NOWCASTING ASTRONOMICAL SEEING AND FORECASTING TELESCOPE ENVIRONMENT FOR THE ESO VLT

Fionn Murtagh
Space Telescope – European Coordinating Facility – European Southern Observatory
(Affiliated to Astrophysics Division, Space Science Department, European Space Agency)
and
Marc Sarazin
European Southern Observatory, D-85748 Garching, Germany

November 21, 1995

ABSTRACT
We use extensive seeing and low-atmosphere meteorological data obtained at the European Southern Observatory's La Silla complex in Chile, and at the site of the Very Large Telescope at Cerro Paranal which is currently under construction.
Predictability of good seeing would allow appropriate instrumentation to be mounted in time, to make best use of the observing conditions. In line with aviation-related prediction, we have been led in the short-term to cast this problem into *nowcasting* terms, rather than forecasting.
A prototype system is set up to nowcast seeing values, given knowledge of the low-atmosphere meteorological environment. *Correspondence analysis* is used to investigate seasonal variation: in particular we seek to know what environmental variables are most important for determining seeing. Following this, a system based on *fuzzy set theory* (used in order to facilitate interpretation of the results) is used to determine *possibilities* of good seeing.
Temperature prediction would permit air conditioning in the telescope enclosure to be preset as a function of the next night's minimum expected temperature. We report results of prediction of univariate temperature measurements. These were obtained using a non-parametric (nearest neighbor) method, which we contrast to another non-parametric model (multilayer perceptron) and to an autoregressive model with exogenous inputs.

KEY WORDS Astronomy, Meteorology, Seeing, Prediction

1 The VLT Astronomical Weather Station

The VLT Astronomical Weather Station is an automatic unit operated 24 hours a day to gather informa-

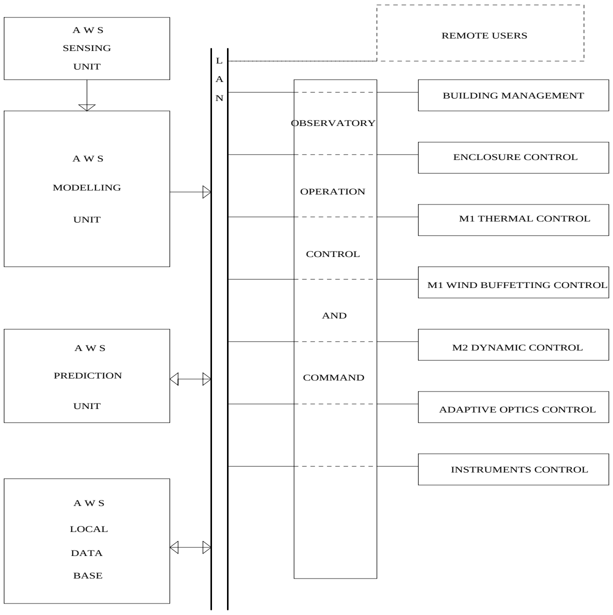

Figure 1: The function of the AWS in the VLT Observatory.

tion about the environment of the VLT observatory. It includes a sensing unit composed of a large number of stand-alone sensors and two expert systems in charge of modeling and prediction tasks. It is linked to the observatory via its local control unit performing local data base management and remote maintenance tasks. It is has three basic functions illustrated in Fig. 1, namely **sensing, modeling and predicting**.

1. The sensing unit includes all meteorological and atmospheric local sensors with their interfaces to the local control unit. A communication link with the outside world also allows to collect data from non-local sensors, i.e.: satellite data, radiosondes, professional synopsis bulletins, etc.

2. The modeling unit filters the input flow of information and produces an accurate picture of the current environmental status. In particular,

it computes secondary parameters such as water vapour, isoplanatic angle, lifetime, etc., according to specific user requirements.

3. The predicting unit takes care of the short term predictions of the variables for which local predicting models are available. It also handles the incoming external global forecasts.

