RevMexAA (Serie de Conferencias), **41**, 36–41 (2011)

THE QUEUE SCHEDULING SITE SELECTION MERIT FUNCTION FOR THE ESO EXTREMELY LARGE TELESCOPE

J. Melnick[1] and G. Monnet[2]

RESUMEN

Los instrumentos robóticos utilizados en campañas de evaluación de sitios generan una gran cantidad de información, esencialmente acerca de todos los parámentros relevantes de la atmósfera. Comenzando por suposiciones relativamente genéricas, es posible capturar esta riqueza de información en una figura sencilla de mérito para cada sitio, lo cual simplifica algunas de las etapas del proceso de evaluación de sitio. Esta contribución presenta dos formalismos diferentes que fueron usados para evaluar la función de mérito de selección de sitio para el E-ELT. Ambos formalismos recaen en suposiciones acerca de las formas en que se usará el telescopio –los modos científicos de operación– pero mientras un algoritmo calcula las figuras de mérito promediadas sobre todo el tiempo de la campaña de evaluación de sitio (típicamente 2 años), el otro explora la variabilidad de las condiciones de observación durante la noche, y de noche a noche durante la campaña. Se encontró que en general, los dos métodos arrojan resultados diferentes, señalando la importancia de incluir la variabilidad como un parámetro fundamental para caracterizar los sitios astronómicos para telescopios grandes operados en modo de programación de cola. Sin embargo, los dos mejores sitios potenciales para E-ELT están clasificados como mejores por ambos métodos.

ABSTRACT

Robotic instruments deployed by modern site-testing campaigns generate enormous amounts of information about essentially all relevant parameters of the atmosphere. Starting from relatively generic assumptions, it is possible to capture this wealth of information into a single *figure of merit* for each site, which simplifies some of the stages of the site testing process. This contributions presents two different formalisms that were used to evaluate the site-selection merit function for the E-ELT. Both formalisms rely on assumptions about the ways in which the telescope will be used –the science operations modes– but while one algorithm calculates figures of merit averaged over the whole site-testing campaign (typically 2 years), the other explores the variability of the observing conditions during the night, and from night-to-night during the campaign. We find that in general the two methods yield different results, signaling the importance of including variability as a key parameter to characterize astronomical sites for large telescopes operated in Queue-scheduling mode. However, the two best potential sites for the E-ELT are ranked best by both methods.

Key Words: site testing — telescopes

1. INTRODUCTION

The characterization of sites for astronomical observation has evolved from being an art to become a bona-fide branch of astronomy. Modern site-testing instruments produce a wealth of data describing a large number of parameters ranging from temperature, humidity, atmospheric pressure, and wind velocity, to the vertical distribution of optical turbulence and the transparency of the atmosphere as a function of wavelength. In parallel, the engineering design of the largest telescopes requires, as initial input, to know the scientific goals of the telescope. Thus, during the conceptual phase, telescope design and instrument development proceed in parallel, together with the characterization of potential astronomical sites. The net result of this is that by the end of the study phase, the project has developed not only a robust engineering design for the telescope, but also a detailed plan of the scientific problems that will be addressed with the telescope, and the observing techniques that will be used to address these problems. This is very useful for choosing the most suitable site because the prior knowledge of how the telescope will be used allows to optimize the site selection process. One possible optimization method, which was used for the VLT site-selection process, is to assign to each site a figure of merit that allows to compare sites by comparing a single

[1]European Southern Observatory, D85748 Garching, Germany, (jmelnick@eso.org).

[2]European Southern Observatory and Australian Astronomical Observatory, PO Box 296 Epping NSW 1710, Australia (gmonet@aao.gov.au).

number (Sarazin et al. 1990). Clearly, condensing a vast amount of information into a single number requires a number of simplifying assumptions and the result invariably depends on these assumptions: the answer is not unique. It is very important, therefore, to understand the dependence of the figures of merit on these assumptions, while keeping in mind that the merit function is only one of the many considerations that enter in the site selection process.

The general methodology to compute the site merit function, that was also used by TMT as described by Matthias Schöck in this conference, is to assume that the telescope will be operated in a number of generic 'science modes', and to compute a merit coefficient for each of these generic modes. These merit coefficients reflect the (scientific) efficacy of operating the same telescope on different sites and the simplest metric one can imagine is just the integration time needed to reach a given signal-to-noise on a point source.

