Execution of queue-scheduled observations with the Gemini 8m telescopes

Phil Puxley

Gemini 8m Telescopes Project
950 N. Cherry Ave., Tucson, AZ 85741

ABSTRACT

Queue scheduling will be one of the major modes of operation of the Gemini 8m telescopes in which scientific programs will be carried out on behalf of applicants by Gemini staff. Use of a substantial fraction of the available telescope time in this manner will permit access to the exquisite conditions of image quality and background which the telescopes are designed to exploit as well as matching the demands of individual observations to the current conditions. In a previous presentation (SPIE 2871) the classical scheduling and loading of the queue were described. In this paper we discuss the philosophy and parameters which define its execution. Results from detailed simulations of the queue execution process are presented.

Keywords: Gemini telescopes, telescope scheduling

1. INTRODUCTION : QUEUE OBSERVATIONS ON GEMINI

The Gemini telescopes have been designed to deliver exquisite images and to have a low background flux in the thermal infrared. The formal requirement[1] is for a system image quality with an 85% encircled energy diameter of 0.25 arcsec at a wavelength of 2.2mm; the expectation is that the total emissivity will be close to the goal of 2%. In order to exploit these capabilities a significant fraction of the observing time on Gemini will be operated in a queue-scheduled mode. In this mode, applicants whose scientific programs have been recommended for time by their national time allocation committees (TAC) will define the details of each observation to be executed on their behalf by Gemini staff. Similar flexible scheduling schemes are now being studied experimentally or planned for other large telescopes[2, 3, 7].

In this paper we describe the two phases of the time application procedure (section 2), the decision process by which the next queue observation is drawn from the pool of potential observations (section 3), simulations of this process (section 4) and outline aspects of its implementation (section 5). This paper complements descriptions of the procedure for merging programs from the various Gemini partners into a single ranked list, the allocation of classically-scheduled time and the loading of queue programs which were presented at an earlier meeting ([4], hereafter Paper I). The implementation of pre-planned observing is described in more detail in [5].

2. THE TIME APPLICATION PROCESS

Gemini queue applications will be made via a two-stage process. In Phase I, applications are made to the responsible body (TAC or national office) within each of the Gemini partners. These applications will closely resemble those commonly in use at other national facilities and include the scientific case, technical and time justification, target list and scheduling constraints. In this case the scheduling constraints will define the minimum requested site conditions (e.g. image quality, IR background, cloud cover, sky brightness) as well as any time constraints. For many programs it will also be necessary to verify that suitable guide stars are available to ensure that the requested image quality is achievable. Given that only a fraction of Phase I applications will be successful, the intent is to require applicants to provide only that information necessary for scientific and technical evaluation of the program and to establish whether the program can be loaded into the queue. The ranked lists of proposals produced by each TAC are merged into a single queue as described in Paper I.

To provide a responsive and efficient application process, all Phase I applications recommended for time by the partner TACs will be transmitted to Gemini in a standard electronic format. Relevant information can then be readily extracted and will be used to populate the observing database[5]. In Phase II, successful applicants interact directly with the observing database via a graphical user interface to completely define their observations, select and sequence instrument and telescope configurations, refine positions etc.

The nominal scheduling period for the Gemini telescopes is expected to be a 6-month semester. However, given the flexibility inherent in supporting queue-mode observing, Gemini will be an excellent tool for carrying out target-of-opportunity and other short-timescale observations[10]. Consequently it will be possible for partners to set aside queue time at the start of the semester for as yet unspecified targets. (Gemini partners may also allocate blocks of classically scheduled time in support of classical service-mode observations. These would be pre-defined like queue observations but be carried out by visiting observers on specific dates in the telescope schedule). The partners could process such quick-response applications more frequently or, alternatively, external triggers for submitted proposals (e.g. detection of a g-ray burst) may be set.

3. QUEUE WEIGHTING FUNCTION PARAMETERS

There are many factors that might be incorporated into a decision of the best observation to execute next. These include the scientific ranking of the program, the closeness to which the requested and current site conditions match, the position of the target in the sky, the amount of time remaining in the semester when target can be acquired, the status of other observations in the program, the relative usage of time amongst the partners. In this section we describe the baseline properties adopted for each of these factors and how they might be combined in an overall weighting function.

