Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

MNRAS **492,** 140–152 (2020)
Advance Access publication 2019 November 29

doi:10.1093/mnras/stz3342

# Filtering techniques to enhance optical turbulence forecast performances at short time-scales

E. Masciadri ,★ G. Martelloni and A. Turchi

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

Accepted 2019 November 26. Received 2019 November 13; in original form 2019 September 25

## ABSTRACT

The efficiency of the management of top-class ground-based astronomical facilities supported by adaptive optics (AO) relies on our ability to forecast the optical turbulence (OT) and a set of relevant atmospheric parameters. Indeed, in spite of the fact that the AO is able to achieve, at present, excellent levels of wavefront corrections (a Strehl ratio up to 90 per cent in *H* band), its performances strongly depend on the atmospheric conditions. Knowing in advance the atmospheric turbulence conditions allows an optimization of the AO use. It has already been proven that it is possible to provide reliable forecasts of the OT ($C_N^2$ profiles and integrated astroclimatic parameters such as seeing, isoplanatic angle, wavefront coherence time, etc.) for the next night. In this paper, we prove that it is possible to improve the forecast performances on shorter time-scales (order of 1 or 2 h) with consistent gains (order of 2–8) employing filtering techniques that make use of real-time measurements. This has permitted us to achieve forecasts accuracies never obtained before and reach a fundamental milestone for the astronomical applications. The time-scale of 1 or 2 h is the most critical one for an efficient management of the ground-based telescopes supported by AO. We implemented this method in the operational forecast system of the Large Binocular Telescope (LBT), named Advanced LBT Turbulence and Atmosphere (ALTA) Center that is, at our knowledge, the first operational system providing forecasts of turbulence and atmospheric parameters at short time-scales to support science operations.

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

## 1 INTRODUCTION

In spite of the fact that the adaptive optics (AO) is able to achieve, at present, excellent levels of correction of the perturbed wavefront [Strehl ratio up to 90 per cent in *H* band on high contrast imaging single-conjugate adaptive optics (SCAO) systems], the AO performances are strongly dependent on the atmospheric conditions. A couple of examples are emblematic in this respect. Performances of the best SCAO systems for 8–10-m class telescopes can achieve a Strehl ratio (SR) in *H* band of 90 per cent with a seeing of the order of 0.4 arcsec but the SR can drastically decreases to 20 per cent if the seeing is of the order of 1.2 arcsec. Looking at the problem from a different point of view, if the seeing improves from 1 to 0.6 arcsec, the limit magnitude of the AO guide stars with which we obtain a SR of 30 per cent can move from 13 to 15 mag for the same instrument. Such a better seeing strongly increases the sky coverage and opens new observational windows and new perspectives in terms of scientific programs. This gain in magnitude should permit, for example, to increase by a factor of 10 the number of accessible active galactic nuclei (AGNs) from the ground (from the order of 10 to the order of 100).

The efficiency of modern ground-based astronomy, particularly if supported by AO and interferometry, is, therefore, strongly dependent on the ability to select the scientific programs to be run during a night and the set-up of instrumentation to be used during each night. This selection and management depends on the atmospheric conditions and in particular on the optical turbulence (OT) conditions ($C_N^2$ profiles) and is called, in the astronomical context, 'flexible scheduling'. All the top-class telescopes and future generation telescopes (Extremely Large Telescopes – ELTs) are planning to use the Service Mode to schedule the scientific programs. Such a mode takes into account the status of the atmospheric conditions besides the quality of the scientific programs and this permits to concretely perform the flexible scheduling. As extensively explained precedently (Masciadri, Lascaux & Fini 2013), the Service Mode is crucial and mandatory for an efficient exploitation of the best ground-based astronomical facilities.

The idea to reconstruct $C_N^2$ profiles with mesoscale non-hydrostatical models has been originally proposed by Masciadri,

★ E-mail: masciadri@arcetri.astro.it

Vernin & Bougeault (1999). The authors proposed a parametrization of the OT employing the prognostic equation of the turbulent kinetic energy (TKE). These models are certainly the most suitable models to be used for this kind of applications mainly because the general circulation models (GCMs), which are extended on the whole globe, have necessarily a lower horizontal resolution as extensively explained in Masciadri et al. (2013). This approach has been followed by successive developments in many other studies using the ASTRO-MESO-NH code (Masciadri & Jabouille 2001; Masciadri et al. 2002, 2004; Masciadri & Egner 2006; Lascaux, Masciadri & Hagelin 2010, 2011; Hagelin et al. 2011; Masciadri, Lascaux & Fini 2017) that, over the years, contributed to prove that $C_N^2$ and integrated astroclimatic parameters can be reliably forecasted for astronomical applications (at mid-latitudes and at polar latitudes) following this approach.

In most recent years other studies concerning the OT forecast on the whole atmosphere have been published using other mesoscale models (Trinquet & Vernin 2007; Cherubini et al. 2008, 2011; Giordano et al. 2013; Liu et al. 2015) or using GCMs (Ye 2011; Osborn & Sarazin 2018). Methods employed include the TKE prognostic equation approach and empirical approaches based on description of the $C_N^2$ as a function of the temperature and wind speed. Studies with GCMs are all performed with empirical approaches. A very recent analysis (Masciadri et al. 2019) confirmed the thesis that mesoscale models provide better performances than GCMs in the estimate of the seeing. A different study (Turchi et al. 2019) showed that a gain is obtained with mesoscale models with respect to GCMs by forecasting the precipitable water vapour (PWV). We remind that mesoscale models have been invented exactly to bypass intrinsic limitations of the GCMs. This is not therefore surprising.

The most recent version of the ASTRO-MESO-NH model has been used to set-up an automatic and operational forecast system for the OT and some relevant atmospheric parameters with the goal to support the observations of the Large Binocular Telescope (LBT) located at Mt. Graham (USA; Masciadri et al. 1999, 2017). LBT is a binocular telescopes, with two 8.4-m primary mirrors working in interferometric configuration; it is therefore considered the precursor of the ELTs. The operational forecast system is called the Advanced LBT Turbulence and Atmosphere (ALTA) Center,[1] it is running since a couple of years and it is in continuous evolution. The Mauna Kea Weather Center[2] is, at our knowledge, the only other similar tool existing at present time.

The approach that our team followed so far, for an operational application (see ALTA Center), consists on calculating the forecast of the OT for the next night taking care to provide the forecast a few hours before the beginning of the night, typically in the early afternoon. Hereafter, we will call this as 'standard' strategy or 'standard' configuration. Results obtained so far with this approach are very promising (Masciadri et al. 2017). The technique we proposed and implemented in ALTA Center has a few important appealing characteristics.

(i) The accuracy of the forecast system for the OT [or equivalently the root-mean-square error (RMSE)] is of the same order of the accuracy attainable with instruments. In other words, the dispersion between prediction and observations is comparable to the dispersion of observations obtained with different instruments (e.g. Masciadri et al. 2017).

(ii) It permits to have a temporal frequency of the forecast of 2 min (but this can be further reduced in case of necessity). This feature makes mesoscale models more attractive with respect to GCMs having a frequency from 1 to 6 h.

(iii) It permits to implement operational forecast systems with mesoscale models without the necessity of expensive clusters, i.e. with a relative cheap approach preserving the best model performances.

In this paper, we started from the consideration that, if we take into account the overhead necessary to carry out a scientific program and/or the logistic to switch the beam from an instrument to another one on top-class telescopes, the most critical time-scale on which to optimize observations supported by AO is of 1 or 2 h. It should be, therefore, very useful to have forecasts on this time-scale and to know if we can improve model performances with respect to the standard strategy (characterized by longer time-scales). The question is therefore: Is it possible to achieve this goal using filtering techniques such as autoregression, Kalman filter, or neural networks (also known as machine learning techniques)? The idea behind this is that the knowledge of *in situ* measurement might help in eliminating some short time-scales biases and trends that affect the forecast of atmospheric models at longer time-scales. In a preliminary analysis (Turchi et al. 2018) our team showed that such an approach might be promising.

In this paper, we concentrated our attention on the autoregressive (AR) technique that depends simultaneously on a continuous data stream of real-time measurements taken *in situ* and on time series of the atmospherical model outputs. We decided to start with the AR method because the astronomical application implies the interest for a specific point, the location of the telescope. It is highly possible therefore that observations done in just one location can be enough to achieve our objective. We considered here the ASTRO-MESO-NH forecasts done using the standard strategy available in the early afternoon. We defined the algorithm for the AR forecast, we carried out a complete quantitative analysis on the impact of such technique on the forecasts of the seeing and other relevant atmospheric parameters, and we defined the best configuration to obtain the highest gain, i.e. the best model performances. We finally implemented this system in the automatic and operational forecast system ALTA Center that has therefore, now, the possibility to provide forecasts at different time-scales: the forecasts of the next night on a time-scale of the order of 6–15 h and a forecast at short time-scale, i.e. order of 1 h.

The plan of the paper is synthesized here. In Section 2, we described the observations and in Section 3, the configuration of the atmospherical model used for this study. In Section 4, it is reported the principle of the autoregression method proposed and analysed in this paper. Section 5 reports the results of the AR technique performances in forecasting the various parameter using statistical operators of different nature applied to a statistical sample of 1 yr. In order to quantify the impact of the AR method on the forecasting performances, we compare these results with respect of the 'method by persistence', i.e. the simple use of real-time measurements. In Section 6, we describe the implementation of the method in the operational forecast system ALTA Center and finally, in Section 7, the conclusions and perspectives are reported.