Ignoring for now instrumental noise, the signal-to-noise ratio (S/N) can be generically expressed as

$$\mathrm{S/N} = \frac{F \times t}{\sqrt{F \times t + B \times t}}\,, \tag{1}$$

where t is the integration time, F is the flux of the star, and B is the flux of the background. This S/N equation tells us that the merit coefficients (i.e. t^{-1}) scale as F in the photon-limited regime and as F^2 in the background limited case. In order to calculate the relative merit coefficients it is therefore convenient to consider observing modes where the noise is either dominated by the background, or by the photon statistics. Otherwise the calculations become somewhat more complicated, but certainly not impossible. For example, detector noise must be included to estimate the relative site merits for near-IR spectroscopy between OH lines.

1.1. *E-ELT Science Modes*

We have assumed that the E-ELT will be operated in five generic modes, which were defined considering: (a) the suite of eight instruments that were studied during the Phase A of the project; (b) the Design Reference Mission (DRM) that was prepared by an ad-hoc committee of ESO and community astronomers to define the level-1 requirements for the telescope; (c) the scientific interests of the community expressed in about 150 dry-proposals to observe with the telescope that were submitted in response to a special call issued by ESO (Kissler-Patig et al. 2009).

The five modes thus selected are,

TABLE 1

ADOPTED E-ELT OBSERVING MODES

Mode	Merit	Instruments	Weight
1	F	OPTIMOS; CODEX	0.30
2	F^2	EAGLE; HARMONI	0.15
3	F	HARMONI; MICADO; SIMPLE	0.25
4	$F^2/(1-F)$	EPICS	0.15
5	F^2	METIS	0.15

1. Seeing limited observations at optical wavelengths (the 'visible light-bucket' mode);
2. Active Optics (AO) in the near IR with large (∼50 mas) pixels, as required to observe very faint targets over relatively wide fields;
3. AO in the near IR with small pixels, as required to deliver excellent images on relatively bright targets, also over wide fields;
4. Extreme Adaptive Optics (X-AO) in the near IR to deliver the highest possible image quality over narrow fields. This mode will be mostly used to detect planets around nearby bright stars;
5. Active Optics in the thermal IR (mid-IR AO).

Table 1 summarizes the flux dependence of the merit coefficients for the 5 modes together with the instruments from the Phase A studies relevant for each mode. In the X-AO mode the noise is dominated by the residual flux in the halo of the Airy function. Since we are 'observing' point sources, the flux is proportional to the encircled energy (EE), which is determined by the Strehl ratio (S) for the AO modes, and by the seeing in the visible light-bucket mode. Thus, F must be replaced by EE or by S in Table 1 as relevant. (Notice that for the seeing-limited mode we have assumed that most of the time will be devoted to CODEX.)

In order to assign a single figure of merit to each site, and thus compute the site-selection Merit Function, we need to assign weights to each of the modes. These weights are simply the fraction of the total observing time of the telescope that will be devoted to a given mode. We used the DRM and the dry observing proposals described above to estimate these fractions that are listed in the last column of Table 1.

2. THE MERIT FUNCTION

The figure of merit for each site is calculated as,

$$M = C_0 \sum_{i=1}^{5} w_i C_i\,, \tag{2}$$

where C_i is the merit coefficient of mode i, and w_i is the weight of that mode given in Table 1. C_0 is the fraction of good weather to which we will return in the following section. In order to calculate the merit coefficients we need to use the full expression for the S/N,

$$\mathrm{S/N} = \frac{F_\lambda T_\lambda t}{\sqrt{F_\lambda T_\lambda t + B_{\rm th} t + B_{\rm OH} T_\lambda t + D_n t + \mathrm{ron}^2}}\,, \tag{3}$$

where t is the integration time, F_λ is the stellar flux at a given wavelength, and $T_\lambda(T, \mathrm{PWV})$ is the atmospheric transmission, which is a function of wavelength λ, temperature T, and precipitable water vapor PWV. $B_{\rm th}(\lambda, T)$ is the thermal IR background radiance calculated as,