3.1 Scientific ranking bands

As described in Paper I, the programs recommended for time by the partner TACs are merged into a single list ranked on the basis of scientific quality. If observations were to be drawn from the queue in strict rotation starting at the top and executed provided that conditions were at least as good as those requested then there would be good maintenance of the TAC rankings but a poor match to site conditions. Conversely if the queue were treated as a single pool from which any observation could be drawn then there would be a good correspondence between requested and current conditions but at the expense of the scientific ranking. A solution is to divide the queue into a number of "scientific ranking bands" which contain programs of (assumed) equal scientific quality. (This also helps negate the difficulty that TACs might experience of distinguishing uniquely between different scientific programs). In this scheme, observations are drawn from the top ranking band provided that current conditions are at least good enough. Only if no observation meets this constraint does the search proceed down to the next band. Thus the bands effectively have infinite weight.

In addition to matching site conditions, there is also a trade-off between the number of programs in a band and the number of completed programs at the end of the semester. Results from the simulations described in section 4 indicate that the optimum number of programs per band is 15-20, for the average of 6 observations/program and with 3 out of 4 instruments generally available. This band length has the additional benefit of typically enclosing at least one program from each partner in each band.

3.2 Match to conditions

Matching the requested observing constraints to the current conditions is not limited solely to the basic site properties of image quality, IR background, cloud cover and sky brightness (i.e. lunar phase and distance). Observations may have a temporal constraint (e.g. occultations or periodic monitoring) or some other limitation. One specific example is multi-object spectroscopy with GMOS in which differential atmospheric refraction and instrumental flexure impose restrictions on the zenith distance over which a particular slit mask may be used.

Extensive site monitoring equipment[6] will be installed to provide the input data against which requests are tested. As an illustration, the image quality delivered in the telescope focal plane will depend on the wind speed and direction, telescope zenith distance (due to static aberrations), temperature differential between mirror and ambient as well as the natural seeing. Furthermore, the requested image quality will depend on the wavelength of observation; Gemini is expected to be diffraction limited at 10 and 20mm under almost all conditions for example.

In general the weight adopted for the match to conditions is an inverse proportionality to their frequency of occurrence. This ensures that observations requiring the best image quality or IR background, conditions which the telescopes are designed to exploit, receive the highest weight. Example weights for the basic site properties used for the simulations in section 4 are given in Table 1.

Site Property	Conditions	Relative Weight	Time Category
Image quality	Best 20%-ile	25	A
	Best 50%-ile	4	B
	Unconstrained	1	C
IR background	Best 20%-ile	25	A
	Best 50%-ile	4	B
	Unconstrained	1	C
Cloud cover	Photometric	2	B
	Non-photometric	1	C
Moon	Dark	2	B
	Bright	1	C

Table 1: weights for site properties

The final column in Table 1 shows the observing time categories used for tracking the time usage amongst partners (see section 3.5). Since the definition of which conditions are 'desirable' may differ, they have been grouped according to their frequency of occurrence. Category "A" time is the most desirable and corresponds to the best image quality or IR background (approximately 36% of usable time if they are uncorrelated). "B" time is also desirable in at least one of the site properties with better than median image quality or IR background or is dark or photometric (approximately 90% of the usable time). "C" time is the least desirable with conditions that have worse than median image quality and IR background, are bright and non-photometric.

3.3 Target position

The location of an object in the sky affects not only the delivered image quality (see section 3.2) but also defines the duration it is visible on any given night, in Gemini's case above an elevation of 15 degrees, and for what fraction of the semester. This leads to two parameters that we refer to as the declination and long-term visibility weights.

The declination weight arises because southern objects (when observed from Gemini North) are visible for a shorter period of time than northern objects and therefore require an enhanced weight if they are to have a comparable likelihood of execution. The simple scheme adopted for the simulations reported in section 4 is shown in Fig. 1.

This figure illustrates how the weights of two targets with the same RA but different declinations change with their trajectory across the sky. Below an elevation of 15 degrees the targets are inaccessible. Above this elevation they have a weight which is a function of their declination with values chosen to produce approximately equal areas of (weight * duration available) i.e. comparable average likelihoods of execution. Note that execution in practice would also depend on the delivered image quality meeting or exceeding the set constraint. The weights adopted for the simulations are given in Table 2.

