity to implement an automatic system for the forecast of these parameters at the VLT Observatory at Cerro Paranal and at the E-ELT Observatory at Cerro Armazones.

In a previous paper (Lascaux, Masciadri & Fini 2013), we presented results of a deep analysis of the bias and root-mean-square error (RMSE) between observations and model outputs of absolute temperature, wind speed and wind direction above both astronomical sites (Paranal and Armazones). These statistical operators provide fundamental information on systematic and statistical model errors. This statistical study was performed on a sample of 20 nights, the same for each site. Some of these nights come from the PAR2007 campaign (Dali Ali et al. 2010), and the sample was completed with nights having available observations from the same period of the year. In the same study, we also provided a detailed analysis on the model performances on individual nights. The most important results we obtained can be summarized as follows. The model shows a very good score of success in terms of bias and RMSE for absolute temperature – median bias in the [0.03,0.64] °C range and median RMSE in the [0.64,0.93] °C range – and wind direction – median bias in the [−1.01,−8.55] ° range and median RMSE in the [30,41]° range for a wind speed threshold of 2 m s^{-1} equivalent to $RMSE_{\rm rel}$ in the [17,23] per cent range – when it is used in grid-nesting configuration and the horizontal resolution of the innermost model domain is 500 m. To obtain equivalent satisfactory

★ E-mail: lascaux@arcetri.astro.it

results for the wind speed, we need to use a horizontal resolution of the innermost model domain equal to 100 m. Under this condition, we obtained a median bias within 0.93 m s^{-1} and a median RMSE within 2.18 m s^{-1}. These results were, in conclusion, very promising.

In this paper, the preliminary analysis carried out by Lascaux et al. (2013) was extended to a sample of 129 nights uniformly distributed between 2007 and 2011. The analysis is performed for Cerro Paranal. It was only possible to have access to homogeneous measurements distributed over such a long period of time at this site.[1] All 129 nights belong to different periods of the year, and not only to summer (as for the sample of 20 nights from Lascaux et al. 2013). This permits us to analyse definitely the robustness of the model on a rich statistical sample, independently from the period of the year considered. All the measurements, at Cerro Paranal, are part of the VLT Astronomical Site Monitor (Sandrock & Amestica 1999).

Moreover, for the same sample of nights, we present a complementary statistical analysis based on a different approach. Indeed, it is worth noting that the bias and RMSE, in spite of being fundamental statistical operators providing key information on model performances, do not provide the necessary information we would like to have in terms of model performances (Thornes & Stephenson 2001; Nurmi 2003; Jolliffe & Stephenson 2003). A key rule in the complex art of model-prediction estimation is that it makes no sense to quantify the model performance with one parameter only or one method only. To investigate the quality of a model prediction, a method widely used in the physics of the atmosphere, as well as in other fields such as economy and medicine, consists of constructing and analysing contingency tables. From these tables, it is possible to derive a number of different parameters that describe the quality of the model performances.

A contingency table allows for the analysis of the relationship between two or more categorical variables. One of the first classic examples of a forecast verification using contingency tables is the tornado forecast from Finley (1884). A contingency table is a table with $n \times n$ entries that displays the distribution of modelled outputs and observations in terms of frequencies or relative frequencies. Here, the variables we considered are the observations of different categories of temperature, wind speed, wind direction, near the surface, and the respective reconstructed parameters by the MESO-NH model.

In this part of Chile, the weather is particularly dry, the RH in the surface layer is typically below 30 per cent. The frequency of nights with RH above 80 per cent, the level at which the dome of the telescope has to be closed and observations are not allowed, is definitely very low (typically fewer than five nights each year). For all other nights, the RH does not really represent a critical issue in terms of scheduling of observations because the RH value is not high enough to affect the observations. The study of RH close to the ground is therefore not very critical or important for telescopes in this region. Also, it has been observed for the TMT site characterization in this region (T. Travouillon, private communication) that the reliability of RH measurements, when values are so low, can hardly be assured. This makes it difficult (and mostly useless) to investigate the model performances in reconstructing the RH on the same sample of nights selected for the analysis of other parameters. A more interesting analysis might be to check the model performances in reconstructing the RH values on those few nights each year where RH is higher than the threshold that makes it necessary to close the dome of the telescope. In Appendix B, we report satisfactory results obtained in this case.

In Section 2, we describe the model configuration defined for this study. In Section 3, we report the bias and RMSE obtained for the new sample of 129 nights, at Cerro Paranal. In Section 4, we define a contingency table for different discretized values n of observations and predictions, and we explain which parameters qualifying the model behaviour can be retrieved from them. In Section 5, we describe the criteria we used to identify the thresholds separating the n categories. In Sections 6, 7 and 8, we provide the results in terms of model performances (reconstruction of the absolute temperature, the wind speed and wind direction in the 0–30 m range). Finally, conclusions are drawn in Section 9.

2 BRIEF OVERVIEW OF THE NUMERICAL CONFIGURATION

2.1 Sample selection

The 129 nights have been selected starting with the 20 nights of Lascaux et al. (2013). Those 20 nights were all concentrated in 2007 November and December; the reader can refer to Lascaux et al. (2013) to understand how this first sample was constructed. When results obtained with the sample of 20 nights indicated a very promising model behaviour, in order to have a better evaluation of the model performances in statistical terms, it was decided to extend this sample to a more consistent sample. Two different periods, distant in time, were arbitrarily chosen (the only constraint was the availability of continuous measurements): (i) from 2007 January to 2007 December; (ii) from 2010 June to 2011 May. We tried to select nights that were regularly spaced in time, as much as possible, in order to cover all four seasons (which are characterized by slightly different meteorological behaviours). The result was 55 nights in 2007 (the 20 nights mentioned above, plus around three nights per month), and 74 nights in 2010–2011 (around six nights per month). No specific criterion has been used on the goodness of the weather conditions. In other words, nights include good and/or bad weather conditions, so no specific bias affects the sample. All these nights were characterized by different conditions of wind speed, temperature and RH (see the discussion in Section 3).

2.2 Model configuration

All the numerical simulations of the nights (129 in total) presented in this study have been performed with the mesoscale numerical weather model MESO-NH[2] (Lafore et al. 1998) and the ASTRO-MESO-NH code for the optical turbulence (Masciadri, Vernin & Bougeault 1999). The model has been developed by the Centre National des Recherches Météorologiques (CNRM) and Laboratoire d'Aérologie (LA) from Université Paul Sabatier (Toulouse). The MESO-NH model can simulate the temporal evolution of three-dimensional meteorological parameters over a selected finite area

[1] At Cerro Armazones during this period, some of the measurements were done with Thirty Meter Telescope (TMT) instrumentation and some with ESO instrumentation with sensors located at a different number of levels and heights above the ground. This lack of homogeneity in measurements did not permit us to perform the analysis at Cerro Armazones using the same procedure. Therefore, we have decided in this paper to focus our attention only on Cerro Paranal.

[2] http://mesonh.aero.obs-mip.fr/mesonh/

Table 1. MESO-NH model configurations. Columns 2–4 show the horizontal resolution ΔX, the number of grid points and the horizontal surface covered by the model domain, respectively.

Domain	ΔX (km)	Grid points	Domain size (km)
Domain 1	10	80 × 80	800 × 800
Domain 2	2.5	64 × 64	160 × 160
Domain 3	0.5	150 × 100	75 × 50
Domain 4	0.1	100 × 100	10 × 10

of the globe. We refer the reader to Masciadri et al. (2013), Section 3.3, for the general model configuration and the physical packages used for this study. We just recall here that we used the grid-nesting technique (Stein et al. 2000), which consists of using different imbricated domains of the digital elevation models (DEMs, i.e. orography) extended on smaller and smaller surfaces, with increasing horizontal resolution but with the same vertical grid. In this study, we use two different configurations. The first grid-nesting configuration employs three domains (Table 1) and the innermost resolution is $\Delta X = 500$ m. The second configuration is made of four imbricated domains, the first three the same as the previous configuration, plus one centred at the Cerro Paranal site, with a horizontal resolution of $\Delta X = 100$ m (all domains of Table 1). Along the z-axis, we have 62 levels distributed as follows: a first vertical grid point equal to 5 m, a logarithmic stretching of 20 per cent up to 3.5 km above the ground, and an almost constant vertical grid size of ∼600 m up to 23.8 km.

All the simulations performed for the analyses discussed in this paper were initialized the day before at 18 UT and forced every 6 h with analyses from the European Centre for Medium-Range Weather Forecasts (ECMWF), and finished at 09 UT (05 LT) of the simulated day (for a total duration of 15 h). The reader can refer to Lascaux et al. (2013) for more information about the site coordinates and ground altitude.