$$\begin{aligned} B_{\rm th}(\lambda, T) &= \eta B(\lambda, T_A)\,, & (4)\\ &+ [1 - T_\lambda(T, \mathrm{PWV})](1-\eta) B(\lambda, T_S)\,, \end{aligned}$$

where $B(\lambda, T)$ is the Planck function and η is the telescope emissivity. T_A is the air temperature, and T_S is the equivalent Black-Body temperature of the sky, generally assumed to be $T_S = 250°$K. Since the S/N equation contains the sky background, and detector read noise and dark current (ron & D_n), it is necessary to use physical units to evaluate the Planck functions, thus,

$$\begin{aligned} B(\lambda, T) &= \frac{2\mathrm{hc}^2/\lambda^5}{e^{\mathrm{hc}/\lambda k T} - 1}\,, & (5)\\ &= 2.40 \times 10^{-7} \frac{\lambda^5_{\mu\mathrm{m}}}{e^{14388/\lambda_{\mu\mathrm{m}} T} - 1}\,, \end{aligned}$$

in units of Watts cm^{-2} μm^{-1} arcsec^{-1}.

In the near IR the sky background is dominated by the airglow emission, $B_{\rm OH}$, which can be assumed to be constant and the same for all our tested sites according to measurements done at a number of mid-latitude observatories in the world (Sánchez et al. 2008). As we have only considered the photon limited regime (CODEX) for the visible light-bucket mode ($m = 1$), the optical dark-sky background does not appear explicitly in these equations. However, the computations are done in physical units so it should be straightforward to include the background limited regime case, although the dark-sky brightness is a strong function of Solar activity and must be measured contemporaneously at all sites.

The AO end-to-end simulations of the telescope provide convenient interpolation formulas to compute the Strehl ratio as a function of the seeing (ϵ_0) and the coherence time of the atmospheric turbulence (τ_0) for the various AO modes. In the seeing limited mode the stellar profiles are assumed to be Gaussians of width given by the seeing and an outer-scale of turbulence $L_0 = 30$ m for all sites (Tokovinin 2002). Since the visible extinction and OH airglow emission were assumed to be constant during the E-ELT site-testing campaign (although the extinction does vary significantly from one site to another, mostly as a function of elevation), our merit coefficients only depend on T_A, ϵ_0, τ_0, and PWV. The first three parameters were measured with a cadence of minutes, while PWV was measured using calibrated satellite data with a typical cadence of hours (the methodology used to calibrate the satellite PWV data is described by Florian Kerber in these proceedings together with the models used to calculate the atmospheric transmission). Thus, by interpolating the PWV measurements we obtain of the order of 10^5 independent measurements of the merit coefficients for each site. Thus, the values of C_i that enter in equation 2 are the averages of all the data and therefore carry very small statistical errors, although the standard deviations are quite large. The merit coefficients are calculated for the effective wavelengths of the relevant photometric band of each mode ($UBVRI$; JHKs; $LMNQ$) and averaged assuming that an equal amount of time will be spent observing at each wavelength.

2.1. *Weather*

The weather coefficient C_0 that appears in equation 2 is the probability that the weather will be clear for astronomical observations. This means no clouds, low wind, low humidity, and low dust content. The fraction of nights free of clouds were estimated using satellites, while the wind, humidity, and dust statistics were obtained from measurements from in-situ sensors. For the E-ELT the wind speed limit is 18 m/s and the humidity limit requires the air temperature to be 2.5°C above the dew-point. There is not hard design limit for dust, so we used a rule of thumb limit, partly based on the experience of the TNG and CAMC on La Palma, which requires the density of dust particles measured by sensors near the ground to be less than 19 micro-grams per cubic meter.

Our clear weather statistics, therefore, should actually measure photometric nights, and indeed, for sites with existing observatories (Paranal, La Silla, and La Palma), there is a very good match between our site-testing values and the actual fractions of photometric nights reported for the same period by the science operations teams. Table 2 summarizes the weather statistics for our seven sites showing that the best astronomical sites on the planet have