As the semester progresses, any object not yet observed has a decreasing opportunity for observation before it is no longer accessible. Thus it is desirable to have a weight which gradually increases for each object as time advances.

Care must be taken that this weight does not result in objects tending only to be observed at high western air masses. In the simulations the values of the long-term visibility weight ranged from 1-1.5.

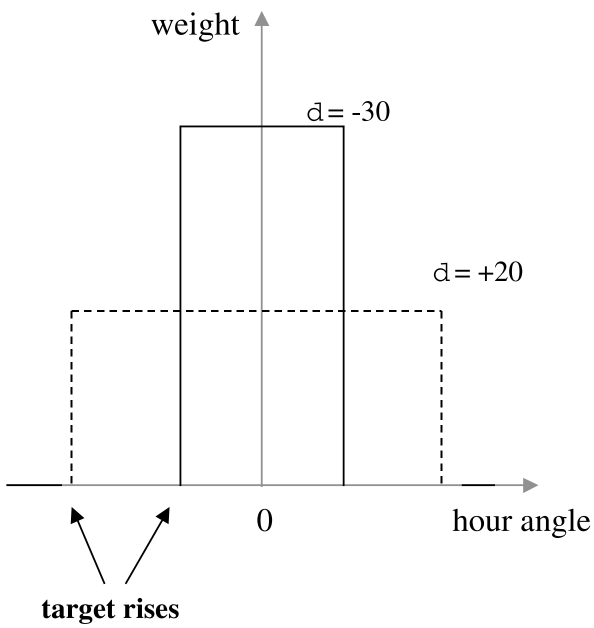

Figure 1: weights for two targets of different declination as a function of position on the sky

Declination	Weight
d ‡ −10	1.0
−10 > d ‡ −25	1.3
−25 > d ‡ −30	1.6
−30 > d ‡ −35	2.0

Table 2: weights for declination dependence

3.4 Program status

It is generally considered desirable to complete all of the observations in fewer programs rather than several observations in more programs. However, one of the early results from queue scheduling experiments on the WIYN telescope was that an inordinate amount of time, compared to the scientific return, can be expended attempting to complete the last few observations in programs which contain long target lists[8]. Thus a weight was introduced into the simulations to favour the execution of observations in programs which had already been started, but with a monotonic decrease in weight after completion of some fraction of the program (for long programs only). Fig. 2 illustrates the weighting function adopted for the simulations. The weight has an initially neutral value, is boosted after the first observation has been executed, then, if the total number of observations is larger than some adopted typical value (10, in this case), decreases after completion of one half of the observations back to the initial value.

3.5 Partner balance

One of the central goals[4] of the Gemini allocation and scheduling process is to ensure equitable access to the various observing modes and conditions amongst the partners. One metric that can be used to demonstrate this is the

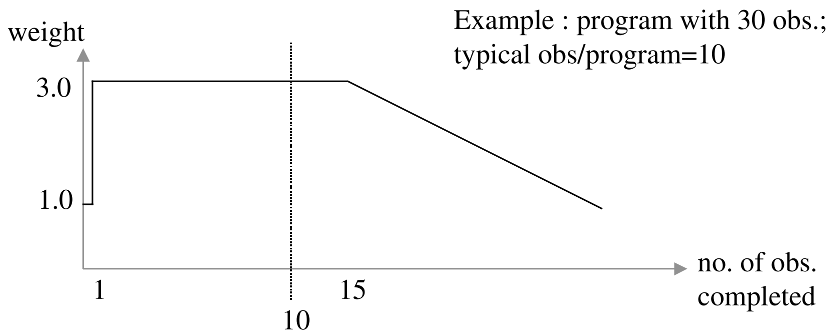

Figure 2: weight favouring completion of programs

historical record of time usage. Paper I described how time is to be charged to individual programs, the possible mechanisms for adjustment of the national shares and showed that the merging of programs was naturally quite well balanced. A finer level of adjustment may be achieved by applying a different weight to each partner based upon their time usage relative to their request or nominal allocation and the total usage amongst all partners.