Before computing bias and RMSE, and before constructing the contingency tables, all the investigated meteorological parameters (measurements and MESO-NH outputs) have been first averaged with a moving average of 1 h to cut the high frequencies, and then resampled over a 20-min interval. The moving average of 1 h has been selected because astronomers are, in reality, more interested in the trend of the prediction. It has been observed that the moving average is more efficient to identify the model and measurement trends than a simple resampling. The high-frequency variability on a time-scale of 5 or 10 min is less relevant and useless to predict because the scheduling of scientific programs could not be tuned with such high frequency. Astronomers are interested in identifying whether the trend of a parameter is increasing, decreasing or is stationary, in order to be able to take a decision about changing a modality of observation or a scientific program. The selection of the interval of 20 min for the resampling of measurements is justified by the fact that this is more or less the effective time necessary to switch from one modality of observation to another. The values of 1 h and 20 min have been selected in agreement with ESO staff. In the analysis presented in this paper, we consider the model's abilities to reconstruct the atmospherical parameters during the night-time only.[3]

[3] To perform daytime predictions, a different model configuration should be selected.

3 STATISTICAL ANALYSIS

In this section, we estimate the statistical model reliability using three statistical operators: the bias, the RMSE and the bias-corrected RMSE:

$$BIAS = \sum_{i=1}^{N} \frac{(\Delta_i)}{N} \tag{1}$$

and

$$RMSE = \sqrt{\sum_{i=1}^{N} \frac{(\Delta_i)^2}{N}}. \tag{2}$$

Here, $\Delta_i = Y_i - X_i$, where X_i are the individual observations, Y_i are the individual simulation parameters calculated at the same time and N is the number of times for which a couple (X_i,Y_i) is available with both X_i and Y_i different from zero. Because the wind direction is a circular variable, we define Δ_i for the wind direction as

$$\Delta_i = \begin{cases} Y_i - X_i & \text{if } |Y_i - X_i| \leq 180^\circ \\ Y_i - X_i - 360^\circ & \text{if } Y_i - X_i > 180^\circ \\ Y_i - X_i + 360^\circ & \text{if } Y_i - X_i < -180^\circ. \end{cases} \tag{3}$$

From the bias and the RMSE, we deduce the bias-corrected RMSE (σ),

$$\sigma = \sqrt{RMSE^2 - BIAS^2}, \tag{4}$$

which represents the error of the model once the systematic error (the bias) is removed. For the wind direction, we add another operator, the relative RMSE, with RMSE expressed in degrees:

$$RMSE_{\rm rel} = \frac{RMSE}{180} \times 100 \,\text{per cent}. \tag{5}$$

This is justified because the wind direction is a circular variable and the maximum error is obtained when RMSE is equal to 180°.

Fig. 1 shows the scattered plots for the temperature, the wind speed (with both $\Delta X = 500$ m and $\Delta X = 100$ m configurations) and the wind direction (with and without filtering the lowest wind speed). The first thing to notice is the total extent of the observed values for all the meteorological parameters. The temperature goes from around 0 °C to 20 °C, the wind speed reaches 20 m s^{-1} and the wind direction covers the whole [0–360]° range. This is in agreement with the minimum and maximum observed values during a whole year at the site, as reported on the Astroclimatology of Paranal web site.[4] This assures us that all possible different meteorological conditions were encountered for the 129 simulated nights.

The forecasts of the absolute temperature and the wind direction near the surface at Cerro Paranal, with the MESO-NH model, are excellent. Considering the whole sample (Fig. 1 and Table 2), the bias in the temperature is very small (no larger than 0.37 °C in absolute value) and the RMSE is no larger than 1 °C.

The bias in the wind direction is almost zero, and the RMSE is around 37° (the data for which the wind velocity was inferior to 3 m s^{-1} have been excluded from the computations, because these data are associated, in general, with a great dispersion for the wind direction but are not interesting data for astronomers). This corresponds to $RMSE_{\rm rel}$ of around 20 per cent.

For the wind speed, we observe that the conclusions found by Lascaux et al. (2013) are confirmed. The high horizontal resolution ($\Delta X = 100$ m) is mandatory to reduce the bias to acceptable values

[4] http://www.eso.org/gen-fac/pubs/astclim/paranal

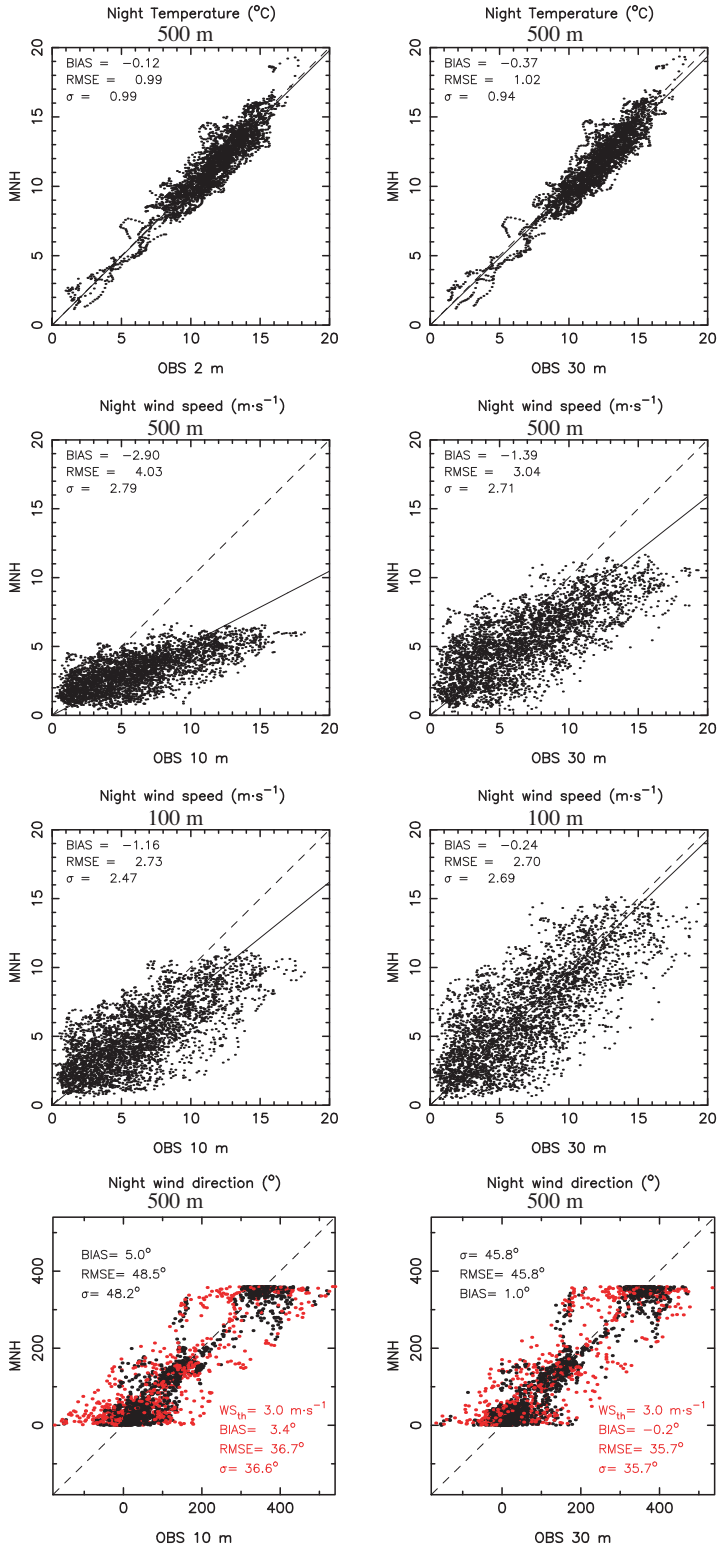


Figure 1. Scattered plots of the absolute temperature (top row), of the wind speed with the $\Delta X = 500$ m configuration (second row), of the wind speed with the $\Delta X = 100$ m configuration (third row) and of the wind direction (bottom row). A moving average (1 h) has been applied to both measurements and MESO-NH outputs (see the text). Each point represents an average of 20 min. The thin black line is the regression line passing by the origin. The red dots correspond to points for which the observed wind was inferior to $WS_{\rm th} = 3$ m s^{-1}. The corresponding values of bias and RMSE of the wind direction, for the sample without these dots, are reported in red at the bottom-right of the wind direction scattered plots. The values of the bias and the RMSE are reported in Table 2.

Table 2. Near surface bias, RMSE, bias-corrected RMSE σ (MESO-NH minus observations), at Cerro Paranal, for the sample of 129 nights (see Fig. 1). The relative RMSE is also reported in the case of the wind direction. The computations have been made considering the whole sample of data points. For the wind direction, data corresponding to a wind speed inferior to 3 m s^{-1} have been discarded from the computations.

	Temperature (°C) $\Delta X = 500$ m		Wind speed (m s^{-1}) $\Delta X = 500$ m		Wind speed (m s^{-1}) $\Delta X = 100$ m		Wind direction (°) $\Delta X = 500$ m	
Cerro Paranal	2 m	30 m	10 m	30 m	10 m	30 m	10 m	30 m
BIAS	− 0.12	− 0.37	− 2.90	− 1.39	− 1.16	− 0.24	3.4	− 0.2
RMSE	0.99	1.02	4.03	3.04	2.73	2.70	36.7	35.7
σ	0.98	0.95	2.80	2.70	2.47	2.69	36.5	35.7
$RMSE_{rel}$							20.4 per cent	19.8 per cent

Table 3. Model performance for individual nights. Near surface median bias, RMSE and bias-corrected RMSE (MESO-NH minus observations). The subscripts and superscripts give the first and third quartiles, respectively. The values for the wind speed were obtained with the $\Delta X = 100$ m configuration for MESO-NH, whereas the values obtained for the temperature and wind direction were obtained with the $\Delta X = 500$ m configuration.