## 2 OBSERVATIONS

Different typologies of observations have been considered as a reference. For the atmospheric parameters we considered the real-

[1] http://alta.arcetri.inaf.it. Also accessible through lbto.org
[2] http://mkwc.ifa.hawaii.edu/

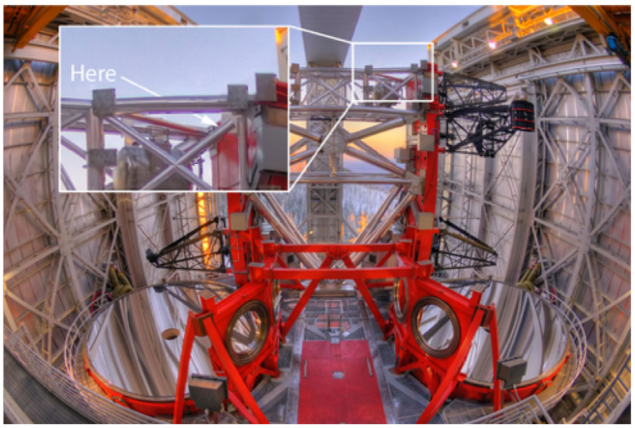


**Figure 1.** Location of the DIMM running nightly at Mt. Graham and measuring the seeing. The instrument is located on the top of the LBT dome (see zoom in the square with white frame).

time measurements routinely done with sensors placed on the roof of the telescope dome and successively stored in the LBT telemetry. As described in Turchi, Masciadri & Fini (2017), the sensors are installed on masts having different heights and located on the LBT roof (53 m above the ground). Temperature ($T$) and relative humidity (RH) sensors are located at 55.5 m above the ground (sensor at 2.5 m above the roof). Wind speed (WS) and wind direction (WD) are both measured by two different anemometers placed in two different locations on the roof that we call 'front' and 'rear' (at 56 and 58 m above the ground and 3 and 5 m above the roof). WS measurements are computed using a combination of measurements taken by the two sensors using an algorithm that takes into account the relative position of the telescope line of sight with respect to the wind direction. We refer the reader to Turchi et al. (2017) for a detailed description of the algorithm. We considered only WD measurements taken from the rear anemometer because we verified that rear and front WD measurements are statistically equivalent. Observations are stored with a frequency of around 1 s in the LBT telemetry. For the seeing, i.e. the integral of the $C_N^2$ on the whole atmosphere, we considered the measurements taken with a Differential Image Motion Monitor (DIMM), a monitor installed inside the LBT dome, close to the roof (Fig. 1), that nightly monitors the turbulence affecting the quality of images on the scientific camera. Looking at the position of the DIMM inside the telescope dome, we deduce that this instrument necessarily measures also the dome seeing (if any). On the other side, this represents the real turbulence affecting the images obtained at the focus of the telescope. Even if the LBT-DIMM does not measure the 'pure' atmospheric turbulence, it provides a more realistic estimate of the turbulence affecting the images. It has been decided therefore to assume the DIMM measurements as our reference and use these estimates for the model validation. The elimination of the dome seeing should not be trivial considering the information that are accessible but it is visibly not really very relevant. Assuming the DIMM as a reference means that we are calibrating the model to take into account a surplus of turbulence due to the dome so that the predicted turbulence is equivalent to the total turbulence affecting in reality the camera. This is obviously done in statistical terms. It is important to note that, at present, at Mt. Graham there is not a vertical profiler running nightly. This means that there are no real-time measurements of the wavefront coherence time ($\tau_0$) and the isoplanatic angle ($\theta_0$). Both parameters depend, indeed, on the integral of the $C_N^2$ on the atmosphere. We will treat therefore in this study only the seeing as integrated astroclimatic parameter.

**Table 1.** ASTRO-MESO-NH model grid-nesting configuration. In the second column the number of horizontal grid points, in the third column the domain extension, and in the fourth column the horizontal resolution $\Delta X$.

| Domain | Grid points | Domain size (km) | $\Delta X$ (km) |
|---|---|---|---|
| Domain 1 | 80 × 80 | 800 × 800 | $\Delta X = 10$ |
| Domain 2 | 64 × 64 | 160 × 160 | $\Delta X = 2.5$ |
| Domain 3 | 120 × 120 | 60 × 60 | $\Delta X = 0.5$ |
| Domain 4 | 100 × 100 | 10 × 10 | $\Delta X = 0.1$ |

## 3 MODEL

The atmospherical mesoscale model MESO-NH[3] (Lafore et al. 1998; Lac et al. 2018) has been used in this study for the forecast of the atmospheric parameters ($T$, RH, WS, and WD), while the ASTRO-MESO-NH code (Masciadri et al. 1999, 2017) has been used for the forecast of the OT, i.e. the seeing. In both cases it is possible to retrieve the spatiotemporal evolution of three-, two-, or mono-dimensional parameters over a specific limited area of the Earth. In the case of the seeing, the model calculates first a 3D map of the $C_N^2$ in a region around the telescope, and afterwards, the $C_N^2$ is integrated on the whole atmosphere (∼20 km above ground level, a.g.l.) to obtain the seeing, i.e. a 2D map. The same model configuration described in Turchi et al. (2017) has been implemented. We synthesize here the main elements to permit the readers to follow. For what concerns the MESO-NH model, the system of hydrodynamic equations is based upon an anelastic formulation that permits an effective filtering of acoustic waves. The model uses the Gal-Chen & Sommerville (1975) coordinates system on the vertical and the C-grid in the formulation of Arakawa & Messinger (1976) for the spatial digitalization. In this study, we used in the wind advection scheme the 'forward-in-time' (FIT) numerical integrator instead of the 'leap-frog' one. Such a solution allows for longer time steps and therefore shorter computing time. The model employs a one-dimensional 1.5 turbulence closure scheme (Cuxart, Bougeault & Redelsperger 2000) and we used a one-dimensional mixing length proposed by Bougeault & Lacarrere (1989). The surface exchanges are computed using the interaction soil biosphere atmosphere (ISBA) module (Noilhan & Planton 1989). The seeing ($\varepsilon$) is calculated with the ASTRO-MESO-NH code developed by Masciadri et al. (1999) and since there in continuous development by our group. The geographic coordinates of Mt. Graham are (32.70131, −109.88906) and the height of the summit is 3221 m above the sea level. We used a grid-nesting technique (Stein et al. 2000) consisting in using different embedded domains of the digital elevation models (DEM, i.e. orography) extended on smaller surfaces, with progressively higher horizontal resolution but with the same vertical grid. Simulations of the OT are performed on three embedded domains centred on the summit where the horizontal resolution of the innermost domain is $\Delta X = 500$ m (Table 1). We used four domains and a highest resolution of 100 m (Table 1) for the WS because such a configuration better reconstructs the WS close to the surface when the WS is strong. The model is initialized with analyses provided by the general circulation

[3] http://mesonh.aero.obs-mip.fr/mesonh52 – we used the MASDEV5.2 version of the code.

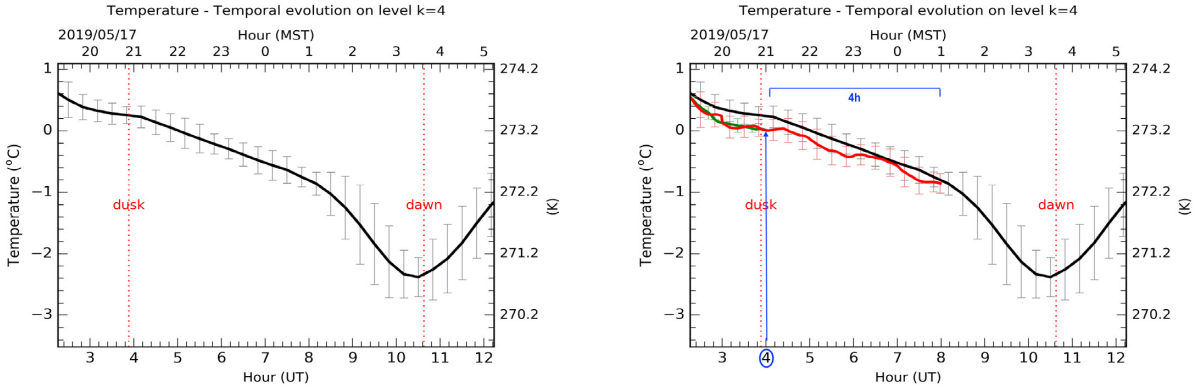


**Figure 2.** Left: temporal evolution of the forecast of the temperature for the whole night (2019 May 17) at Mt. Graham. The forecast is available at 14:00 MST (Local Time) of the day before. On the *x*-axes is reported the time expressed in UT (bottom) and local time (top). Right: temporal evolution of the forecast of the temperature available at 14:00 MST of the day before (black line); real-time measurements *in situ* (green line); forecast of the temperature using the AR technique (red line). The latter is calculated at 04:00 UT and extended on the successive 4 h (see text).

model (GCM) HRES of the European Center for Medium Weather Forecast (ECMWF) having an intrinsic horizontal resolution of around 9 km. All simulations we performed with ASTRO-MESO-NH start at 00:00 UT of the day $J$ and we simulate in total 15 h.[4] We consider data starting from the sunset up to the sunrise. During the 15 h the model is forced each 6 h (synoptic hours) with the forecasts provided by the GCM related to the correspondent hours. We consider the $C_N^2$ outputs with a temporal frequency of 2 min. All the other atmospheric parameters have a temporal frequency of the order of the second. In all cases the simulated data are extracted from the innermost domain (domain 3 or 4 – see discussion on the wind speed a few line above). We have a 54 vertical levels with a first grid point of 20 m, a logarithmic stretching of 20 per cent up to 3.5 km above the ground and almost constant vertical grid size of ∼600 m up to 23.57 km. The height of the first grid point has been fixed to be able to resolve the *in situ* measurements of the various parameters analysed in this study.