Formally, let the time used by some partner *i* up to now in the semester be $QUsage_i$. Thus the fraction of time used relative to the other partners is simply

$$FractionUsed_i = \frac{QUsage_i}{\sum_i QUsage} \tag{1}$$

The benchmark that the time used is tested against is either the total allocation requested by that partner or their nominal share, whichever is the smaller. We can construct a benchmark fraction in the same way as for the time usage in eqn. (1) and then the raw partner weight is

$$RawWeight_i = \frac{BenchmarkFraction_i}{FractionUsed_i} \tag{2}$$

In practice, to avoid infinities if there has been no usage to date of time by a partner and to provide a gain control for defining the impact of the time usage feedback, we can re-cast eqn. (2) as

$$Weight_i = \left(\frac{BenchmarkFraction_i}{FractionUsed_i + \left(\frac{BenchmarkFraction_i}{G1} \right)} \right)^{G2} \tag{3}$$

where G1 and G2 are the control parameters. Weights as defined in eqn. (3) are constructed for category A- and B-time (see section 3.2) separately for each partner and compared between partners to arrive at a series of relative weights. In the simulations described in section 4 the parameters have values G1=G2=2 resulting in a typical range of the relative weights during execution of 1-3.

Even with the inclusion of this feedback mechanism, it is not expected that detailed equity between the partners will be achieved within one semester. In particular, for the category A time and for the smaller partners, the relatively small shares may be subject to significant statistical fluctuation (see section 4).

3.6 Overall weighting function policies

The weights described in this section are analogous to the suitability functions used to constrain scheduling of the Hubble Space Telescope (see [9] for the theoretical and conceptual foundations of this approach). In the Gemini observatory the multiplicative combination of weights forms a "policy" which defines the preferred emphasis of the overall scientific program. Several variant policies may be constructed by adjusting the relative weights and their results compared in terms of scientific return, fairness etc. when subjected to the external constraints of simulated or forecast weather conditions, partner time preferences, actual program target distributions and so forth.

The baseline policy adopted for the queue execution simulations was to emphasize the properties which the Gemini telescopes are designed to exploit and thus stress the match between requested and current conditions within individual scientific ranking bands. Under typical circumstances, the favourable combination of all of the other weighting factors was required to exceed the site condition matching. As an example, consider two observations: observation A the first observation in a program of high angular resolution near-IR imaging of modestly bright targets, and observation B, visible spectrophotometry through a wide slit of a target required to complete a program for a partner whose usage is well below its nominal share. The specific and combined weights for this example are given in Table 3 together with a summary of the weighting parameters. In this case observation A would be favoured.

Parameter		Weight or typical range	Observation A weights	Observation B weights
Scientific band		Infinite	Same band	
Image quality	20%-ile	25	25	
	50%-ile	4		
	Unconstrained	1		1
IR background	20%-ile	25		
	50%-ile	4		
	Unconstrained	1	1	1
Cloud cover	Photometric	2		2
	Non-photometric	1	1	
Moon	Dark	2		2
	Bright	1	1	
Declination		1.0 - 2.0	1.5	1.5
Long-term visibility		1.0 - 1.5	1.3	1.4
Program completion		1.0 - 3.0	1	3
Partner balance		1.0 - 2.0	1	1.5
Total weight			48.75	37.8

Table 3: typical weights for queue execution parameters used in the simulations and two example observations

4. QUEUE EXECUTION SIMULATIONS

To investigate the behaviour of the queue execution model described in section 3, a software simulation was constructed. To provide a flexible, graphical means of analysing the results the simulation was implemented within an Excel spreadsheet.

4.1 Input parameters

Simulating the output from the partner time allocation committees and, to some extent, any partner bias intrinsic to the distribution of target, instrument or observing condition requests, ranked lists of programs were constructed by the Gemini project scientist team and national project scientists. These were merged into a single queue as described in Paper I. Within each program, each observation was defined by a set of instrument, site condition, target co-ordinate and integration time constraints consistent with the scientific aims. Thus (i) a mid-infrared observation would typically request good IR background and photometric conditions (no clouds), but would be uncaring about the image quality (always diffraction limited) or lunar phase, whereas (ii) narrow-slit near-IR spectroscopy might request median visible-wavelength image quality (i.e. good image quality in the near-IR). For computational simplicity the integration times were defined in integer hours. The final queue comprised 68 programs with a total of 412 observations requesting 906 hours of telescope time.