	Temperature		Wind speed		Wind direction	
	2 m	30 m	10 m	30 m	10 m	30 m
BIAS	$-0.18^{+0.40}_{-0.73}$	$-0.48^{+0.12}_{-0.92}$	$-0.85^{+0.35}_{-2.50}$	$-0.14^{+0.91}_{-1.86}$	$+3.90^{+17.84}_{-12.22}$	$-0.32^{+12.06}_{-15.15}$
RMSE	$0.91^{+1.19}_{+0.60}$	$0.92^{+1.25}_{+0.58}$	$2.06^{+3.09}_{+1.41}$	$2.30^{+3.09}_{+1.63}$	$29.89^{+43.94}_{+17.20}$	$27.29^{+42.75}_{+14.57}$
σ	$0.54^{+0.76}_{+0.38}$	$0.48^{+0.69}_{+0.34}$	$1.25^{+1.70}_{+0.96}$	$1.45^{+2.01}_{+1.09}$	$15.94^{+27.32}_{+9.44}$	$15.30^{+31.22}_{+9.06}$

Table 4. Winter: near surface median bias, RMSE and bias-corrected RMSE (MESO-NH minus observations). Only the 54 winter nights are considered. The subscripts and superscripts give the first and third quartiles, respectively. The values for the wind speed were obtained with the $\Delta X = 100$ m configuration for MESO-NH, whereas the values obtained for the temperature and wind direction were obtained with the $\Delta X = 500$ m configuration.

	Temperature		Wind speed		Wind direction	
	2 m	30 m	10 m	30 m	10 m	30 m
BIAS	$-0.33^{+0.25}_{-0.81}$	$-0.53^{+0.03}_{-1.04}$	$-1.37^{+0.34}_{-2.69}$	$+0.00^{+1.02}_{-1.33}$	$+9.36^{+18.35}_{-0.52}$	$+5.02^{+13.01}_{-5.08}$
RMSE	$0.79^{+1.16}_{+0.54}$	$0.80^{+1.26}_{+0.52}$	$2.21^{+3.28}_{+1.34}$	$2.26^{+3.45}_{+1.44}$	$18.65^{+39.83}_{+14.61}$	$16.22^{+33.09}_{+11.78}$
σ	$0.41^{+0.60}_{+0.32}$	$0.39^{+0.53}_{+0.30}$	$1.08^{+1.67}_{+0.82}$	$1.32^{+1.95}_{+1.03}$	$11.12^{+16.32}_{+7.97}$	$10.11^{+17.41}_{+7.46}$

(from 2.90 to 1.16 m s^{-1} at 10 m and from 1.39 to 0.24 m s^{-1} at 30 m, in absolute values). With the $\Delta X = 100$ m resolution, the RMSE is around 2.7 m s^{-1}.

When looking at the statistics for single nights, the performances are even better. We computed the bias, RMSE and bias-corrected RMSE for every night of the sample, and constructed the cumulative distributions at every level for every meteorological parameter. From the cumulative distributions, we can extract the median and the first and third quartiles. Figs A1 and A2 in Appendix A show the cumulative distributions at all levels. Table 3 summarizes all the median values, together with the first and third quartiles, for the single-night bias, RMSE and σ, for the absolute temperature, the wind speed with the $\Delta X = 100$ m configuration and the wind direction. The median RMSE for the absolute temperature is less than 0.92 °C. It is inferior to 2.30 m s^{-1} for the wind speed, and inferior to 29° for the wind direction.

To study the seasonal variation and to check whether some systematic better/worst model performance is observed in summer and winter, we also calculated bias, RMSE and bias-corrected RMSE for the two periods of [April–September], which we call winter, and [October–March], which we call summer (Tables 4 and 5). The two periods consist of 54 and 75 nights, respectively. The summer period has more nights because of the presence of the first 20 nights analysed by Lascaux et al. (2013), which were all concentrated in the months of 2007 November and December. We observe that, for the temperature, in summer we have a slightly better median bias but a slightly larger median RMSE as well as σ (bias-corrected RMSE) than in winter. In any case, the median RMSE remains below 1 °C. For the wind direction, the median bias of the absolute temperature is sightly better in summer than in winter but the median RMSE is larger (almost double, i.e. 33–35° instead of 16–19°) in summer than in winter. Also, σ shows the same trend, even if the difference between summer and winter is less marked. For the wind speed, negligible differences are observed for the median bias and RMSE.

Now that we have characterized the performances of the model using standard statistical parameters (bias, RMSE and σ), we want to know how good or bad the model is at predicting some specific events (such as strong wind speed) or categories (such as the quadrants from which the wind is blowing). For this, we have constructed contingency tables.

Table 5. Summer: near surface median bias, RMSE and bias-corrected RMSE (MESO-NH minus observations). Only the 75 summer nights are considered. The subscripts and superscripts give the first and third quartiles, respectively. The values for the wind speed were obtained with the $\Delta X = 100$ m configuration for MESO-NH, whereas the values obtained for the temperature and wind direction were obtained with the $\Delta X = 500$ m configuration.

	Temperature		Wind speed		Wind direction	
	2 m	30 m	10 m	30 m	10 m	30 m
BIAS	$-0.10^{+0.53}_{-0.54}$	$-0.34^{+0.16}_{-0.88}$	$-0.81^{+0.39}_{-1.96}$	$-0.31^{+0.71}_{-2.06}$	$-2.51^{+14.58}_{-17.37}$	$-6.74^{+10.51}_{-17.43}$
RMSE	$0.94^{+1.20}_{+0.66}$	$0.92^{+1.25}_{+0.60}$	$1.91^{+2.88}_{+1.47}$	$2.42^{+2.99}_{+1.78}$	$35.15^{+48.89}_{+20.63}$	$32.65^{+51.23}_{+18.56}$
σ	$0.62^{+0.85}_{+0.46}$	$0.59^{+0.73}_{+0.43}$	$1.31^{+1.71}_{+1.03}$	$1.54^{+2.05}_{+1.18}$	$19.07^{+37.22}_{+13.72}$	$20.66^{+40.00}_{+13.66}$

Table 6. Generic 2 × 2 contingency table.

Event		**Observations**		
		YES	NO	Total
Model				
	YES	a (hit)	b (false alarm)	$a+b$ Yes (Model)
	NO	c (miss)	d (correct rejection)	$c+d$ No (Model)
	Total	$a+c$ Yes (OBS)	$b+d$ No (OBS)	$N=a+b+c+d$ Total of events

4 WHAT IS A CONTINGENCY TABLE?

As mentioned in the Introduction, a contingency table allows for the analysis of the relationship between two or more categorical variables. Table 6 is an example of a generic 2 × 2 contingency table, where the observations are divided into two categories (YES and NO), which could, for example, correspond to a wind above a given threshold (YES) and a wind below this threshold (NO). In the 2 × 2 contingency table, four combinations are possible:

a, the number of times an event was well reproduced by the model (hit);

b, the number of times an event was erroneously forecasted by the model (false alarms);

c, the number of times an event was not reproduced by the model (miss);

d, the number of times the absence of an event was well reproduced by the model (correct rejection).

Using a, b, c and d (and $N = a + b + c + d$), we can compute different probabilities useful to have an insight on how well (or bad) the model performed for a particular event.

In the following, we list the different simple scores that we use in the rest of the paper (to define them we use a, b, c, d and N from the generic 2 × 2 contingency table of Table 6).

The per cent of correct detections *PC* (in per cent) is defined as

$$PC = \frac{a+d}{N} \times 100; \quad 0 \leq PC \leq 100 \,\text{per cent}, \tag{6}$$

where $PC = 100$ per cent is the best score, and corresponds to a perfect forecast. *PC* alone is not sufficient because sometimes it can be influenced by the most common category and it does not provide us with information on the ability of the model to detect observations in the different categories. Therefore, it is always preferable to provide also the probability of detection (*POD*) together with *PC*.

The probability of detection (*POD*, in per cent) of a given event, or hit rate, is defined as

$$POD(event_1) = \frac{a}{a+c} \times 100; \quad 0 \leq POD \leq 100 \,\text{per cent}, \tag{7}$$

where $POD = 100$ per cent is the best score. *POD* measures the proportion of observed events that have been correctly predicted by the model. In the 2 × 2 contingency table, only one event is generally considered, and thus only one *POD* is calculated. However, we could also consider the non-occurrence of the event as an event itself, and then define a second *POD* as

$$POD(event_2) = \frac{d}{b+d} \times 100; \quad 0 \leq POD \leq 100 \,\text{per cent}. \tag{8}$$

It might be of interest to highlight that, for a total random prediction and in the case of a 2 × 2 contingency table, $a = b = c = d$. This means that all *POD* are equal to 50 per cent, and $PC = 50$ per cent. The model is useful if it performs better than this random case.