TABLE 2

SITE MERIT COEFFICIENTS

Site	Clear Weather C_0	Seeing Limited C_1	NIR AO large pix. C_2	NIR AO diff. lim. C_3	X-AO C_4	MIR AO C_5	Average Merit
Site 1	1.000	0.652	0.619	0.887	0.882	0.664	0.836
Site 2	0.996	0.655	0.657	0.906	0.895	0.701	0.854
Site 3	0.971	0.907	0.845	0.998	1.000	0.771	1.000
Site 4	0.845	1.000	1.000	1.000	0.981	1.000	0.949
Site 5	0.742	0.778	0.717	0.941	0.937	0.649	0.681
Site 6	0.892	0.723	0.649	0.899	0.867	0.644	0.769
Site 7	0.903	0.735	0.823	0.909	0.888	0.953	0.863

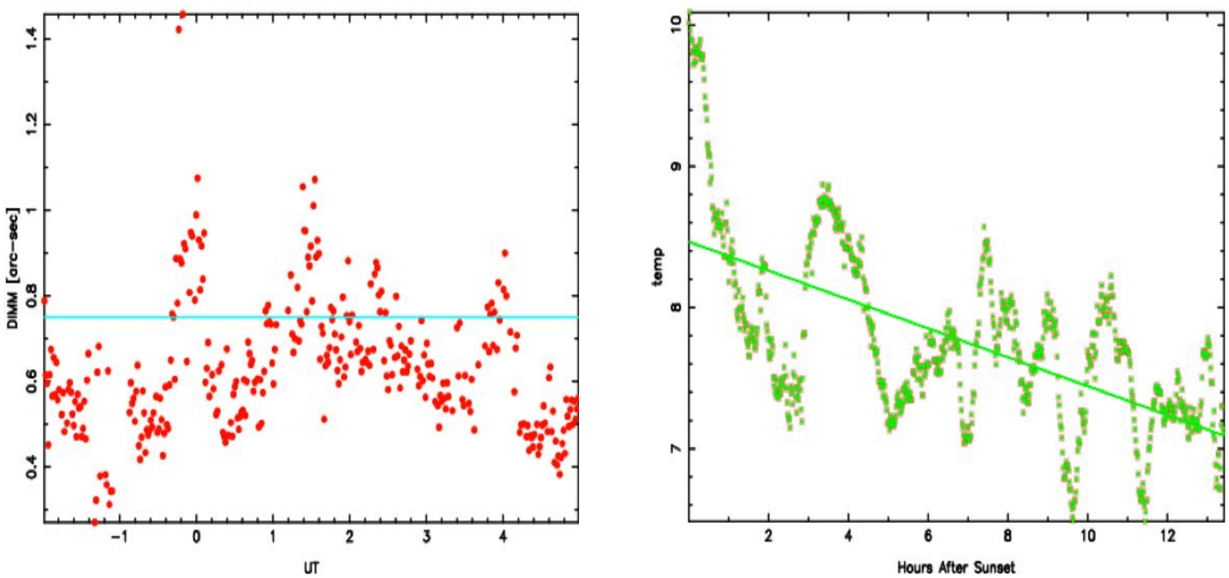

Fig. 1. One night of seeing (left) and temperature (right) measurements at one of the best astronomical sites in the world. The plot illustrates the short-term variability that characterize the atmospheric conditions of even the best sites on the planet.

photometric conditions about 75% of the time. We did not measure fractions of usable (spectroscopic) nights, but statistics from Paranal, for example, indicate that the domes are open more than 90% of the time.

2.2. *The Average Merit Function*

Table 2 presents the normalized (1.0 for the best site) merit coefficients and the combined merit figures for the 7 sites tested during the E-ELT campaign. For reasons of confidentiality we cannot disclose the actual names of the sites, so the sites are identified by proxies in arbitrary order. Since the merit coefficients are averages of a large number of measurements, they are statistically significant to 0.1%. We would have been hard-pressed, however, to discriminate between sites differing by as much as 1%, let aside ten times less. It is therefore reassuring that even the best two sites in Table 2 differ from each other by more than 5%, which is statistically highly significant (although at this level the relative merits still depend on our choice of weights and other assumptions).

3. VARIABILITY

The large variances of the merit coefficients reflect the variability of the atmospheric conditions at all the tested sites. This is illustrated in Figure 1 that shows the evolution of the seeing and the air temperature in a random night at one of our best sites. The median seeing of that night ($\sim 0.75''$ indicated by the horizontal line) was quite good, but the figure shows that at times the seeing exceeded $1''$ and reached as much as $\sim 1.4''$ for a short period of time. The air temperature also shows large excursions on top of the secular cooling from Sunset to Sunrise.