## 4 AUTOREGRESSIVE METHOD

As we said in Section 1, the goal of this study is to verify if we can improve the model performances of forecasts on time-scales of a few hours. The method that we propose to use in this paper is based on the autoregressive (AR) technique. We chose a formulation inspired by Dzhaparidze et al. (1994). The method is based on a function that depends on the difference between the real-time observations taken *in situ* that we take as a reference (i.e. we assume to be the 'truth') and on the forecasts performed by the atmospheric model. When we say 'atmospherical model' we are referring to the forecast of the model in standard configuration (see Section 3) that is available early in the afternoon of the day before.[5] The AR model $X^*_{t+1}$ calculated at the $(t+1)$ is

$$X^*_{t+1} = M_{t+1} + X_{t+1}, \tag{1}$$

where $M$ is the model output at the time $(t+1)$ and the function $X$ at the time $(t+1)$ depends on the difference between the observations and the atmospherical model outputs calculated on a polynomial function built with the addition of $P$ terms characterized by $P$ coefficient $a_i$ (called regressors) in the form

$$X_{t+1} = \sum_{i=1}^{P} a_i(\mathrm{OBS}_{t-i+1} - \mathrm{MOD}_{t-i+1}), \tag{2}$$

where the variable OBS indicates the real-time measurements and MOD the atmospheric model outputs in the standard configuration. From one side, the larger is $P$, the larger is the number of the regressors, the more accurate is the fit to the trend of the past observations. On the other side, we have interest in limiting the number of the coefficients $a_i$ to limit the computation time. We identified an optimal trade-off $P = 50$ for a temporal frequency of 1 min.

The values of the 50 regressors are obtained through a least mean square (LSM) method applied to a finite number of nights in the past, for example, the last 3, 4, 5, etc. nights.

Fig. 2 shows how the AR method works. The figure shows, as an example, the forecast of the temperature but the same procedure can be used for whatever parameter. On the left-hand side is shown the standard forecast of the night of 2019 May 17[6] that is available early in the afternoon. On the right-hand side is reported an example of the AR method applied at 04:00 UT. The black line is the standard forecast with the atmospheric model (same as the left-hand side), the green line represents the real-time measurements up to 04:00 UT (that is the present time), the red line represents the forecast calculated at 04:00 UT with the AR method for the successive 4 h. As we are interested here on studying the forecast performances on time-scales of 1 or 2 h, we considered therefore an AR forecast of 4 h that certainly covers this time-scale. We expect that the effect of the data assimilation of the local measurements provides an improvement of the forecast that is maximum close to the present time (nowcasting) and it decreases with the time up to disappear. The positive effect of the AR method vanishes after a $\Delta T$, i.e. when the performance of the AR method is equal to the performance of the atmospheric model in standard configuration. Later on, in Section 5, this aspect will be treated in a more detailed way. If the same procedure described in Fig. 2 is repeated with the

[4]To avoid misunderstandings found in the literature, we highlight that a simulation of 15 h does not mean that we need 15 h to simulate that period. It means that we reconstruct the atmospheric evolution of 15 h. The simulated time and the effective calculation time required to perform a calculation are two different concepts of 'time'.

[5]For simplicity, we will call hereafter simply 'model' the MESO-NH or the ASTRO-MESO-NH models, knowing that the first one is used for the atmospheric parameters, the second one for the OT.

[6]The date refer to the start of the night.

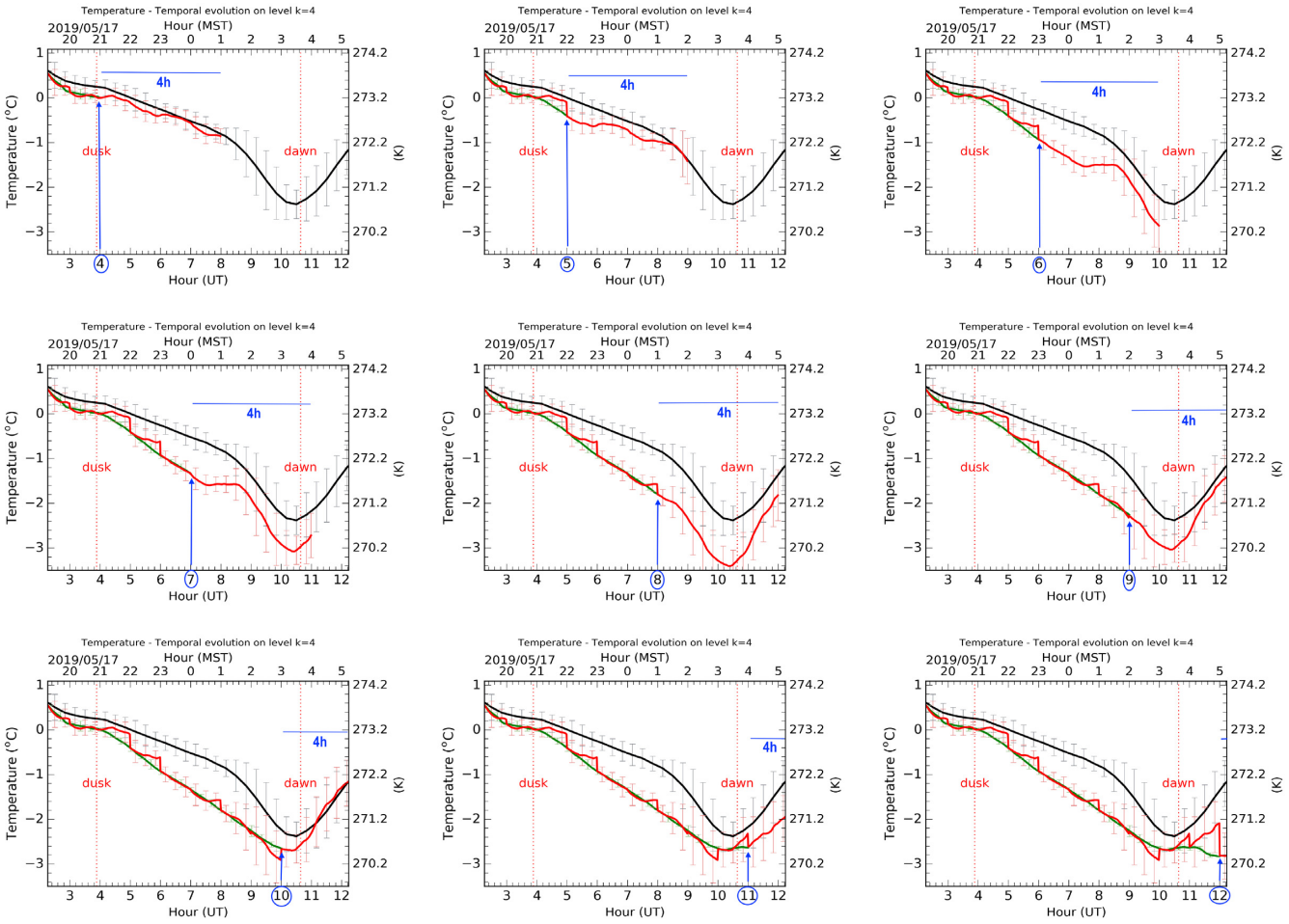


**Figure 3.** Temporal sequence of the updated forecasts of the temperature with 1 h step during the night 2019 May 17. First image on top-left is the situation at 04:00 UT of 2019 May 17, last image on bottom-right is the situation at 12:00 UT of 2019 May 17. The sequence is read by rows, from the left to the right. The black line is the forecast of the temperature available at 14:00 local time of the day before. It is therefore always the same in all the pictures. The green line is the real-time measurements. In each picture the end of the green line ends at the time in which the AR forecast is calculated. The red line is the forecast of the temperature obtained with the AR technique. The red line in the last picture (bottom-right) represents the model forecast at 1 h for the whole night.

suitable frequency during the whole night, it is possible to obtain a forecast on a time-scale of 1 h.

Fig. 3 reports the sequence of successive AR forecasts that are recalculated at each full hour during the night. The sequence has to be read from the top to the bottom, from the left to the right, following the different rows. We observe that, in each successive picture of the sequence, the green line becomes longer of 1 h and the red line, showing the forecast related to the successive 4 h, shifts of 1 h on the right. If we consider the red line of the last picture (bottom-right) extended on the whole night, we have the performance of the system on a time-scale of 1 h. We highlight that, in this computation and procedure, we take into account only data between the sunset and the sunrise.

As said previously, the unique free parameter remains the number of nights ($N$) on the past on which to calculate the values of the regressors. As we will see later on, $N = 5$ is a suitable number for our application. The whole analysis presented in Section 5 has been performed assuming this value of $N$.