A huge number of other statistical parameters can be deduced from a contingency table (such as false alarm ratio, probability of false detection, etc.). We limit the current study to the analysis of the parameters mentioned above, which already provide a consistent panorama of the model potentialities.

For our purpose, a 2 × 2 table is, however, too simplified an analysis. A 3 × 3 table is definitely more appropriate. It consists of dividing the observed values into three categories delimited by some thresholds. An example of a 3 × 3 contingency table is shown in Table 7. The equivalent *PC* and *POD* can be defined using a, b, c, d, e, f, g, h, i and N from Table 7, and in the case of the 3 × 3 contingency table, we also add a third parameter, which we call *EBD* (for extremely bad detection):

$$PC = \frac{a+e+i}{N} \times 100 \tag{9}$$

$$POD(event_1) = \frac{a}{a+d+g} \times 100 \tag{10}$$

$$POD(event_2) = \frac{e}{b+e+h} \times 100 \tag{11}$$

$$POD(event_3) = \frac{i}{c+f+i} \times 100 \tag{12}$$

$$EBD = \frac{c+g}{N} \times 100. \tag{13}$$

EBD represents the per cent of the most distant predictions, by the model, from the observations. For a perfect forecast, it is equal to 0 per cent. In the case of a perfectly random forecast ($a = b = \ldots = i = N/9$), all *POD* are equal to 33 per cent,

Table 7. Generic 3 × 3 contingency table.

Intervals	Observations 1	2	3	Total
Model				
1	a (hit 1)	b	c	$a+b+c$ 1 (Model)
2	d	e (hit 2)	f	$d+e+f$ 2 (Model)
3	g	h	i (hit 3)	$g+h+i$ 3 (Model)
Total	$a+d+g$ 1 (OBS)	$b+e+h$ 2 (OBS)	$c+f+i$ 3 (OBS)	$N=a+b+c+d+e+f+g+h+i$ Total of events

Table 8. Generic 4 × 4 contingency table.

Intervals	Observations 1	2	3	4	Total
Model					
1	a (hit 1)	b	c	d	$a+b+c+d$ 1 (Model)
2	e	f (hit 2)	g	h	$e+f+g+h$ 2 (Model)
3	i	j	k (hit 3)	l	$i+j+k+l$ 3 (Model)
4	m	n	o	p (hit 4)	$m+n+o+p$ 4 (Model)
Total	$a+e+i+m$ 1 (OBS)	$b+f+j+n$ 2 (OBS)	$c+g+k+o$ 3 (OBS)	$d+h+l+p$ 4 (OBS)	$N=a+b+c+d+e+f+g+h+i+j+k+l+m+n+o+p$ Total of events

$PC = 33$ per cent and $EBD = 22.2$ per cent. These values are a good reference to evaluate the performances of the model.

For the wind direction, a 3 × 3 table is not appropriate because this is a circular variable. A 4 × 4 (divided into four quadrants, with 90° each) is therefore a more suitable solution (see Table 8). *PC* and *POD* are defined in this case as

$$PC = 100 \times \frac{a+f+k+p}{N}, \tag{14}$$

$$POD(event_1) = 100 \times \frac{a}{a+e+i+m}, \tag{15}$$

$$POD(event_2) = 100 \times \frac{f}{b+f+j+n}, \tag{16}$$

$$POD(event_3) = 100 \times \frac{k}{c+g+k+o}, \tag{17}$$

$$POD(event_4) = 100 \times \frac{p}{d+h+l+p} \tag{18}$$

and

$$EBD = 100 \times \frac{c+h+i+n}{N}. \tag{19}$$

If we consider the random case ($a = b = \cdots = p = N/16$), $PC = 25$ per cent, all *POD* are equal to 25 per cent, and $EBD = 25$ per cent.

From this point and until the end of the paper, we write POD_i instead of $POD(event_i)$ with i being the event considered.

Table 9. Climatological tertiles for the wind speed at Cerro Paranal, at 33, 50 (median) and 66 per cent.

	Wind speed (in m s^{-1}) 33 per cent	Median (50 per cent)	66 per cent
10 m	3.98	5.47	7.25
30 m	4.40	6.10	8.10

5 CLIMATOLOGICAL TERTILES AT CERRO PARANAL

To construct the contingency tables that we discuss in the next sections, it is necessary to define some thresholds limiting the discretized intervals of the observed values and to divide the sample into three categories. For the 3 × 3 contingency tables, we need two thresholds. In this study, we decided to use the climatological tertiles computed using the available measurements. We used six complete years of measurements from a mast, from 2006 January 1 to 2012 December 31, which are part of the VLT Astronomical Site Monitor (Sandrock & Amestica 1999). At each level, and for each parameter, we computed the median value, and the tertiles, of all the available measurements. Table 9 shows the results for the wind speed at Cerro Paranal. Table 10 shows the results for the absolute temperature at Cerro Paranal.

6 ABSOLUTE TEMPERATURE

The first meteorological parameter reconstructed by the model that we investigate is the absolute temperature near the ground. We report the results only for the standard model ($\Delta X = 500$ m). We highlight

Table 10. Climatological tertiles for the absolute temperature at Cerro Paranal, at 33, 50 (median) and 66 per cent.

	Temperature (in °C)		
	33 per cent	Median (50 per cent)	66 per cent
2 m	11.03	12.24	13.27
30 m	11.40	12.58	13.64

the fact that results obtained with the high horizontal resolution model ($\Delta X = 100$ m; not shown here) are very similar. Tables 11 and 12 are the contingency tables for the absolute temperature at Cerro Paranal, at 2 m and 30 m, respectively. Each estimate corresponds to the average over a 20-min interval (as indicated in Section 3). For the sake of simplification, we used, for thresholds, the rounded values taken from the climatological study of Section 5. We can observe that the per cent of good forecast by the model (characterized by *PC*) is very good at every level. *PC* is between 73 and 75 per cent. Moreover, an *EBD* very close to 0 per cent at both levels is the sign that the model never produces extremely bad forecasts for the temperature. If we look at the probability of detection of the temperature in a given interval (characterized by *POD*), once again the model demonstrates its accuracy. *POD* is between 55.2 per cent (detection of temperature between 11.5 and 13.5 °C at 30 m) and 93.7 per cent (detection of a temperature inferior to 11.5 °C at 30 m). In all cases, *PC* and *POD* are larger than 33 per cent (the value of a random case). Therefore, this proves the utility of the model.

7 WIND SPEED

Considering the tendency of the model to underestimate the wind velocity slightly (see Section 3) – even if this underestimation is obviously quantitatively different between $\Delta X = 500$ and 100 m – it was decided to evaluate the predictive power of the model using a forecasted wind speed corrected by the multiplicative bias $B_{\rm mult}$ computed as

$$B_{\rm mult}(k) = \frac{\overline{WS_{\rm mod}}(k)}{\overline{WS_{\rm obs}}(k)}, \tag{20}$$

where $\overline{WS_{\rm mod}}(k)$ is the average of the forecasted wind speed over the whole sample of 129 nights, and $\overline{WS_{\rm obs}}(k)$ is the average of the observed wind speed from the same sample, at the level k (10 and 30 m for Cerro Paranal). The different values are summarized in Table 13. The choice of the multiplicative bias is justified by the fact that we want the highest values of the wind speed to be more corrected than the lowest values. It is important to note that the values of $B_{\rm mult}(k)$ can be considered as constant, whatever the subsample considered. Indeed, we have computed it with the initial sample of 20 nights, with another subsample of 73 nights and with the final sample of 129 nights, and in all cases its values were the one reported in Table 13 (variations of the second digit only are noticeable, and thus not reported here). This tells us that is useless to consider a calibration subsample and a complementary and independent validation or testing subsample. It is therefore more appropriate to apply this multiplicative factor on the whole sample of nights.

Table 13. Wind speed multiplicative biases (from equation 20). Values are rounded.

Cerro Paranal	10 m	30 m
$\Delta X = 500$ m	0.5	0.8
$\Delta X = 100$ m	0.8	0.9

7.1 Sensibility of the model detection to the wind speed threshold

Considering that the most critical issue for the wind speed is the model performances in detecting the strong wind speed, we

Table 11. A 3 × 3 contingency table for the absolute temperature during the night, at 2 m above ground level (a.g.l.) at Cerro Paranal, for the sample of 129 nights. We use the MESO-NH $\Delta X = 500$ m configuration. Total points = 3483; $PC = 75.5$ per cent; $EBD = 0.2$ per cent; $POD_1 = 91.8$ per cent; $POD_2 = 66.3$ per cent; $POD_3 = 65.9$ per cent.

Division by tertiles (climatology)	**Observations**		
	$T < 11°C$	$11°C < T < 13.5°C$	$T > 13.5°C$
Model			
$T < 11°C$	1157	315	0
$11°C < T < 13.5°C$	96	972	258
$T > 13.5°C$	8	178	499

Table 12. A 3 × 3 contingency table for the absolute temperature during the night, at 30 m a.g.l. at Cerro Paranal, for the sample of 129 nights. We use the MESO-NH $\Delta X = 500$ m configuration. Total points = 3483; $PC = 73.7$ per cent; $EBD = 0.5$ per cent; $POD_1 = 93.7$ per cent; $POD_2 = 55.2$ per cent; $POD_3 = 69.0$ per cent.