It seems relevant, therefore, to try to capture the variability of the atmosphere in the site-selection process and various ways doing this were extensively discussed during the E-ELT site testing campaign. The approach we finally adopted was to simulate the execution of realistic service observing Queues using the site-testing data, which will be discussed below.

3.1. *Queue scheduling simulations*

Having already computed the average Merit Function as discussed in the preceding sections, we wanted to explore additional constraints resulting from the inclusion of variability. We therefore assumed the simplest possible Queue scheduling situation where,

- The Queues are uniformly filled with programs exploiting all atmospheric conditions for the 5 science operations modes;
- Associate each mode with a particular set of atmospheric conditions, but including only seeing, coherence time, and precipitable water vapor. Temperature was not considered;
- Mode changes do not take any time: No observing time is lost when the observing mode is changed.

The (realistic albeit arbitrary) *orthogonal* atmospheric conditions that we associated with each mode are the following

1. Seeing Limited: $\epsilon_0 > 0.75$ & PWV > 2 mm or $\tau_0 < 2$ ms;
2. NIR AO big pixels: $0.6 < \epsilon_0 \leq 0.75$ & $\tau_0 > 2$ ms;
3. NIR AO small pixels: $0.6 < \epsilon_0 \leq 0.75$ & $\tau_0 > 2$ ms (same as Mode 2);
4. NIR X-AO: $\epsilon_0 \leq 0.6$ & $\tau_0 > 2$ ms;
5. MIR AO: $\epsilon_0 > 0.75$ & $\tau_0 > 2$ ms & PWV < 2 mm.

The simulations are performed executing 1-hour observations (OBs) according to the atmospheric conditions prevailing at the beginning of each exposure. Every 30 minutes we check whether the average seeing, τ_0, and PWV, have changed and, if appropriate, we switch mode to reflect these changes. Thus, using the average merit coefficient during a 30 min integration we quantify the scientific quality of the data, while comparing the total time spent observing in each mode with the design specifications (i.e. the weights w_i discussed above) we quantify the stability of the site.

3.2. *The Queue scheduling site-selection Merit Function*

Since we preselect the atmospheric conditions at the beginning of each exposure, there is actually relatively little difference in the 30 min-averaged merit coefficients between the different sites: the conditions to observe in a given mode are *a fortiori* the same at all sites. This is not a problem because what we actually want to measure is how well a given site matches the scientific program of the E-ELT. In other words, we want to compare the fraction of the observing programs that are completed in each of the 5 modes to the design fractions for the telescope as given by the weights w_i of each mode. Thus, the Queue scheduling Merit Function M_Q is computed as,

$$M_Q = (1 - f_{\rm ch}) C_0 \sum_{i=1}^{5} \mathfrak{m}_i(f_i|w_i) C_i^{30} \,, \qquad (6)$$

where $f_{\rm ch}$ is the fraction of incomplete OBs, w_i are the target fractions of observing time (or weights) for each mode defined above, and C_i^{30} is the merit coefficient of mode-i averaged over the 30 minute integration. The function $\mathfrak{m}_i(f_i|w_i)$ is defined as,

$$\begin{aligned} \mathfrak{m}_1 &= \min\{f_1, 0.30\} \,, \qquad (7) \\ \mathfrak{m}_2 &= \min\{f_2 + {\rm pos}(f_4 - 0.15), 0.15\} \,, \\ \mathfrak{m}_3 &= \min\{{\rm pos}[f_3 + {\rm pos}(f_4 - 0.15) - m_2], 0.25\} \,, \\ \mathfrak{m}_4 &= \min\{f_4, 0.15\} \,, \\ \mathfrak{m}_5 &= \min\{f_5, 0.15\} \,, \end{aligned}$$

where $\min(x, y)$ is the minimum of x and y, and ${\rm pos}(x) = x$ if $x \geq 0$ and ${\rm pos}(x) = 0$ if $x < 0$. Equation 6 assumes that every time we change modes in the middle of an OB we loose 30 minutes of observing time (i.e. we abort the OB). However, our simple simulations do not distinguish whether the mode has changed because conditions improved or because they degraded, whereas in realistic observing conditions an integration would be aborted only if the conditions degrade. In part this is considered by equation 7 where we redistribute the extra time in Mode-4 (when available) to Modes 2 & 3 that require less stringent seeing conditions, but in general our values of $f_{\rm ch}$ tend to overestimate the frequency of changes. On the other hand, we have assumed that it takes no time to change modes, so even if $f_{\rm ch}$ is overestimated, the total amount of time lost