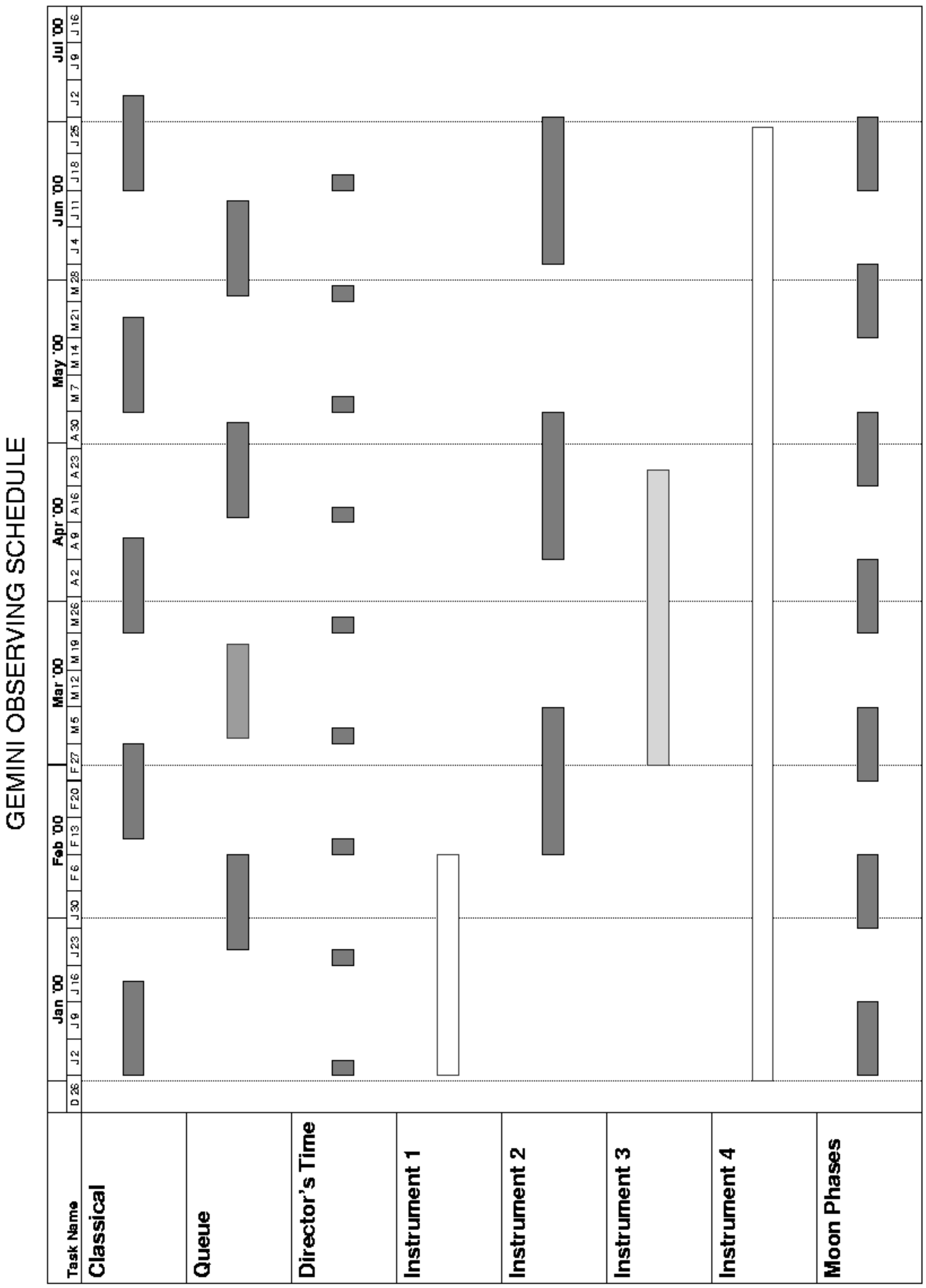

Figure 3: Example blocked observing schedule defining classical, queue and engineering time and (in this instance, limited) instrument availability

A blocked observing schedule (see Fig. 3), which would in practice be issued ahead of the call for proposals, defines the specific nights available to queue observing, instrument availability and lunar phase. For simplicity it was assumed that nights were entirely bright or dark. No daytime observation blocks were scheduled.