Division by tertiles (climatology)	**Observations**		
	$T < 11.5°C$	$11.5°C < T < 13.5°C$	$T > 13.5°C$
Model			
$T < 11.5°C$	1250	422	2
$11.5°C < T < 13.5°C$	68	663	292
$T > 13.5°C$	16	117	653

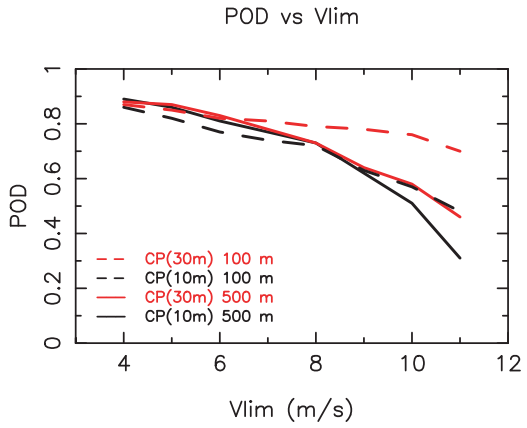


Figure 2. Evolution of the probability of detection (*POD*) from 2 × 2 contingency tables, by MESO-NH, of winds stronger than $V_{\rm lim}$, in function of $V_{\rm lim}$. Dashed lines, with $\Delta X = 100$ m configuration. Continuous lines, with $\Delta X = 500$ m configuration. In all cases, the MESO-NH wind speed was corrected by the multiplicative bias of Table 13. CP stands for Cerro Paranal.

Table 14. This contingency table considers the event: winds stronger than 6 m s^{-1}. The site is Cerro Paranal. The altitude is 10 m a.g.l. The wind speed is computed with the MESO-NH $\Delta X = 100$ m configuration and corrected by the multiplicative bias from Table 13. Total points = 3456; *PC* = 77.4 per cent; $POD_1(WS > 6\,{\rm m\,s^{-1}})$ = 77.2 per cent; $POD_2(WS < 6\,{\rm m\,s^{-1}})$ = 77.6 per cent.

		Observations	
		YES	NO
Model	YES	1230	418
	NO	363	1445

performed the following exercise before calculating the 3 × 3 contingency tables. To see how the model outputs depend on the choice of the threshold used for the wind speed, we plot in Fig. 2 the dependence of POD_1 from the 2 × 2 contingency tables, on the threshold $V_{\rm lim}$, for both $\Delta X = 500$ and 100 m configurations. $V_{\rm lim}$ varies between 4 and 11 m s^{-1}.

The two MESO-NH configurations ($\Delta X = 500$ and 100 m) were used with the correction of the wind by the multiplicative bias. The first thing we can see in Fig. 2 is that the performance of the simulations with the $\Delta X = 100$ m configuration depends less on the threshold than the performance of the $\Delta X = 500$ m configuration. Up to $V_{\rm lim} = 10$ m s^{-1}, with $\Delta X = 100$ m, all POD_1 are very good with the worst PODs at 10 m s^{-1} between 70 and 80 per cent. In contrast, the model performance decreases faster with $V_{\rm lim}$, for the $\Delta X = 500$ m configuration. At $V_{\rm lim} = 9$ m s^{-1}, the performance is still satisfactory (all POD_1 are around 60 per cent), but for $V_{\rm lim} > 9$ m s^{-1}, the performances are degraded (all POD_1 are inferior to 50 per cent, down to ∼30 per cent for $V_{\rm lim} = 11$ m s^{-1}). This tells us that the 100-m configuration guarantees a better model performance than the 500-m configuration. As an example, if we consider, at Cerro Paranal, a reasonable value of $V_{\rm lim} = 6$ m s^{-1} and we study the probability to detect wind speed stronger than this threshold, we obtain very satisfactory results (Table 14) for the PODs that are typically of the order of 77 per cent.

The interest for astronomers is mainly to be able to identify conditions of strong wind, and this result tells us that such a model in this configuration answers this necessity well.

7.2 Wind speed 3 × 3 contingency tables

In this section, we investigate the ability of the model to forecast the wind speed, as we did in Section 6. As we have seen in Section 7.1, we consider the best configuration for the model, that is, the high horizontal ($\Delta X = 100$ m) resolution configuration. The contingency tables have been divided into three intervals, for each level concerned, using the rounded climatological tertiles from Section 5 as thresholds. They are reported in Tables 15 and 16, at 10 and 30 m, respectively.

PC is significatively better than 33 per cent (random case): it is 60.0 per cent at 10 m and 60.6 per cent at 30 m. The *POD* also are very satisfactory, between 47.5 and 79.0 per cent. Moreover, the *EBD* are negligible (less than 4.4 per cent). If we consider POD_3

Table 15. A 3 × 3 contingency table for the wind speed during the night, at 10 m a.g.l. at Cerro Paranal. We use the MESO-NH $\Delta X = 100$ m configuration with the wind corrected by the multiplicative bias. Total points = 3456; *PC* = 60.0 per cent; *EBD* = 4.4 per cent; POD_1 = 56.2 per cent; POD_2 = 47.5 per cent; POD_3 = 74.3 per cent.

Division by tertiles (climatology)		Observations		
		$T < 4$ m s^{-1}	4 m s^{-1} $< T <$ 7 m s^{-1}	$T > 7$ m s^{-1}
Model				
	$T < 4$ m s^{-1}	641	283	50
	4 m s^{-1} $< T <$ 7 m s^{-1}	395	510	269
	$T > 7$ m s^{-1}	103	281	924

Table 16. A 3 × 3 contingency table for the wind speed during the night, at 30 m a.g.l. at Cerro Paranal. We use the MESO-NH $\Delta X = 100$ m configuration with the wind corrected by the multiplicative bias. Total points = 3456; *PC* = 60.6 per cent; *EBD* = 3.7 per cent; POD_1 = 52.5 per cent; POD_2 = 48.2 per cent; POD_3 = 79.0 per cent.

Division by tertiles (climatology)		Observations		
		$T < 4$ m s^{-1}	4 m s^{-1} $< T <$ 8 m s^{-1}	$T > 8$ m s^{-1}
Model				
	$T < 4$ m s^{-1}	554	289	27
	4 m s^{-1} $< T <$ 8 m s^{-1}	400	559	234
	$T > 8$ m s^{-1}	101	311	981

Table 17. A 4 × 4 contingency table for the wind direction α during the night, at 10 m, and 30 m a.g.l. at Cerro Paranal. We use the MESO-NH $\Delta X = 500$ m configuration. We filter out the observed wind inferior to 3 m s^{-1}. NE corresponds to $0^\circ < \alpha < 90^\circ$; SE corresponds to $90^\circ < \alpha < 180^\circ$; SW corresponds to $180^\circ < \alpha < 270^\circ$; NW corresponds to $270^\circ < \alpha < 360^\circ$. α is the wind direction (see Fig. 3). Total points = 6253; *PC* = 67.5 per cent; *EBD* = 2.8 per cent; *POD*(*NE*) = 81.7 per cent; *POD*(*SE*) = 65.5 per cent; *POD*(*SW*) = 27.9 per cent; *POD*(*NW*) = 41.8 per cent.

		Observations			
Cerro Paranal: all levels		NE	SE	SW	NW
Model					
	NE	2809	212	35	815
	SE	102	734	109	28
	SW	23	87	63	11
	NW	505	87	19	614

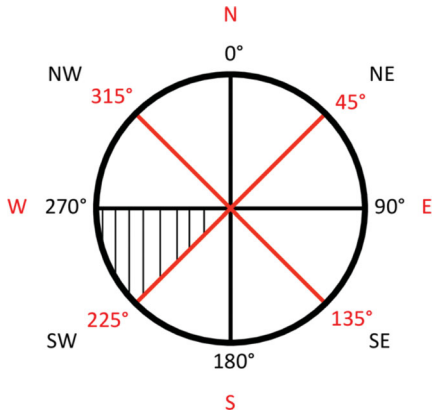


Figure 3. Convention chosen for the wind direction ($0^\circ < \alpha < 360^\circ$). The black quadrants correspond to the intervals of 90° chosen for the 4 × 4 contingency tables of Table 17; the red quadrants correspond to the intervals of 90° chosen for the 4 × 4 contingency tables of Table 18. The hatched section is discussed in the text.

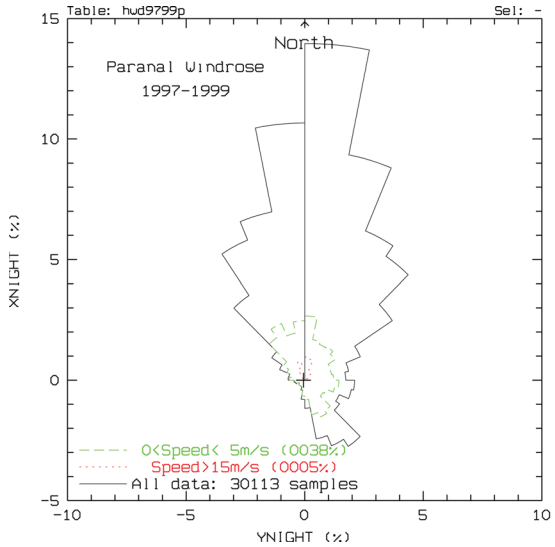
only (i.e. detection of wind superior to a fixed threshold, the second climatological tertile), which is the most interesting information for astronomers, then the results are even better (∼79.0 per cent).