TABLE 3

QUEUE SCHEDULING SIMULATIONS

Site	f_1 (0.30)	f_2 (0.15)	f_3 (0.25)	f_4 (0.15)	f_5 (0.15)	$f_{\rm ch}$	Queue Merit with $f_{\rm ch}$	Queue Merit $f_{\rm ch}=0$	Average Merit
Site 1	0.581	0.142	0.142	0.081	0.195	0.136	0.611	0.566	0.836
Site 2	0.517	0.151	0.151	0.055	0.276	0.156	0.600	0.569	0.854
Site 3	0.324	0.216	0.216	0.329	0.130	0.200	1.000	1.000	1.000
Site 4	0.325	0.200	0.200	0.345	0.130	0.205	0.932	0.937	0.949
Site 5	0.520	0.201	0.201	0.231	0.047	0.148	0.582	0.546	0.681
Site 6	0.624	0.246	0.246	0.098	0.032	0.155	0.588	0.557	0.769
Site 7	0.479	0.191	0.191	0.064	0.265	0.177	0.667	0.648	0.863

may actually be underestimated. This is a complicated issue and for the sake of simplicity we have computed the Merit Function with and without including $f_{\rm ch}$ under the assumption that more realistic scheduling algorithms should yield merits between these two extremes[3].

The results are presented in Table 3 where the fractions of time spent in each mode are compared with the target frequencies. The last three columns give the simulated (Queue) Merits and the Average Merits from Table 2. We remark that we had to alter some of the numbers in Table 3 in order to convey the correct information without disclosing the names of the sites. Table 3 shows that, except for the best two sites, the differences between the Average Merits and the Queue Merits can be substantial, although, as already mentioned, the results are rather sensitive to our choice of thresholds between the different modes. The information about variability is conveyed by the losses of observing time due to incomplete (or rejected) OBs – $f_{\rm ch}$. Interestingly, the two highest ranked sites also appear to be the ones most affected by variability. This is partly due to our choice of thresholds for mode changes, and partly due to the fact that our simplistic simulations do not distinguish whether modes have changed because conditions degraded or improved. Clearly, the fraction of incomplete OBs in underpopulated modes ($f_i \ll w_i$) will be very low. Table 2 shows that variability does not change the overall rankings significantly, but can lead to substantial losses of observing time, which in principle can be estimated using more realistic scheduling strategies.

[3]An alternative method of simulating the overheads is to always expose for 60 min and to reject OBs that do not meet the specifications. We find that both methods gives essentially similar results.

4. CONCLUSIONS

We used two fundamentally different formalisms to compute the site-selection merit function for the E-ELT. The first approach, which is similar to that adopted by TMT, makes use of the average merit coefficients over the site-selection campaign. Instead of using averages, the second method explores how changes in the atmospheric conditions affect the ability of the observatory to fulfill its scientific mission. Both methods make the bold assumption that we know today how the telescope will be used ten years from now, which may not be realistic but is the best we can do. We find that both methods give basically identical results for the two top-ranked sites, and similar overall rankings, but with a larger dispersion for the Queue method. Our simplistic Queue-scheduling simulations indicate that, even under very idealized conditions, a substantial fraction of the observing time will be lost due to incomplete or rejected OBs if short time scheduling reacts to changes in the atmosphere. The ability to predict the seeing over time scales comparable to the typical exposure times, therefore, would be required to optimize the short-term scheduling. Therein lies the ultimate challenge of mesoscale atmospheric models to predict C_n^2.

REFERENCES

Kissler-Patig, M., Küpcü-Yoldaş, A., & Liske, J. 2009, The Messenger, 138, 11
Tokovinin, A. 2002, PASP, 114, 1156
Sánchez, S. F., et al. 2008, PASP, 120, 1244
Sarazin, M. 1990, ESO VLT Report No. 62