The AWS is interfaced locally to expert systems composing the *Observatory Operation Control and Command* with several levels of capabilities which are, by order of complexity and priority:
a) The display and archival of information: current conditions as well as archived data will be accessible from the telescopes, the remote control center in Garching and from astronomical institutes. Each astronomical observation will be archived at the instrument level with a data header including, among other parameters, the relevant environmental conditions averaged over the exposure time.
b) The direct command of some telescope functions, e.g.: the thermal control of the enclosures, light ducts and primary mirror (with a prediction of the air temperature of the next night/hour).
c) The definition of the optimal instrumental set-up, to help the astronomer tuning its observations with the particular constraints of the environment at the telescope (e.g.: wind load, local seeing, reference star and bandwidth selection for adaptive optics) and at the instrument (e.g.: optimal exposure time for image recenter and add-up techniques, and for interferometry)
d) The definition of the optimal observing schedule, e.g.: to be able to benefit from the optimal atmospheric conditions when they occur (flexible scheduling).

2 The VLT AWS Database

The AWS sensing unit is composed of independent subsystems including complex instruments developed during the site evaluation campaign (which led to the selection of Cerro Paranal for the VLT observatory, see (Sarazin, 1990)). The site survey database which is currently used for the prediction mode is beeing continuously extended during the VLT construction phase. New monitoring instruments are currently under develpoement. The AWS will reach a final state when the VLT starts operating. It will then include:

1. Seeing and Coherence Monitor: measuring seeing with the differential image motion method. In the coherence mode, it monitors optical wavefront temporal characteristics defined as average velocity, average direction and dispersion velocity. In this mode, the index of scintillation is also available.

2. Acoustic Sounder: measuring the 3D wind pattern (horizontal velocity and direction, vertical celocity) as a function of height above ground from 50 m up to several hundred meters. It delivers also information on the vertical profile of thermal turbulence ($C_T^2(h)$) as well as its ground level value (monitored with an independent sensor).

3. Sky Monitor: determines the cloud map of the night sky (photometric or otherwise), measures infrared emissivity at zenith at several representative wavelengths and monitors the atmospheric extinction index.

4. Ground Meteorological Station: measuring classical meteorological parameters (wind speed and direction, ground and air temperature, relative humidity, pressure) and monitoring the airborne particle density (cleanliness class) of the site.

5. External Data Sources: consisting of all available information from meteorological institutes in terms of meteorological forecasts and satellite charts.

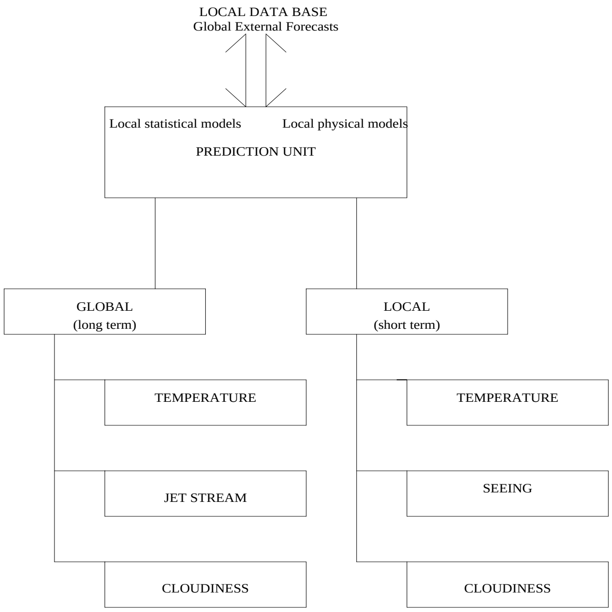

Figure 2: AWS Prediction functional Diagram.

3 The VLT AWS Prediction Unit

The prediction unit is a software item detailed in Fig. 2 where the distinction is made between between long term forecasts obtained at least one day in advance and short term forecasts from half a day to one hour in advance. It is assumed that long term or global forecasts are provided by specialized meteorological institutes. Because the observatory lies in a part of the globe poorly covered by ground stations, the spatial grid used for such forecasts will be of the order of tens of kilometers and will allow sharp temperature gradients, wind storms and large scale cloud banks to be foreseen.

Short term or local forecasts might ideally rely on physical models trained on the site orography for the micrometeorological variables (again the scarcity of local measurement points impedes proper tuning of the models). Astronomical seeing has escaped all models up to now as it depends on the vertical distribution of several parameters which do not belong to the same spatial scales. The thermal turbulence in the surface layer was successfully represented (see de Baas, Sarazin, 1991). Some efforts are currently underway at ESO, using 3D wind and turbulence vertical profiles delivered by an acoustic sounder placed either at the site or at a distance of a few kilometers upwind, to predict the behaviour of the nocturnal atmospheric boundary layer. For the higher atmosphere the first steps have been taken towards an identification of the relevant parameters (see Coulman & al., 1988).