A strong wind is indeed the main cause of vibrations of the primary mirror and adaptive secondary (a critical element of the adaptive optics system). All POD_3 for both sites and at all altitudes are excellent, between 74.3 and 79.0 per cent. This demonstrates the ability of the model to predict, at Cerro Paranal, critical wind speed values for adaptive optics applications.

8 WIND DIRECTION

In this section, we investigate the reconstruction of the wind direction by the MESO-NH model with the $\Delta X = 500$ m configuration. We have constructed 4 × 4 contingency tables divided into four quadrants of 90° each.

In Table 17, the samples are divided into the following four quadrants: (i) wind blowing from the direction between 0° and 90° (NE winds); (ii) wind blowing from the direction between 90° and 180° (SE winds); (iii) wind blowing from the direction between 180° and 270° (SW winds); (iv) wind blowing from the direction between 270° and 360° (NW winds). These quadrants are reported in black in Fig. 3. *PC* is equal to 67.7 per cent. The probability of detection of the wind between 0° and 90° (NE winds) is around 82 per cent; between 90° and 180° (SE winds), it is 65.5 per cent; between 180° and 270° (SW winds) it is 27.9 per cent; between 270° and 360° (NW winds), it is around 42 per cent. The model performance is very good in the quadrant in which the wind speed flows more frequently (between 0° and 90° with 82 per cent). The high frequency of occurrence of the NE winds at Cerro Paranal is confirmed by the windrose of Fig. 4 (left), between 1997 and 1999. The worst *POD* is found for the third quadrant (27.9 per cent), but



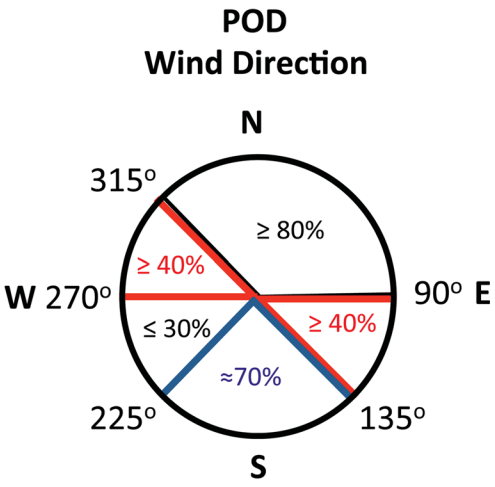


Figure 4. Left panel: observed wind direction at 10 m a.g.l. at Cerro Paranal. The sectors are in steps of 11°.25, and the length of the sector represents the fraction of time when the wind comes from this direction. The figure is extracted from the ESO astroclimatic web site at https://www.eso.org/gen-fac/pubs/astclim/paranal/wind/. Right panel: the *POD* of the wind direction by the model is summarized by sectors. The circle has been constructed by combining the results from Tables 17 and 18.

Table 18. A 4 × 4 contingency table for the wind direction α during the night, at 10 m, and 30 m a.g.l. at Cerro Paranal. We use the MESO-NH $\Delta X = 500$ m configuration. We filter out the observed wind inferior to 3 m s^{-1}. N corresponds to $-45^\circ < \alpha < 45^\circ$; E corresponds to $45^\circ < \alpha < 135^\circ$; S corresponds to $135^\circ < \alpha < 225^\circ$; W corresponds to $225^\circ < \alpha < 315^\circ$. α is the wind direction (see Fig. 3). Total points = 6253; *PC* = 77.7 per cent; *EBD* = 1.9 per cent; *POD*(*N*) = 92.8 per cent; *POD*(*E*) = 40.1 per cent; *POD*(*S*) = 69.5 per cent; *POD*(*W*) = 27.7 per cent.

		Observations			
Cerro Paranal: all levels		N	E	S	W
Model					
	N	3807	536	68	138
	E	196	473	119	4
	S	18	141	512	25
	W	83	30	38	64

this corresponds to only less than 4 per cent of the observed wind total sample at Cerro Paranal. This low frequency of occurrence is also confirmed by the windrose of Fig. 4 (left). In the other quadrants, the model performance is, in most cases, well above the random case (25 per cent). We note that the low frequency in the quadrant [180°, 270°] is confirmed by the climatology, and is not limited to our sample of 129 nights. If we divide now the sample into four different intervals (N, E, S and W, i.e. rotated by 45° with respect to the previous one in order to better constrain the angular sectors), we observe that the results remains very similar. In Table 18, this time the samples are divided into the following four quadrants: (i) wind blowing from the direction between −45° and 45° (N winds); (ii) wind blowing from the direction between 45° and 135° (E winds); (iii) wind blowing from the direction between 135° and 225° (S winds); (iv) wind blowing from the direction between 225° and 315° (W winds). These quadrants are reported in red in Fig. 3. We notice that *PC* remain high, as expected, with a value of 77.7 per cent. *POD* are very good – as high as 92.8 per cent for *POD*(*N*) – except for *POD*(*W*) (less than 28 per cent). In all cases, *EBD* is very low, always inferior to 2.8 per cent, which confirms the goodness of the prediction by the model.

POD(*SW*) (Table 17) and *POD*(*W*) (Table 18) are the only *POD*s giving values closer than the random case, so we can deduce that the wind blowing between 225° and 270° (hatched section of Fig. 3) is reconstructed with the most apparent difficulty. However, as previously mentioned, the observed frequency of these events is very low.

The results for the probability of detection of the wind direction by MESO-NH are summarized in Fig. 4 (right).

9 CONCLUSIONS

In this study, we have quantified the quality of the predictions of the absolute temperature, the wind speed and the wind direction in the surface layer 0–30 m at the ESO site of Cerro Paranal, using the atmospherical non-hydrostatical mesoscale model MESO-NH. First, we identified the bias and RMSE between model and observations. With respect to the study of Lascaux et al. (2013), the sample is more homogeneous (nights distributed in different periods of the year) and richer (129 nights instead of 20). This permitted us to obtain very reliable and robust values for the bias, the RMSE and the bias-corrected RMSE σ.

To quantify the quality of the model predictions, simulations of 129 nights (distributed between 2007 and 2011) have been performed. In addition to this, we also have constructed contingency tables and analysed different parameters deduced from these tables, more precisely the *PC* (per cent of correct detection), the *POD* (probability of detection of a single event) and the *EBD* (per cent of extremely bad detections). These contingency tables (and the associated statistical parameters) allowed us to quantify the performance of the MESO-NH model above Cerro Paranal.

For the absolute temperature, the per cent of correct detection (*PC*) computed from 3 × 3 contingency tables is excellent. It is in the range of around 75 per cent, depending on the height above the ground (2 and 30 m). *POD* are mostly larger than 66 per cent. It is as good as 93.7 per cent at 30 m for the detection of temperature inferior to 11.5 °C. The worst performances are observed in detecting the temperature in the narrow range [11.5 °,13.5 °] at 30 m (*POD* = 55.2 per cent). At every level, the *EBD* are almost all equal to 0 per cent. In all cases, we can note that the *PC* and the *POD* are greater than 33 per cent (a typical value for a random distribution for a 3 × 3 contingency table).

For the wind speed, the per cent of correct detection (*PC*) computed from 3 × 3 contingency tables is good (around 60 per cent, when the best configuration, $\Delta X = 100$ m, is used). The probability to detect the wind velocity in the three subsamples limited by the tertiles of the cumulative distribution is also good. The strongest winds are especially well detected by the model: POD_3 is equal to 74 per cent (at 10 m) and is equal to 79 per cent (at 30 m).

In all cases, the *PC* and the *POD* are greater than 33 per cent (a typical value for a random distribution). We have also proved that the configuration with the horizontal resolution of 100 m in the innermost domain provides a much better model performance in reconstructing strong wind speed.

For the wind direction, differently from the temperature and the wind speed, we have analysed 4 × 4 contingency tables. If we consider the separation in the four categories (NE, SE, SW, NW), the *PC* is 67.5 per cent. If we consider the separation in the four categories (N, E, S, W), the *PC* is 77.7 per cent. Both values are much better than for a random forecast (*PC* = 25 per cent) of a 4 × 4 contingency table. The *POD* of the wind is very good in the quadrants from which the wind flows more frequently (N and NE) with *POD*s of the order of 92.8 per cent (N) and 81.7 per cent (NE). The *POD* is not satisfactory (∼27 per cent) only in the narrow angular sector [225 °, 270 °]. However, the occurrence of wind blowing from the SW is the less probable (only 4 per cent of the time for the observed sample of 129 nights). Therefore, the impact of the model performances on this case can be considered negligible.

Concerning the RH, we show (in Appendix B) that the model is able to reconstruct the high RH (as high as 80 per cent or more), a condition for which the dome is closed.

Therefore, we conclude that the model performances in reconstructing the absolute temperature, the wind speed and direction in the surface layer [0, 30 m] above the ground is very satisfactory. The results of this study guarantee a concrete practical advantage from the implementation of an automatic system for the forecasts of these parameters.