## 5 RESULTS

In order to quantify the model performances of the forecasts on 1 h time-scale using the AR method built as described in Section 4, it is necessary to consider a very rich statistical sample because the AR method requires a sequence of observed data related to successive nights in which it is important to minimize the number of breaks (lack of measurements). We considered therefore data of all the nights of the whole year 2018 and we calculated the statistical operators (bias, RMSE, and $\sigma$)[7] for temperature, relative humidity, wind speed, wind direction, and the total seeing. Real-time measurements and outputs of the atmospheric model in standard configuration related to these parameters have been treated using the same procedure: we first apply a moving average of 1 h to filter out the high frequencies and put in evidence the forecast trend, we perform a resampling on a time-scale of 20 min[8] and we conclude with the calculation of the various statistical operators.

Fig. 4 shows the scattering plot related to the temperature (left), the wind speed (centre), and the relative humidity (right) obtained with an AR at a time-scale of 1 h. Fig. 5 shows the scattering plot of the WD at the same time-scale of 1 h obtained including all the data (left), filtering out all the data associated with wind speed weaker

[7]We refer the reader to Masciadri et al. (2017) for the definition of the statistical operators.

[8]A resampling on 10 min provides a very similar result.

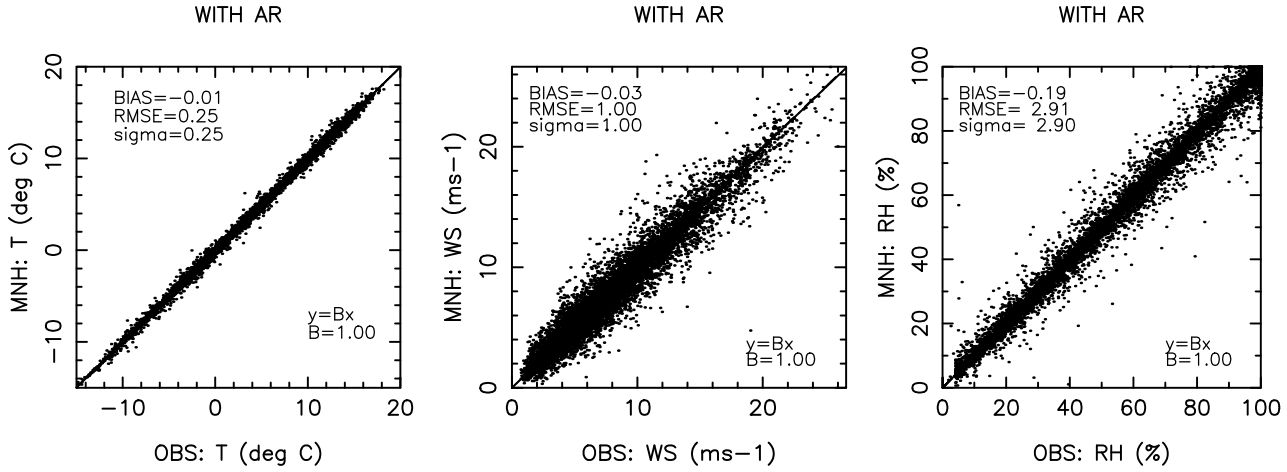


**Figure 4.** Scattering plot between observations and AR method outputs for absolute temperature (left), wind speed (centre), and relative humidity (right). Data are treated with a moving average on 1 h and resampling on 20 min. The number of nights on which the regressors are calculated is $N = 5$.

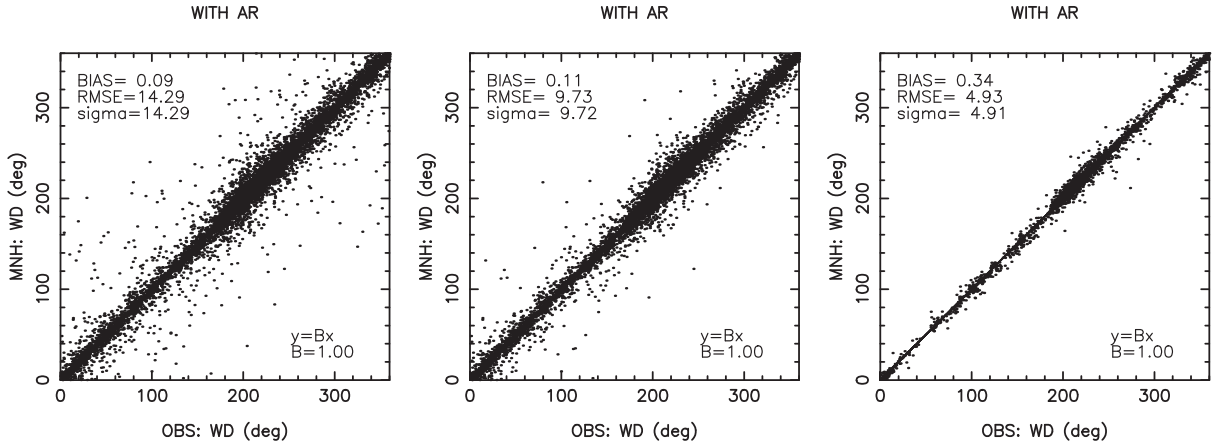


**Figure 5.** Scattering plot between observations and AR method outputs for the wind direction (left). Same results but filtering out all cases in which WS is weaker than 3 m s$^{-1}$ (centre) and weaker than 10 m s$^{-1}$ (right). $N = 5$ as in Fig. 4.

than 3 m s$^{-1}$ (centre) and filtering out all data having a wind speed weaker than 10 m s$^{-1}$. We skipped out the data associated with a WS weaker than 3 m s$^{-1}$ because under this condition it is extremely difficult (and meaningless) to quantify the WD because of the high variability of the WD. The central picture of Fig. 5 is therefore more representative for the WD than the left one. We skipped out data weaker than 10 m s$^{-1}$ to quantify the model performances in those cases that are certainly the most critical one for the ground-based observations, i.e. those in which the WS is very strong. To conclude, Fig. 6 shows the scattering plot for the seeing in the whole year (left), in the summer (April–September) interval (centre), and winter (October–March) interval (right).

Table 2 reports the RMSE obtained for the AR method at a time-scale of 1 h and with the atmospheric model in the standard configuration. As we observe that, in the standard configuration,[9] the dispersion of the seeing increases for large seeing values but the forecasts are less interesting for those cases. We decided, therefore, to consider observations below 1.5 arcsec. From a practical point of view, indeed, in the astronomical context it is poorly interesting to discriminate seeing values between 1.5 arcsec and larger values. We maintained both cases for the AR (Fig. 6) because the RMSE are very similar. We observe that, for all the parameters, the values of RMSE obtained with the AR method at a time-scale of 1 h are definitely better than for the standard configuration with consistent gains that are variable depending on the parameters between a minimum of a factor of 2.7 and a maximum of 4.9 (Table 3, first row). Those gains are definitely consistent and, at our knowledge, these model performances have never been achieved before. In Appendix A is reported a detailed description on the number of nights used to analyse this statistics for each parameter. The extremely small value of the RMSE for the temperature of the order of $0.25°C$ tells us that, with such performances in predicting the temperature close to the ground, the elimination of the dome seeing through a thermalization of the primary mirror temperature and the atmosphere inside the dome with respect to the external temperature is not a dream anymore, as declared by Racine et al. (1991).

It remains to consider how to fix the number of nights on which to calculate the regressors. Fig. 7 shows how the RMSE obtained with the AR method changes as a function of the interval of time $\Delta T$ on which we calculate the forecast and as a function of $N$. We decided to consider $\Delta T = 1$ h as a minimum value because,

[9]This feature has not been observed after the application of the AR technique as one can see in Fig. 6.

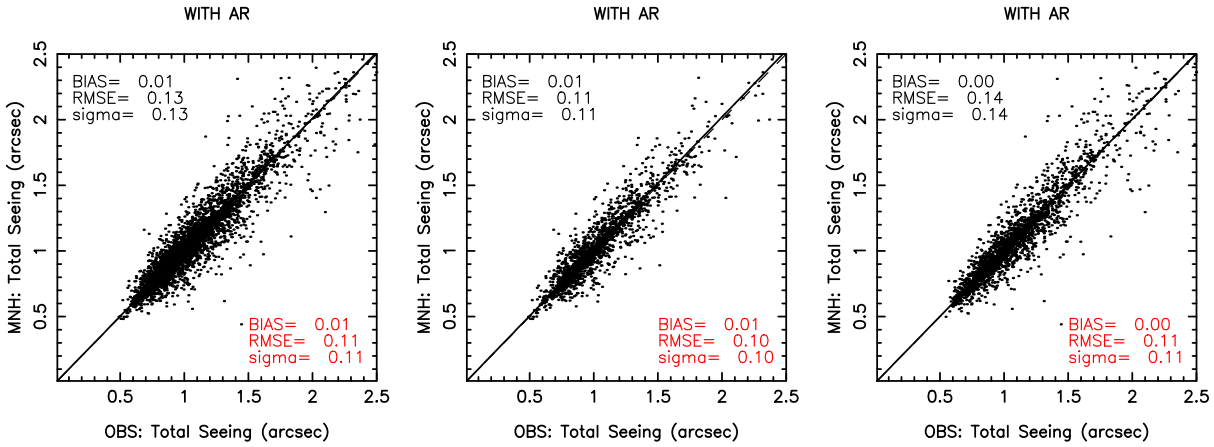


**Figure 6.** Scattering plot between observations and AR method outputs for the total seeing calculated on the whole year (left), in the summer period (centre) and in the winter period (right). Summer period is included in the April–September interval, winter period in the October–March interval. In black results considering all values, in red considering only observations below 1.5 arcsec (see discussion in the text). $N = 5$ as in Fig. 4.