Each of the 79 available queue nights was assumed to be of ten hours duration. The observing conditions of image quality, IR background and cloud cover were generated according to random distributions. These distributions were modifiable so that, for example, the 20%-ile image quality conditions could occur more or less frequently than 20% of the time, reflecting the natural variation in their occurrence (see section 4.2 for further details). Note that this is the image quality defined at visible wavelengths, thus mid-IR observations were executed under any seeing conditions, near-IR observations requesting 50%-ile conditions were assumed executable under 100%-ile visible image quality; likewise, the majority of 20-%ile near-IR observations were considered executable under median conditions. The image quality was assumed constant on a 2-hour timescale but to be uncorrelated over longer periods.

The IR background distribution was similar to the image quality except that it was assumed to be constant for an entire night, reflecting the longer timescale of the associated atmospheric phenomena. Cloud cover was assumed to produce constant nightly conditions that were photometric or non-photometric, usually in equal proportion, or unusable. Long-term averages for Mauna Kea were used to set the nominal cloud distribution fractions at 40%:40%:20%.

Not implemented in the simulation was an efficiency weighting that might, for example, favour no instrument change or target proximity. However we note that given the capability for daytime calibrations with the facility calibration unit, the simultaneous mounting of three science instruments and the stability of the acquisition and guidance unit, the overhead for instrument changes is expected to be small.

4.2 Example results

A simplified flow-chart of the execution simulation is shown in Fig. 4.

- For each queue night in semester
  - For each hour
    - Determine which programs could be executed under current conditions of image quality, IR background, sky and moon, and instrument availability
    - For each band of programs of equal scientific ranking
      - Calculate weighting factor for all observations that match conditions
      - Execute observation with highest weighting factor
    - If no observations match, search next scientific ranking band
  - Next hour
- Next night

Figure 4: queue execution simulation flow-chart

One of the first results from the simulation was that the influence of the specific distribution of site conditions (e.g. whether a particular night was cloudy or not) on the overall time usage and partner shares was minimal (provided that the distribution function itself was maintained). The 1s variation in the usage of desirable time was less than a few percent, presumably due to the large number of queue hours executed and feedback via the partner weights. This meant that a large number of runs in a Monte-Carlo simulation were not necessary to derive meaningful results from adjustments to individual parameters.

Prior to augmentation of the queue by extra programs, significant periods of unused time appeared in the semester summary. These arose principally from an inadequate number of observations capable of being executed under the poorer conditions and because the distribution of target RAs was significantly non-uniform. This serves to highlight two features. Firstly, analysis of the requested distribution of site conditions and targets, or a simulation of the queue execution for the entire semester, is needed early in the time allocation process i.e. at the partner TAC meetings and after queue merging. Adjustments to the recommended program content may result. Secondly, there is an important trade-off in the degree to which the queue must be overloaded between ensuring an adequate pool of programs without requiring applicants to prepare observations that might only have a small likelihood of being executed in the semester. Clearly the natural variation in site conditions and the effectiveness of the queue scheduler each play a major role in defining the magnitude of the overloading required. By varying the number of unusable nights, this simulation, which does not search for optimum solutions by planning ahead, implies an overallocation of about 20% to ensure that all of the available time is used (for a fixed distribution of conditions). It would be of interest to explore to what extent an optimal scheduler could reduce this figure.

The variation in site conditions might naively be expected to be large but averaged over an entire semester this appears not to be the case. In the absence of adequate historical data at optical or near-IR wavelengths, Fig 5. shows the sub-mm seeing cumulative frequency distributions at Mauna Kea for the eight quarters in 1994/95. The variation in the frequency of occurrence of median (50%-ile) and best (20%-ile) conditions is only about 25% between quarters. Given that the sub-mm seeing distribution is broader than that for optical seeing, and that the Gemini telescopes will use an articulated secondary mirror to remove the tip and tilt components of image motion, this result may be considered an upper limit. Since the two terms contributing to queue overloading are not independent, their combined effect might therefore be expected to be in the range 25-40%.