ACKNOWLEDGEMENTS

The meteorological data set from the AWS and mast at Cerro Paranal are from the ESO Astronomical Site Monitor (Doc. N. VLT-MAN-ESO-17440-1773). We are very grateful to the ESO

Board of MOSE (Marc Sarazin, Pierre-Yves Madec, Florian Kerber and Harald Kuntschner) for their constant support of this study. We acknowledge M. Sarazin and F. Kerber for providing us with the ESO data set used in this study. A great part of the simulations are run on the HPCF cluster of the ECMWF (Project SPITFOT). This study is co-funded by the ESO contract, E-SOW-ESO-245-0933.

REFERENCES

Dali Ali W. et al., 2010, A&A, 524, A73
Finley J. P., 1884, Amer. Meteor. J., 1, 85
Jolliffe I. T., Stephenson D. B., 2003, Forecast Verification. New York: Wiley
Lafore J.-P. et al., 1998, Annales Geophysicae, 16, 90
Lascaux F., Masciadri E., Fini L., 2013, MNRAS, 436, 3147
Masciadri E., Vernin J., Bougeault P., 1999, A&ASS, 137, 185
Masciadri E., Lascaux F., Fini L., 2013, MNRAS, 436, 1968
Nurmi P., 2003, Recommendations on the verification of local weather forecasts (at ECMWF Member States), Consultancy Report (ECMWF Operations Department)
Sandrock S., Amestica R., 1999, ASM Data User Manual – 1.0, VLT-MAN-ESO-17440-1773. Munich: ESO
Schoeck M. et al., 2009, PASP, 121, 384
Stein J., Richard E., Lafore J.-P., Pinty J.-P., Asencio N., Cosma S., 2000, Meteorol. Atmos. Phys., 72, 203
Thornes J. E., Stephenson D. B., 2001, Meteorol. Appl., 8, 307

APPENDIX A: CUMULATIVE DISTRIBUTIONS

We report in this appendix the cumulative distributions of bias, RMSE and σ for all parameters and at all levels at Cerro Paranal, for the sample of 129 nights.

APPENDIX B: RELATIVE HUMIDITY

We report in this appendix the forecasted temporal evolutions for three nights in 2011, and four nights in 2013, of the RH. These nights are all the nights of 2011 and 2013 in which RH larger than 80 per cent has been observed. This is the unique interesting condition to be studied for astronomical application in this region

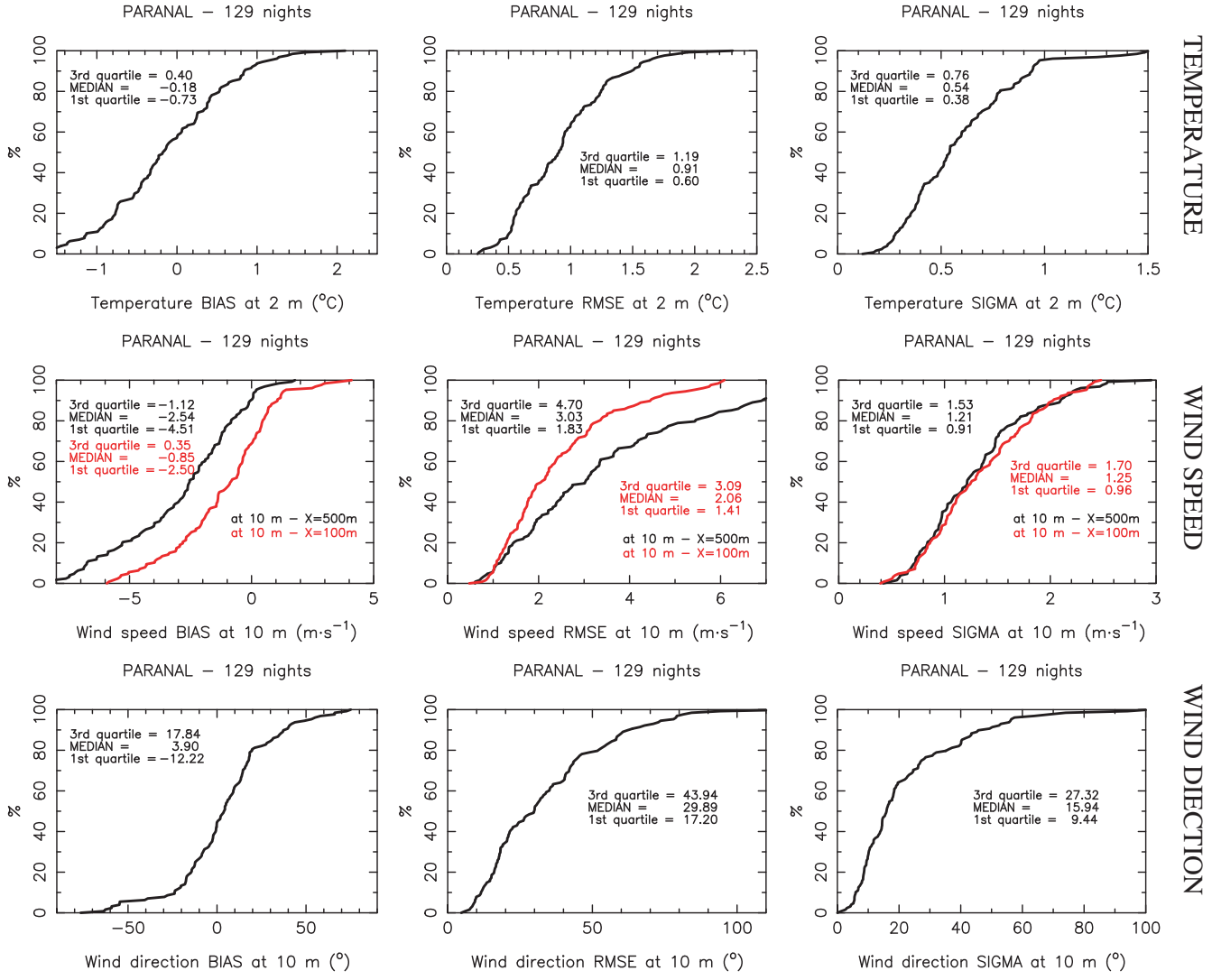


Figure A1. Cumulative distribution of the single-nights bias, RMSE and bias-corrected RMSE for the absolute temperature (top row), the wind speed (middle row, $\Delta X = 500$ m and $\Delta X = 100$ m configurations) and the wind direction (bottom row), at 2 m (temperature) and at 10 m (wind speed and direction).