**Table 2.** RMSE as obtained with the atmospheric model in the standard configuration and as obtained with the AR method on a 1 h time-scale. In the case of the seeing we considered only seeing below 1.5 arcsec. This threshold is more than representative for the AO applications and it guarantees a model performances comparable to the dispersion obtained with measurements.

| RMSE | $T$ (K) | RH (%) | WS (m s$^{-1}$) | WD (>3 m s$^{-1}$) (°) | Seeing (arcsec) |
|---|---|---|---|---|---|
| Atm. model standard config. | 0.98 | 14.17 | 2.81 | 34.71 | 0.30 |
| AR (at 1 h) | 0.25 | 2.91 | 1.00 | 9.73 | 0.11 |

**Table 3.** First row (excluding the table head): gain obtained for the RMSE for the different atmospheric and astroclimatic parameters of the AR method on a time-scale of 1 h with respect to the model standard configuration. Second row: gain using the method by persistence on the same time-scale with respect to the model standard configuration. Third row: gain of the AR method with respect to the method by persistence on the same time-scale.

| Gain | $T$ | RH | WS | WD | Seeing |
|---|---|---|---|---|---|
| AR | 3.90 | 4.90 | 2.80 | 3.60 | 2.70 |
| Persistence | 2.40 | 3.00 | 1.80 | 2.40 | 2.00 |
| AR/persistence | 1.63 | 1.63 | 1.56 | 1.50 | 1.35 |

considering the logistic requiring a change of program or the set-up of an instrument, makes poorly interesting to go below this threshold. As expected, the gain is maximum at 1 h and it decreases as $\Delta T$ increases.[10] The black line represents, for each parameter, the RMSE obtained with atmospherical model in standard configuration that is obviously constant for the whole night. We note that there is a saturation effect for $N$ equal to 4 or 5. We decided therefore to use $N = 5$ in our calculation because no further gain is visible for $N$ larger than 5. The point in which coloured lines cross the black line represents the $\Delta T$ at which the AR stops to present an improvement in the performances with respect to the standard configuration and it starts to diverge. For $\Delta T$ larger than this threshold, the standard configuration is more advantageous than the AR method. This is exactly the expected trend as the *in situ* measurements stop to have a positive influence on the forecast performances for $\Delta T$ too large.

[10]When $\Delta T = 0$, we have the nowcasting.

We can observe that the AR continues to maintain a gain different from zero up to a time-scale of the order of 4–6 h.[11]

Once analysed the gain obtained employing the AR approach, it might be interesting to quantify which is the gain on a time-scale of a few hours if we use just real-time measurements instead of the filtering techniques. We call this approach 'method by persistence'. This means that, at each full hour, the forecast extended on the successive 4 h is obtained by considering the present time measurements as a constant for all its future evolution. Fig. 8 shows the RMSE versus the $\Delta T$ obtained with the optimized AR method ($N = 5$) and the persistence method. It is clearly visible that, as expected, even if the use of pure real-time measurements provides an improvement of the forecast performances on short time-scales with respect to the standard configuration of the model, the AR method that we propose has definitely a more important gain and better performances for all the atmospheric parameters including the OT with differences (with respect to the persistence method) that are quantitatively not negligible. Table 3 (second row) reports the gain of the persistence method with respect to the model forecast in standard configuration. Table 3 (third row) reports the gain of the AR method with respect to the persistence method. Looking at Fig. 8 it is also possible to observe that, in the case of the AR method, the gain persists for a much longer $\Delta T$ with respect to the persistence approach. It is worth to note that, of course, the black line of the model forecast in standard configuration is available

[11]The reason why we display the figure only up to 4 h is to avoid a too large inhomogeneity in the statistical representativity of the samples. The number of samples for each $\Delta T$ decreases indeed, as we increase $\Delta T$.

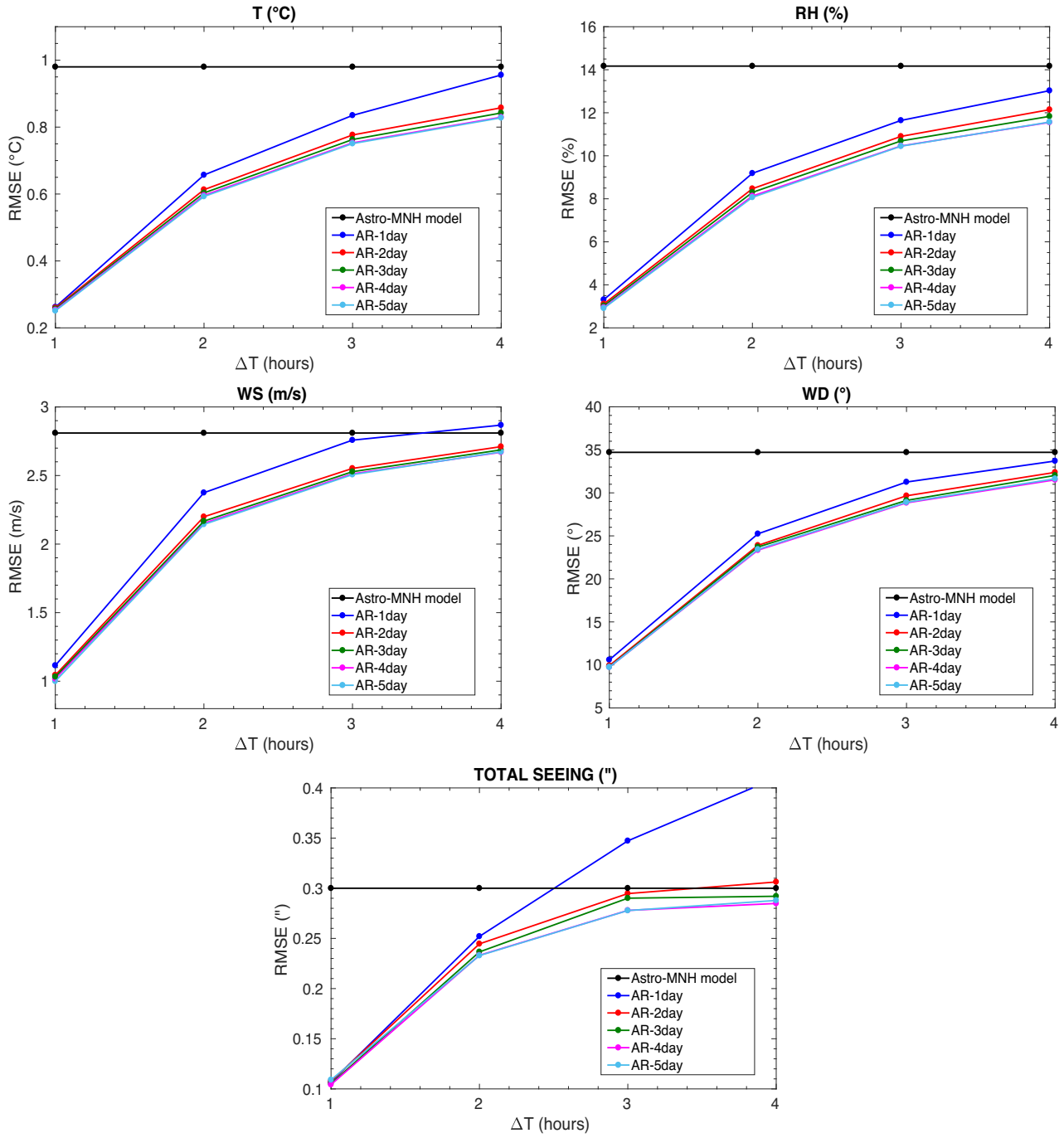


**Figure 7.** Dependency of the RMSE of different atmospheric parameters with respect to different $\Delta T$. On the $x$-axes the 'forecast time' $\Delta T = (T_f - T_i)$, where $T_i$ is the time in which the forecast is calculated and $T_f$ is the time for which the forecast refers to. Example: $\Delta T = 1$ means a forecast at 1 h calculated at $T_i$. Top-left: temperature; top-right: relative humidity; centre-left: wind speed; centre-right: wind direction of all data for which WS $> 3$ m s$^{-1}$; bottom: total seeing (we considered observations below 1.5 arcsec). The horizontal black line represents the RMSE calculated with the model in standard configuration.

much earlier than the start of the observing night. It is therefore obviously worse with respect to the other two methods. The fair comparison is therefore between the red and the blue lines.