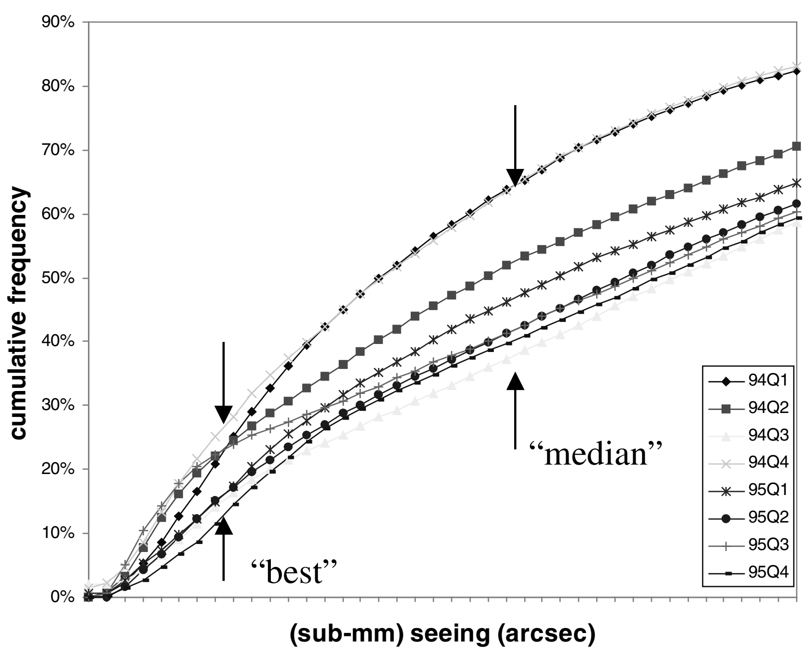

Figure 5: sub-mm seeing cumulative frequency distributions

A summary of the completion status of individual observations and entire programs in the queue after a typical run of the simulation is given in Fig. 6. In this instance, three unequal scientific ranking bands were employed and are indicated by the vertical arrows. The status of each program is indicated in the fourth column. Those which were completed (i.e. all observations executed) are highlighted (dark background). Those programs that were started (only some observations completed) and those not attempted are also indicated. The status of each observation is shown as "done" or by the integration time required.

It can be seen that the scientific bands are effective in ensuring that the most scientifically important programs are carried out. All of the programs in the top-ranking band were completed as were most in the second band. In this example, 75% of all programs were started. The weighting factor which favours completing already started programs is also effective as 84% of started programs were finished.

The observation execution may also be analysed in terms of the time usage by partners (see Table 4). Contrasting the use of the category A (most desirable) and B (desirable) times with the expected partner shares indicates that an adequate, but not precise, balance is achieved within one semester. The result from several runs is that typically the partner usage is within 10-15% of the nominal allocation.

Partner	A - time	B - time	Total time	Nominal goal
US	49%	50%	53%	47%
UK	22%	19%	18%	21%
Ca	15%	15%	14%	16%
UH	0%	0%	0%	none requested
Ch	6%	7%	6%	8%
Ar	3%	5%	4%	4%
Br	5%	4%	5%	4%

Table 4: example partner distributions of queue usage analysed by site conditions

5 IMPLEMENTATION AND SOFTWARE TOOLS

The queue execution simulation described in this paper has highlighted the need for software scheduling tools and analysis at several stages in the time allocation and execution process. The partner TACs must ensure that an appropriate distribution of requested observing conditions and targets in their ranked lists of programs. After electronic transmission of the recommended programs to Gemini and abstraction into the observing database, merging of the lists into a draft queue occurs. Simulation of the entire semester uses as input the draft queue and a statistical expectation of the site conditions. Any conflicts or holes in the overall program thus identified will need to be resolved at the Gemini international time allocation committee (ITAC) meeting, most likely by the replacement of programs. As the semester progresses, more detailed short-term (e.g. weekly and nightly) simulations using statistical distributions, weather forecasts or nowcasts will be employed to identify likely programs to be executed, for the purposes of staff and observatory planning. During the night, weather forecasts and site monitoring equipment will provide the constraints against which observations in the active database are tested. It is anticipated that all of the simulation activities that act on programs in the Gemini database can employ the same scheduling tool.