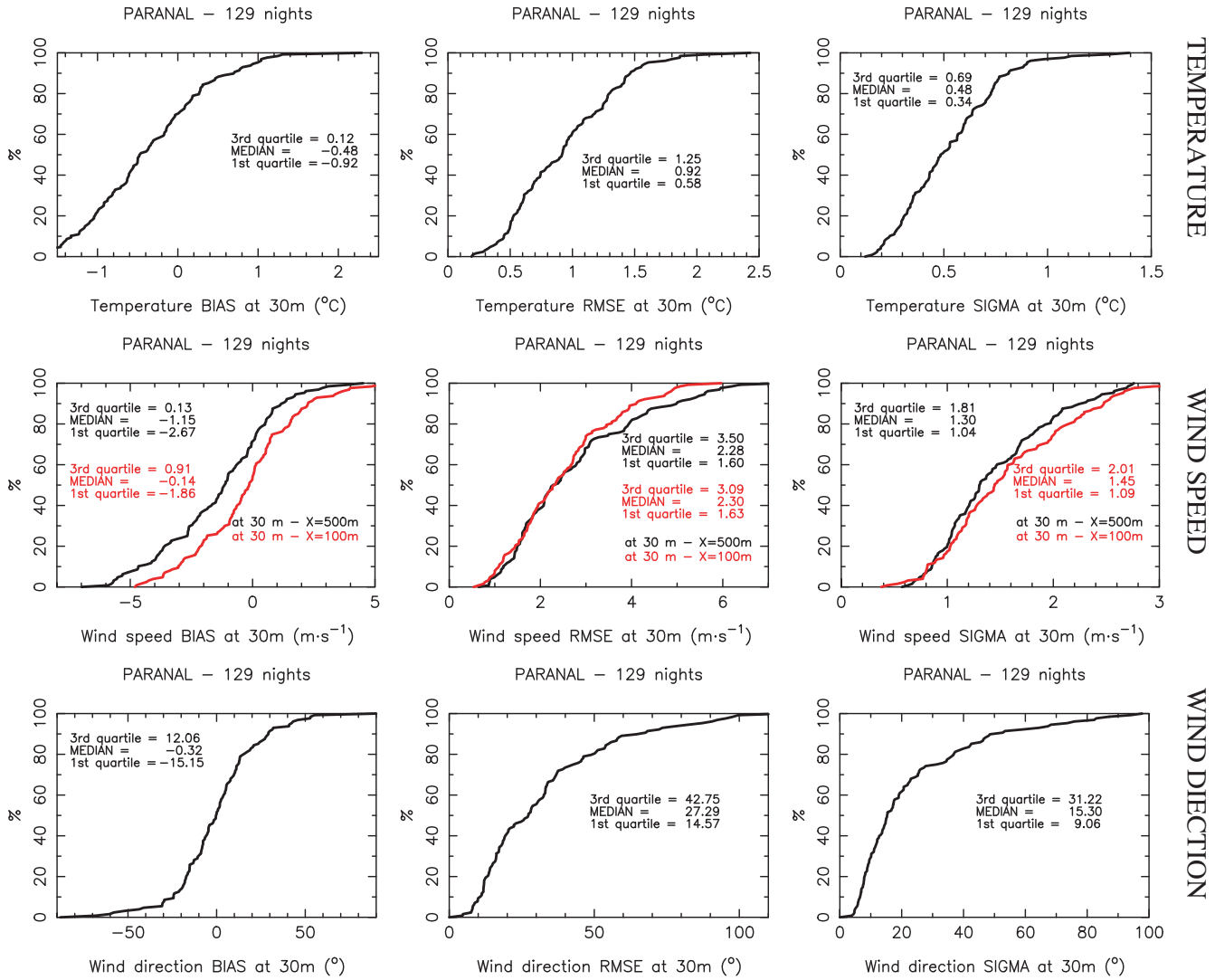


Figure A2. Cumulative distribution of the single-nights bias, RMSE and bias-corrected RMSE for the absolute temperature (top row), the wind speed (middle row, $\Delta X = 500$ m and $\Delta X = 100$ m configurations) and the wind direction (bottom row), at 30 m.

for these wavelengths. As already noted, in this region the RH is mainly of the order of 20–30 per cent and measurements for such low values are not necessarily reliable. Fig. A3 displays the temporal evolution of the RH for these nights. The forecasted RH is, in these cases, very close to the observed RH. In two nights out of three in 2011, the forecasted RH is larger than 80 per cent, as well as the observed RH. On one night (2011 July 5), the forecasted RH is slightly inferior (∼70 per cent) but still very high. The four nights of 2013 with a high observed RH are also well reproduced by the model (predicted values always between 60 and 100 per cent) even if a small dry bias is present. This means that apparently the model succeeds in predicting these critical high values, even though the frequency of occurrence of such an event is very low. Practically, the astronomer needs to know if the humidity during the night will be higher than 80 per cent, a threshold above which the dome must be closed, and thus observations are not allowed. Our results indicate that the model can give a positive forecast of the night RH useful for astronomers. Of course, it would be of interest to identify a richer sample of nights with RH as high as 80 per cent across several years and to study the behaviour of the model on a richer sample. This might permit us to have a more robust conclusion on this specific subject.

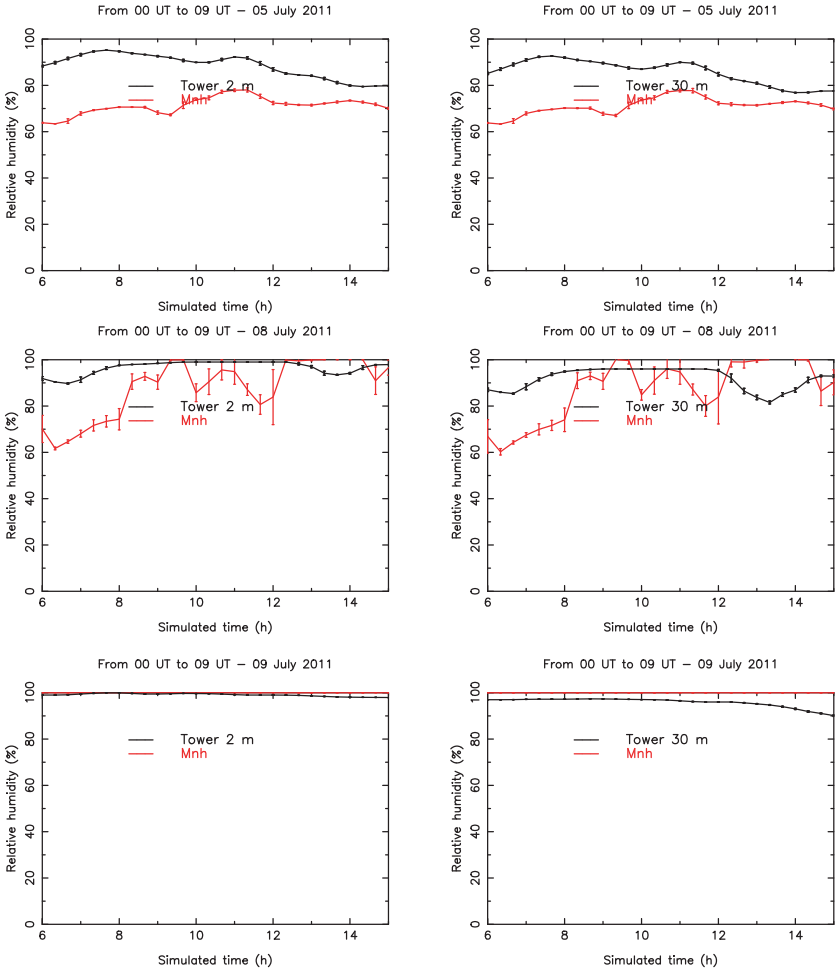


Figure A3. Temporal evolutions during the night of the RH (in per cent) at 2 m (left) and 30 m (right) at Cerro Paranal, for three nights in 2011 July. The observations are shown in black, and the model outputs in red.

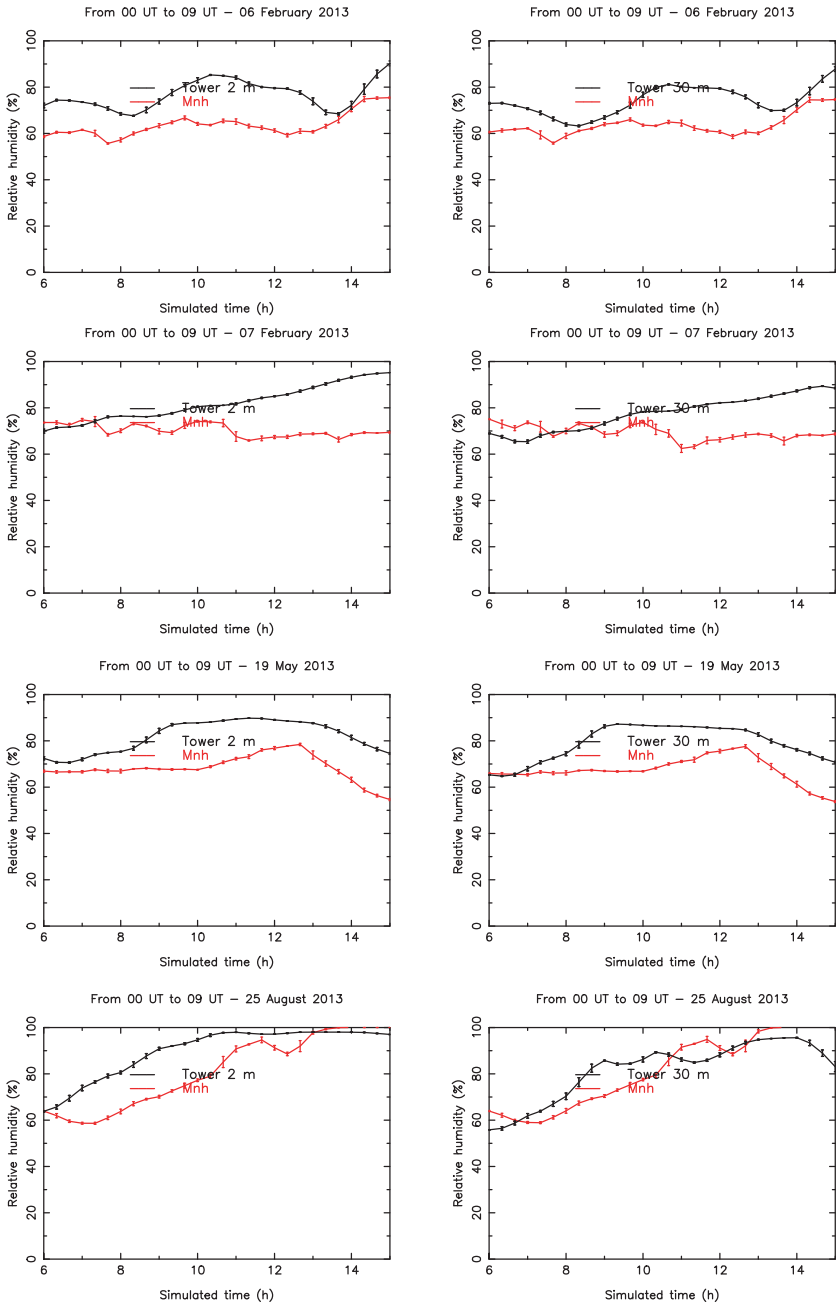


Figure A4. Temporal evolutions during the night of the RH (in per cent) at 2 m (left) and 30 m (right) at Cerro Paranal, for four nights in 2013. The observations are shown in black, and the model outputs in red.

This paper has been typeset from a TEX/LATEX file prepared by the author.