To complete the analysis of the model performances, we finally calculate the contingency tables for each parameter from which we can retrieve the probability of detection (POD), the percentage of correct detection (PC), and the extremely bad detection (EBD). Contingency tables allow for the analysis of the relationship between two or more categorical variables. We refer the readers to Lascaux, Masciadri & Fini (2015) for a detailed definition and description of this tool. To permit the readers to follow the text we refer to Appendix B that contains a synthesis of the definitions of the statistical operators. Here we just remind the principal role of the contingency tables. Given a statistical sample of observations and predictions, the contingency tables permit to calculate the number of times in which observations and predictions fall in the same intervals of values. We used $3 \times 3$ tables for all the parameters with exception of the WD that requires a $4 \times 4$ table as it is a $2\pi$

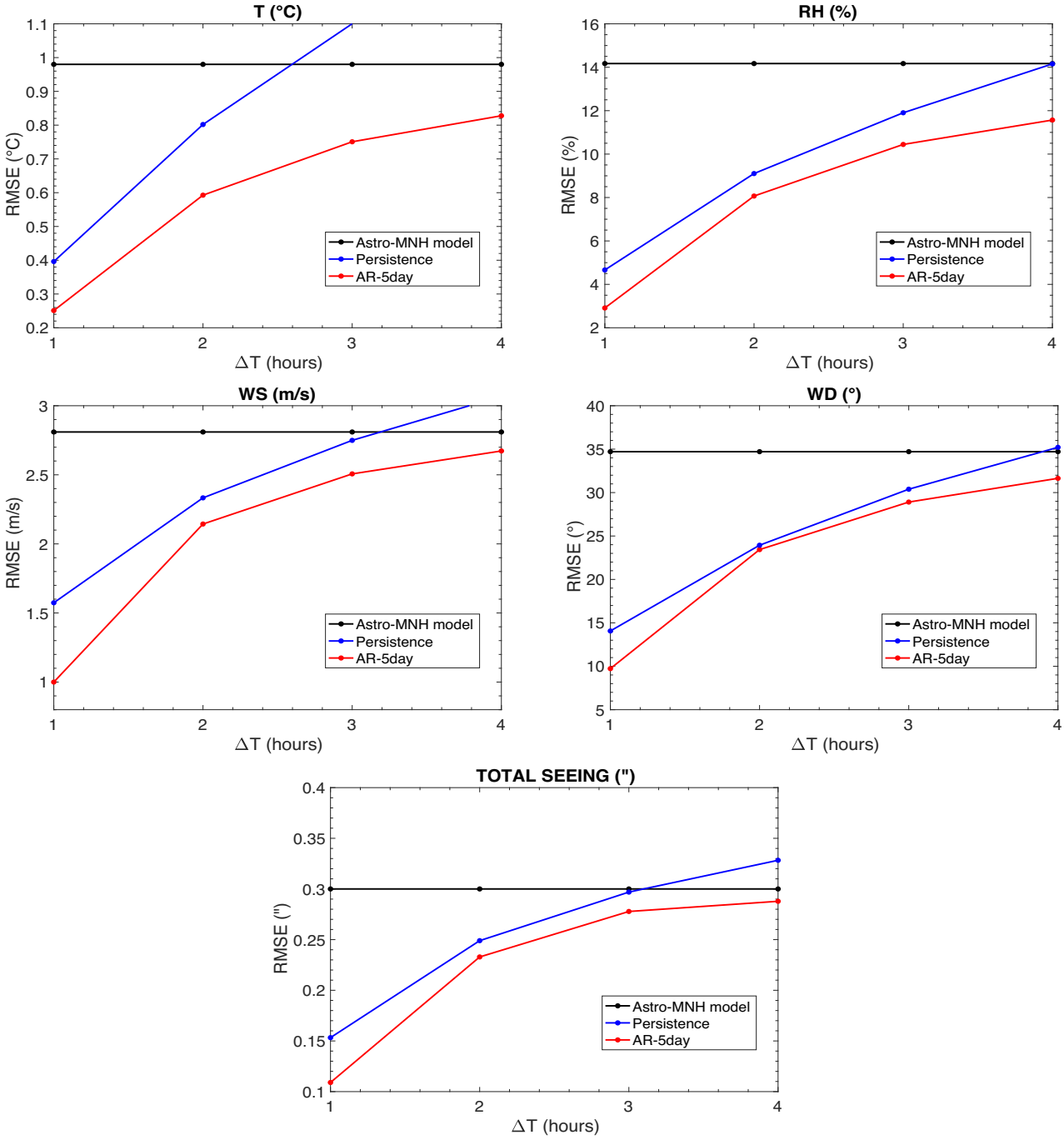


**Figure 8.** Same as Fig. 7, but shown the RMSE for the AR method (red line) and the method per persistence (blue line). The horizontal black line represents the RMSE calculated with the model in standard configuration.

periodic parameter. Starting from this distribution, it is possible to calculate the probability to detect a specific atmospheric parameter in specific intervals of values, the so-called $POD_i$, PC, and EBD. The thresholds of the intervals are calculated from the climatology of *in situ* measurements and they are, usually, the first and third tertiles of the cumulative distribution. Table 4 reports the first and third tertiles calculated on 1 yr (2018) of measurements for the different atmospheric parameters. These values are used as thresholds in this study. Tables 5–10 report the results of $POD_i$, PC, and EBD for temperature, wind speed, relative humidity, wind direction, and seeing in the different configurations: atmospheric model in standard configuration and AR at 1 h time-scale. For temperature, WS, RH, and seeing, we take $i = 1, 2, 3$; $POD_1$ is the probability to detect values smaller than the first tertile; $POD_2$ is the probability to detect values between the first and the third tertiles; $POD_3$ is the probability to detect values larger than the third tertile. For the WD we take $i = 1, 2, 3, 4$ and $POD_1$, $POD_2$, $POD_3$, and $POD_4$ are, respectively, the probability to detect a value in the range [0°, 90°],

**Table 4.** Climatology tertiles calculated on measurements extended on one full solar year (2018) for the absolute temperature $T$, the wind speed WS, the relative humidity RH, and the seeing.

| Param. | 1st tert. | Median | 3rd ter. |
|---|---|---|---|
| $T$ (°C) | 1.00 | 4.11 | 8.23 |
| WS (m s$^{-1}$) | 5.52 | 7.15 | 9.08 |
| RH (%) | 31.63 | 47.48 | 66.67 |
| Seeing (arcsec) | 0.93 | 1.05 | 1.20 |
| Seeing (<1.5 arcsec) | 0.90 | 0.99 | 1.10 |

**Table 5.** Model performances in reconstructing the absolute temperature at different time-scales: at 14 h, i.e. when we provide a forecast early in the afternoon of the day (J-1) for the next night, and at 1 h with AR. $POD_1$, $POD_2$, and $POD_3$ are the probability of detection related to the intervals $T <$ 1st tertile, 1st tertile $< T <$ 3rd tertile, and $T >$ 3rd tertile. The 1st and 3rd tertiles are shown in Table 4.

| Param. | Temperature ($T$) Forecast the day before (%) | Forecast with AR at 1 h (%) |
|---|---|---|
| $POD_1$ | 96 | 99 |
| $POD_2$ | 91 | 98 |
| $POD_3$ | 96 | 99 |
| PC | 94 | 99 |
| EBD | 0 | 0 |

**Table 6.** As Table 5, but for the wind speed (WS).

| Param. | Wind speed (WS) Forecast the day before (%) | Forecast with AR at 1 h (%) |
|---|---|---|
| $POD_1$ | 72 | 91 |
| $POD_2$ | 48 | 83 |
| $POD_3$ | 75 | 93 |
| PC | 65 | 89 |
| EBD | 2 | 0 |

**Table 7.** As Table 5, but for the relative humidity (RH).

| Param. | Relative humidity (RH) Forecast the day before (%) | Forecast with AR at 1 h (%) |
|---|---|---|
| $POD_1$ | 91 | 98 |
| $POD_2$ | 73 | 95 |
| $POD_3$ | 71 | 97 |
| PC | 78 | 97 |
| EBD | 1 | 0 |

[90°, 180°], [180°, 270°], and [270°, 360°]. The same calculation has also been done by rotating the thresholds of 45°, i.e. 45°, 135°, and 225°.

In the case of the seeing, we calculate a contingency table that takes into account an accuracy of 0.2 arcsec. In reality the dispersion between the seeing measured by different and independent instruments (such as Stereo-SCIDAR[12] and DIMM) can reach values as high as 0.29 arcsec (Masciadri et al. 2019), but we decided to use 0.2 arcsec to be more conservative and because this is a technical specification assumed in some top-class telescopes.

We observe that, for the AR forecasts at 1 h, in the case of temperature, RH, WD, and seeing, almost all the $POD_i$ are very close to the saturation (values in the [94, 99 per cent] range), i.e. with small space for further improvements. The PC is also of the same order of magnitude and the EBD basically equal to zero. The wind speed is still very good, only $POD_2 = 83$ per cent but the most important ones ($POD_1$ and $POD_3$), i.e. the probability to detect extremely weak and the extremely strong wind speed, are >90 per cent. This tool might therefore be extremely important to face the so-called *'low wind effect'*, i.e. a significant deterioration of image quality observed with high contrast imaging instruments such as Spectro-Polarimetric High-contrast Exoplanet REsearch (SPHERE)[13] (Milli et al. 2018) when the wind speed is low or absent. This condition enhances the radiative cooling of the spiders that obstruct the big telescopes pupil, creating air temperature inhomogeneities on the phase across the pupil. For WS $\leq 4$ m s$^{-1}$ we calculated that the model is able to reconstruct the WS with an RMSE $= 0.7$ m s$^{-1}$. On the other extreme, we calculated that, for WS $\geq 10$ m s$^{-1}$, the model well reconstructs the WS with a RMSE $= 1.2$ m s$^{-1}$. This means that the method is extremely efficient in predicting the conditions of strong wind speed that represent the main cause of vibration of the adaptive secondaries and/or the primary mirrors.

Looking at the same Tables 5–10, we observe that performances of the atmospherical model in standard configuration are weaker than those of the AR at 1 h as expected, but still very good. We do not comment further results found in this configuration for the atmospheric parameters as a precedent paper has been dedicated to this aspect (Turchi et al. 2017). This calculation has been repeated here (with the same statistical sample used for the AR method at 1 h) to be able to quantify the improvement in terms of model performances on short time-scales. The new result of this paper is however the estimate of statistical operators ($POD_i$, PC, and EBD) for the seeing (Table 10, second column) that reveals to be very promising, i.e. all the $POD_i$ are of the order of 97–98 per cent.