Faced with the difficulties and incompleteness of the physical model approach, purely mathematical methods were investigated. The following sections describe two novel methods for producing local forecasts solely based on the statistical analysis of a large local database.

4 Forecasting Telescope thermal environment

Accurate temperature forecasts were obtained in (Murtagh, Sarazin, Adorf, 1992). 24-hour ahead predictions were carried out, based on temperatures at 0 (current time), 24 and 48 hours previously, and pressures at 0 and 24 hours previously. A nearest neighbor method was used.

The data used in this assessment relate to over 18000 observations carried out in austral summer 1989-1990. Many missing values, and the relative utility of the different variables, required extensive initial exploration in order to arrive at data-sets on which to carry out assessments. The number of tuples used in the results below, with no values missing, was 5260.

The following sets of tuples were derived from this data, for temperature prediction:

$$t_n, t_{n-1}, t_{n-2}, t_{n-3}, p_{n-1}, p_{n-2}$$

where

- t_n is the temperature value to be predicted;
- $t_{n-1}, t_{n-2}, t_{n-3}$ are temperature values at $n-24$, $n-48$, and $n-72$ hours;
- p_{n-1}, p_{n-2} are pressure values at $n-24$ and $n-48$ hours.

The sequence of temperatures is used in order to capture contemporary trend. Pressure is used for its information on quickly changing weather conditions, such as incoming weather fronts.

For production use, a characteristic of the data to be borne in mind is that it is constantly growing. The forecasting approach implemented must be able to cope with such a dynamic data set.

Temperature prediction was carried out using a 1-nearest neighbor method. I.e. the best matching, chronologically prior, tuple was sought. The matching was carried out on the five independent variables in the tuple, i.e. $t_{n-1}, t_{n-2}, t_{n-3}, p_{n-1}, p_{n-2}$, each variable having been range-normalized to the interval $[0, 1]$, and using Euclidean distance. A new five-valued tuple was matched against all chronologically prior prototypes. The value of t_n corresponding to the best-matching prototype provided the forecast.

Results using a multilayer perceptron approach, and results based on autoregressive modeling with exogenous inputs, were both compared with the preferred forecasting method. "Carbon-copy" prediction of temperature within half a degree centigrade of the actual temperature value was obtained in 20.5% of cases. The forecasting method employed by us found predictability within half a degree to be 62.3%. If the accuracy of prediction was relaxed to 2 degrees, then a correct temperature forecast was obtained in 85.1% of cases.

5 Nowcasting Seeing

One use of what has been achieved in (Murtagh, Sarazin, 1993) is to predict temperature, humidity, and the other meteorological variables at a few time steps into the future; and then to use the predicted environmental configuration to determine the seeing. We believe that such variables are more accurately predictable compared to direct prediction of seeing a number of time steps ahead.

The method used in this article for nowcasting is *nearest neighbor regression*, where the closest historical (or archived) meteorological configuration is checked out. The corresponding seeing value is used as an estimate of the seeing (assumed unknown) associated with a given meteorological configuration. Using similar conditions which prevailed in the past is a very intuitive approach. Variables used were the following, together with the abbreviation used in Table 1 below:

- `s1`: Wind velocity (ms^{-1}) at 10m above ground.
- `s1d`: Standard deviation of the windspeed during the averaging period (20 minutes).
- `d1`: Wind direction (degrees) at 10m above ground.
- `rh`: Relative humidity (percent) at 2m above ground.
- `t1`: Air temperature (degrees) at 2m above ground.
- `see`: The seeing is measured at 5m above ground and is defined as the full-width at half-maximum of a stellar image observed with a perfect large telescope, at $0.5\mu m$ wavelength and at zenith. One-hour averaged seeing values are used.

Among the initial data treatment steps undertaken, cases (i.e. meteorological variable, and seeing, values at a given point in time) were selected so that no measurements were missing. This reduced the many tens

Season	Good seeing given...	Conditional possibility of good seeing	Possibility of variable(s) happening
Autumn	rhL	0.614	0.506
	rhL and t1H and d1H	0.743	0.196
Winter	rhH and d1H and s1L and s1dL and t1L	0.930	0.102
Spring	d1H	0.623	0.518
	rhL	0.556	0.527
	d1H and rhL	0.669	0.286
Summer	s1dL and t1H and d1H and rhL	0.699	0.073

Table 1: Cerro Paranal results: some of the more interesting dependencies. The suffix H or L results from the categorical coding. Seasons cover the years 1989–1991.