We put the accent on the most relevant result obtained in this analysis and related to the seeing. The most critical $POD_1$, i.e. the probability to detect a seeing weaker than the first tertile is equal to 81 per cent for the standard configuration and it is equal to 98 per cent for the AR method at 1 h time step. Both are well above the threshold of 33 per cent that is the percentage that corresponds to the random case and the AR method is very close to the saturation in terms of performances. Somehow weaker is the probability to detect the seeing larger than the third tertile (65 per cent) in the standard configuration as the larger is the seeing, the larger is the dispersion between observations and numerical calculation. We have here more space for further improvements of the technique.

## 6 AR FORECASTS IN THE OPERATIONAL SYSTEM

The study presented in this paper quantifies the improvements obtained in terms of performances of the AR method on short time-scales and, in particular, at 1 h. In this section, we describe how this method has been implemented in the ALTA Center, the operational forecast system conceived for the LBTO as we said in the Introduction. We chose to implement the algorithm with $N = 5$

[12]SCIDAR is for SCIntillation Detection And Ranging.

[13]High contrast imaging of the Very Large Telescope (VLT) located at the focus of the UT3.

for all the atmospheric parameters and $N = 3$ for the seeing. The choice done for the seeing aims to minimize the number of breaks in the sequence of data due to the number of nights in which the telescope dome is closed and we do not have measurements of the seeing. The algorithm has been implemented in the automatic operational system that now works nightly providing forecasts at two different time-scales: (a) a forecast of all the classical atmospheric parameters ($T$, WS, WD, and RH) and the PWV and the astroclimatic parameters ($\varepsilon$, $\theta_0$, $\tau_0$) for the successive night that is available early in the afternoon that we call standard configuration; and (b) a forecast at short time-scales. Starting from the sunset, at each full hour (e.g. ..., 02:00 UT, 03:00 UT, 04:00 UT, etc.), an AR forecast is calculated and extended for the successive 4 h. At each full hour, the forecast extended on 4 h is upgraded and shifted of 1 h as indicated in Fig. 3.

Obviously the AR method can be applied only to all the parameters for which we have *in situ* real-time measurements that, at present, are the $T$, RH, WS, WD, and seeing. So far at Mt. Graham there are no monitors that can provide real-time measurements of the isoplanatic angle ($\theta_0$) and the wavefront coherence time ($\tau_0$), as well as real-time measurements of the PWV. In the LBTO plans, it is foreseen the implementation *in situ* of a new generation of Multi Aperture Scintillation Sensor (MASS) that is under development. This should permit to extend the AR forecast at 1 h step also to these two parameters ($\theta_0$ and $\tau_0$) that are extremely important for a set of instruments supported by AO that are running at present such as LUCI[14] with the GLAO[15] system ARGOS[16] (Rabien et al. 2019) and the Large Binocular Telescope Interferometry (LBTI; Hinz et al. 2016) or those that are planned for the near future such as SHARK-VIS[17] (Pedichini et al. 2016), SHARK-NIR (Farinato et al. 2018), iLocator (Crepp et al. 2016) that will be supported by SOUL[18] (Pinna et al. 2016), the AO system that will replace FLAO. At the same time, also an instrument providing real-time measurements of the PWV such as LHATPRO (Kerber et al. 2012) is under evaluation as it should permit an upgrade of the forecasts at short time-scale of a parameter such as PWV that is critical for LBTI scientific programs such as those using the nulling interferometry in $N$ band – see HOST project (Ertel et al. 2018) looking for exozodiacal dust near the habitable zone around nearby, main-sequence stars.

ALTA Center is an operational reality since a couple of years and it is integral part of the operational observing strategy of the LBT (Veillet et al. 2016) and, since 2019 April, it provides forecasts also at short time-scale. We have almost completed the implementation of a similar automatic operational system for Cerro Paranal, the site of the VLT.[19] In this astronomical site we can access also to *in situ* real-time measurements of $\theta_0$, $\tau_0$, and the PWV. We expect therefore to be able to achieve forecasts with an equivalent high level in terms of performances for the three principal astroclimatic parameters: seeing, isoplanatic angle, and wavefront coherence time.

[14]LUCI is for LBT Utility Camera in the Infrared.
[15]GLAO is for Ground Layer Adaptive Optics.
[16]ARGOS is for Advance Rayleigh guided Ground layer adaptive Optics System.
[17]SHARK is for System for coronagraphy with High order Adaptive optics from R to K bands. Originally a unique instrument, in a successive phase it has been decided to develop two different units in the visible (VIS) and in the near-infrared (NIR).
[18]SOUL is for Single conjugated adaptive Optics Upgrade for the LBT.
[19]We point out that the operational forecast system for Cerro Paranal is not, at present, an official ESO tool but it is the result of a research study.

**Table 8.** As Table 5, but for the wind direction (WD) using as a thresholds: 90°, 180°, and 270°.

| Param. | Wind direction (WD: 90°, 180°, and 270°) Forecast the day before (%) | Forecast with AR at 1 h (%) |
|---|---|---|
| $POD_1$ | 78 | 94 |
| $POD_2$ | 75 | 93 |
| $POD_3$ | 88 | 94 |
| $POD_4$ | 57 | 93 |
| PC | 81 | 94 |
| EBD | 2 | 0 |

**Table 9.** As Table 5, but for the wind direction (WD) using as a thresholds: 45°, 135°, and 225°.

| Param. | Wind direction (WD: 45°, 135°, and 225°) Forecast the day before (%) | Forecast with AR at 1 h (%) |
|---|---|---|
| $POD_1$ | 75 | 94 |
| $POD_2$ | 62 | 93 |
| $POD_3$ | 71 | 94 |
| $POD_4$ | 84 | 93 |
| PC | 74 | 94 |
| EBD | 1 | 0 |

**Table 10.** As Table 5, but for the seeing ($\varepsilon$). Values calculated assuming an accuracy of 0.2 arcsec. We considered the seeing <1.5 arcsec.

| Param. | Seeing ($\varepsilon$) Forecast the day before (%) | Forecast with AR at 1 h (%) |
|---|---|---|
| $POD_1$ | 81 | 99 |
| $POD_2$ | 80 | 97 |
| $POD_3$ | 65 | 98 |
| PC | 79 | 98 |
| EBD | 14 | 1 |

## 7 CONCLUSIONS

In this paper, we analyse for the first time the possibility to provide forecasts of a few fundamental atmospheric parameters (temperature, wind speed, wind direction, and relative humidity) and astroclimatic parameters such as the seeing at short time-scales (order of 1 h). This time-scale is by far the most critical one for the science operations of top-class telescopes for all those programs using instrumentation supported by AO. The study is applied to Mt. Graham, the site of the LBT where we have an operational forecast system already running nightly, the ALTA Center. We proposed to use a filtering technique to provide forecasts at short time-scale, more precisely we use an AR technique based on the simultaneous use of temporal series of real-time measurements performed *in situ* and predictions provided by a non-hydrostatic atmospheric model ASTRO-MESO-NH model. We demonstrated that the model performances are improved by a not negligible quantity for all the parameters and that a gain is still visible for a few hours. The gain is maximum at 1 h and it decreases with the time until it vanishes completely when effects of the knowledge of the *in situ* observations looses its positive influence on the future. The values of this threshold are between 4 and 6 h, depending on the parameter (Fig. 7).

The AR technique for the calculation of forecast at a time-scale of 1 h produces a gain on model performances of a factor of 2.7 up to almost 5 (depending on the atmospheric parameter). We quantified the performances of the forecast method using different statistical operators. From one side the bias, RMSE, and $\sigma$. From the other side, the percentage of correct detection (PC), the probability of detection (POD), and the extremely bad detection (EBD) retrieved from the calculation of the contingency tables. For the time-scale of 1 h, using the AR filter we obtain a RMSE $= 0.25^{\circ}C$ for the temperature, a RMSE $= 2.91$ per cent for the relative humidity, a RMSE $= 1$ m s$^{-1}$ for the wind speed, a RMSE $= 9.73$ degrees for the wind direction when we filter out the wind speed weaker than 3 m s$^{-1}$, and a RMSE $= 0.11$ arcsec for the seeing.

Looking at the analysis from the point of view of the contingency tables and probability of detection, we find that all the POD$_i$ are well above 90 per cent reaching in many cases (seeing, temperature, RH, and WD) more than 95 per cent. This condition is already very close to the saturation with small space for further improvement. The WS also presents excellent performances for the POD$_i$ larger than 90 per cent. POD$_2$ is slightly weaker (83 per cent) and tells us that is slightly more difficult to discriminate between the first tertile (5.52 m s$^{-1}$) and the third tertile (9.98 m s$^{-1}$). Besides that, the system is extremely efficient in predicting the weak wind speed (with a RMSE $= 0.7$ m s$^{-1}$) and in predicting the very strong wind speed (with a RMSE $= 1.2$ m s$^{-1}$) that makes the tool very useful to face the *low wind effect* in high contrast imaging instruments (see Section 5) and to identify the interval of time characterized by very strong wind (WS $> 10$ m s$^{-1}$).

Results obtained for the OT, and more precisely with the seeing, are extremely satisfactory. We proved that the AR technique allows us to reach a RMSE of the order of 0.11 arcsec at 1 h and POD$_i$ of the order of 98 per cent. The most relevant result obtained in this study is definitely related to the seeing. The most critical POD$_1$, i.e. the probability to detect a seeing weaker than the first tertile is equal to 81 per cent for the standard configuration and it is equal to 98 per cent for the AR method at 1 h time step. This definitely represents a fundamental milestone for the implementation of the flexible scheduling of ground-based top-class telescopes.

Besides that, we quantified the gain obtained by the AR approach with respect to the use of pure real-time measurements, i.e. the persistence method putting in evidence that the percentage of RMSE gain is between 35 and 63 per cent and it is therefore far from being negligible.

Once validated, we implemented this method in the automatic and operational forecast system conceived for the LBTO named ALTA Center. The outputs of such a forecast system are currently automatically injected into the software driving the science operations at the LBTO. At our knowledge, this is the first automatic operational system providing this kind of information, at least in the astronomical context.

We are implementing a similar automatic and operational system for the VLT and, in this case, we will be able to predict at short time-scales, also the isoplanatic angle and the wavefront coherence time thanks to the presence of instrument providing real-time measurements of these parameters. It will be interesting to quantify the performances of this method on different sites. We point out that, at present, this is not an official operational ESO tool.

In terms of filtering techniques, it is our intention to refine our results to evaluate if other methods such as Kalman or machine learning, and multiple ways to use them, might provide supplementary improvements of the technique.

## ACKNOWLEDGEMENTS

This work is funded by the contract ENV001 (LBTO). The authors acknowledge the LBTO Director, Christian Veillet, and the LBT Board for supporting them in the development of the ALTA Center. The authors thank the LBTO staff, and in particular Matthieu Bec, for their collaboration in keeping synchronized the real-time LBT telemetry and the ALTA Center. The authors thank the MESO-NH users supporter team who constantly works to maintain the model by developing new packages in progressing model versions. Initialization data come from the GCM of the ECMWF.

## REFERENCES

Arakawa A., Messinger F., 1976, GARP Tech. Rep., 17, WMO/ICSU, Geneva, Switzerland
Bougeault P., Lacarrère P., 1989, Mon. Weather Rev., 117, 1872
Cherubini T., Businger S., Lymann R., Chun M., 2008, J. Appl. Meteorology Climatology, 47, 1140
Cherubini T., Businger S., Lyman R., 2011, in Masciadri E., Sarazin M., eds, Optical Turbulence: Astronomy Meets Meteorology. Imperial College Press, London, p. 196
Crepp J. R. et al., 2016, Proc. SPIE, 9908, 990819
Cuxart J., Bougeault P., Redelsperger J.-L., 2000, Q. J. R. Meteorological Soc., 126, 1
Dzhparidze K., Kormos J., Van der Meer T., Zuijlen M. C. A., 1994, Math. Comput. Modelling, 19, 29
Ertel S. et al., 2018, AJ, 155, 194
Farinato J. et al., 2018, Proc. SPIE, 10703, 107030E
Gal-Chen T., Sommerville C. J., 1975, J. Comput. Phys., 17, 209
Giordano C., Vernin J., Vázquez Ramió H., Muñoz-Tuñón C., Varela A. M., Trinquet H., 2013, MNRAS, 430, 3102
Hagelin S., Masciadri E., Lascaux F., 2011, MNRAS, 412, 2695
Hinz P. et al., 2016, Proc. SPIE, 9907, 990704
Kerber F. et al., 2012, Proc. SPIE, 8446, 84460E
Lac C. et al., 2018, Geosci. Model Development, 11, 1929
Lafore J. P. et al., 1998, Ann. Geophys., 16, 90
Lascaux F., Masciadri E., Hagelin S., 2010, MNRAS, 403, 1714
Lascaux F., Masciadri E., Hagelin S., 2011, MNRAS, 411, 693
Lascaux F., Masciadri E., Fini L., 2015, MNRAS, 449, 1664
Liu L. Y., Giordano C., Yao Y.-Q., Vernin J., Chadid M., Wang H.-S., Yin J., Wang Y.-P., 2015, MNRAS, 451, 3299
Masciadri E., Egner S., 2006, PASP, 118, 1604
Masciadri E., Jabouille P., 2001, A&A, 376, 727
Masciadri E., Vernin J., Bougeault P., 1999, A&AS, 137, 185
Masciadri E., Avila R., Sanchez L. J., 2002, A&A, 382, 378
Masciadri E., Avila R., Sanchez L. J., 2004, Rev. Mex. Astron. Astrofis., 40, 3
Masciadri E., Lascaux F., Fini L., 2013, MNRAS, 436, 1968
Masciadri E., Lascaux F., Fini L., 2017, MNRAS, 466, 520
Masciadri E., Turchi A., Martelloni G., 2019, AO4ELT 6th Edition, Quebec City, 9-14 June 2019, preprint (arXiv:1911.02819)
Milli J. et al., 2018, Proc. SPIE, 10703, 107032A
Noilhan J., Planton S., 1989, Mon. Weather Rev., 117, 536
Osborn J., Sarazin M., 2018, MNRAS, 480, 1278
Pedichini F. et al., 2016, Proc. SPIE, 9908, 990832
Pinna E. et al., 2016, Proc. SPIE, 9909, 99093V
Rabien S. et al., 2019, A&A, 621, A4
Racine R., 1991, PASP, 103, 1020
Stein J., Richard E., Lafore J. P., Pinty J. P., Asencio N., Cosma S., 2000, Meteorology Atmos. Phys., 72, 203
Trinquet H., Vernin J., 2007, Environmental Fluid Mech., 7, 397
Turchi A., Masciadri E., Fini L., 2017, MNRAS, 466, 1925
Turchi A., Martelloni G., Masciadri E., 2018, Proc. SPIE, 10703, 107036H
Turchi A., Masciadri E., Kerber F., Martelloni G., 2019, MNRAS, 482, 206
Veillet C. et al., 2016, Proc. SPIE, 9910, 99100S
Ye Q. Z., 2011, PASP, 123, 113

**Table A1.** Number or nights used for the analysis of model performances using the AR method. T hey are extracted from the sample of 365 nights of 2018.

| Year 2018 | *T* | RH | WS | WD | Seeing |
|---|---|---|---|---|---|
| Number of nights | 351 | 351 | 324 | 324 | 229 |

## APPENDIX A: SAMPLE USED FOR THE STATISTICS ANALYSIS

Table A1 summarizes the effective number of nights used for the analysis for each parameter. For temperature and relative humidity we have 351 nights, i.e. just a few nights have been missed, for WS and WD we have a total of 324 nights, and for the seeing we have a total of 229 nights. The reasons for the missing nights are of different nature. For atmospheric parameter, the reason is mainly given by sporadic failure of the sensors that, for some reason, did not work in a few nights. We note that the anemometers (providing WS and WD measurements) failed for a slightly larger number of nights than temperature and relative humidity. In the case of the seeing the justification for a smaller number of nights is mainly due to the fact that measurements are performed with the LBT-DIMM located inside the LBT dome. This means that when the dome is close for whatever reason, measurements are missed. If we consider the shut-down period of LBT (1.5 months in July–August) plus the number of nights lost because of bad weather in 2018 and we subtract to the total number of 365 in 1 yr, we find exactly the number of nights (229) reported in Table A1 that corresponds to the allocated time of LBT on 2018 (∼64 per cent of the total time).

## APPENDIX B: DEFINITIONS OF CONTINGENCY TABLES, PC, POD, AND EBD

Table B1 is an example of a generic 3 × 3 contingency table where the observations and simulations are divided into three categories delimited by two thresholds. PC, $\mathrm{POD}_i$, and EBD can be defined using $a, b, c, d, e, f, g, h, i$ (number of times in which an observation and a simulation fall inside each category) and $N$ (the total events). The percentage of correct detection, PC, is defined in equation (B1), where PC = 100 per cent is the best score; the probability to detect the value of a parameter inside a specific range of values ($\mathrm{POD}_i$) is given by equations (B2)–(B4), where $\mathrm{POD}_i = 100$ per cent is the best score. The extremely bad detection (EBD) probability is given by equation (B5), where EBD = 0 per cent is the best score. For a total random prediction and in case of a 3 × 3 contingency table we have $a = b = \cdots = i = N/9$ and PC = $\mathrm{POD}_i$ = 33 per cent and EBD = 22.2 per cent:

$$\mathrm{PC} = \frac{a+e+i}{N} \times 100;\ 0 \le \mathrm{PC} \le 100 \text{ per cent}, \tag{B1}$$

$$\mathrm{POD}(\mathrm{event}_1) = \frac{a}{a+d+g} \times 100;\ 0 \le \mathrm{POD} \le 100 \text{ per cent}, \tag{B2}$$

$$\mathrm{POD}(\mathrm{event}_2) = \frac{e}{b+e+h} \times 100;\ 0 \le \mathrm{POD} \le 100 \text{ per cent}, \tag{B3}$$

$$\mathrm{POD}(\mathrm{event}_3) = \frac{i}{c+f+i} \times 100;\ 0 \le \mathrm{POD} \le 100 \text{ per cent}, \tag{B4}$$

$$\mathrm{EBD} = \frac{c+g}{N} \times 100;\ 0 \le \mathrm{EBD} \le 100 \text{ per cent}. \tag{B5}$$

**Table B1.** 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 |

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