Paranal		Actual Seeing		
		bad	interm.	good
Estimated Seeing	bad	90	63	41
	interm.	39	67	62
	good	2	4	19

Table 2: Contingency table exemplifying estimated versus actual low, intermediate and high seeing values.

of thousands of cases of the two year database to a few thousand.

A categorical coding of the measurements was carried out: we used the fuzzy coding of a variable into "high", "low" and "intermediate". "High" was coded 1, 0. "Low" was coded 0, 1. "Intermediate" was coded x, y, such that $x + y = 1$, and such that x was linearly interpolated between 1 and 0, necessitating y to be similarly linearly interpolated between 0 and 1. The division points, used to define these three categories, were the 33rd and 67th quantiles.

Variables `s1`, `s1d`, `d1` and seeing were recoded in this way, on the basis of all cases. Finally, variables `rh` and `t1` were recoded with reference to seasonally-related quantile values.

To quantify associations of interest, the first measure which will be used is the foregoing: the conditional possibility of good seeing on one or more meteorological variables. This is a fuzzy success rate, given a particular environment. A second measure will be the possibility of these meteorological variables arising in practice.

Table 1 illustrates that taking multiple meteorological variables into account can increase the possibility of good seeing; but at the expense of the possibility of such environmental conditions actually arising. Notwithstanding the clear limitation involved here, the Table indicates that remarkably positive statements can be made in regard to good seeing.

The categorical coding scheme used ("low", "high", and fuzzy "intermediate") easily allows enumeration of all possible combinations of environmental variables. For this, the "intermediate" category is taken as a third category. We used unique environmental configurations to obtain a best match relative to a given case; and to assess the best match's seeing as a prediction of the given case's seeing.

How often can we *not* make any prediction or nowcast? This translates into the question: how many unique meteorological configurations happened very rarely, – less than or equal to (say) 4 times?

For Cerro Paranal, site of the VLT, we find: Spring, 38.8%; Summer, 46.5%; Autumn, 43.8%; and Winter, 75.4%.

How often did more than 60% of the cases, associated with a unique meteorological configuration, indicate good seeing? 60% is an arbitrary cut-off. What we seek here is the number of occasions in practice that we can – with some confidence – predict good seeing.

For Paranal, we find: Spring, 12.5%; Summer, 8.9%; Autumn, 20.4%; and Winter, 32.2%.

To empirically verify these figures, and to come closer to a production system, we tested the nowcasting ability of our system in the following way. Up to 100 cases were withheld from the remaining cases, and on the basis of the latter only, the unique configurations were determined. Knowing the seeing value of the withheld test case then allowed a comparison. The contingency table shown in Table 2 displays the results obtained. If we look at the reliability of our system as indicated by how many times good seeing actually happens, relative to how many times our system estimated that good seeing should happen, we find values of 76% for Paranal (i.e., with reference to Table 2: 19/(19+4+2)).

REFERENCES

- A.F. de Baas, M. Sarazin; "The Temperature Structure Fonction for Complex Terrain", Eight Symposium on Turbulent Shear Flows, Munich, September 9-11, 1991.
- C.E. Coulman, J. Vernin, Y.Coqueugniot and J.L. Caccia, *Outer scale of turbulence appropriate to modeling refractive-index structure profiles*; App. Opt. Vol.27, n^o1, January 88.
- Murtagh, F., Sarazin, M. and Adorf, H.-M. 1992, "Statistical prediction of astronomical seeing and of telescope thermal environment", in ed. M.-H. Ulrich, *Progress in Telescope and Instrumentation Technologies* (Garching: ESO), 377–380.
- Murtagh, F. and Sarazin, M., 1993, "Nowcasting astronomical seeing: a study of ESO La Silla and Paranal", submitted.
- Sarazin, M. 1990, "ESO site evaluation for the VLT", in eds. F. Sanchez and M. Vazquez, *New Windows to the Universe* (New York: Cambridge University Press), 435–449.