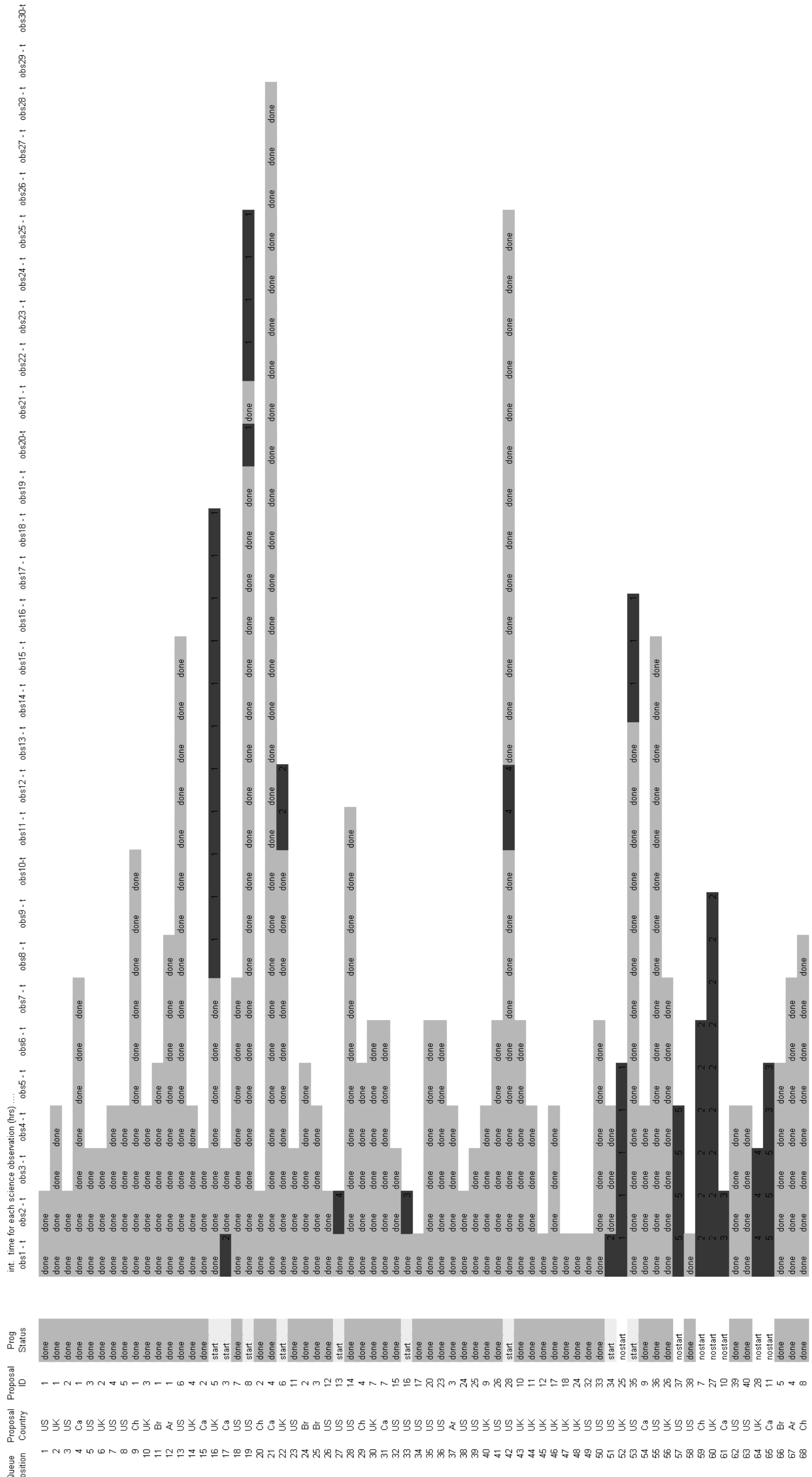

Figure 6: example program execution results

ACKNOWLEDGEMENTS

We thank the Gemini Science Operations Working Group, especially Todd Boroson, John Davies and Bruce Gillespie for considered input during the development of the Gemini queue execution model. Thanks also to Remo Tilanus for extracting the CSO sub-mm seeing data from the JCMT archive.

The Gemini 8-m Telescopes Project is managed by the Association of Universities for Research in Astronomy, for the National Science Foundation and the Gemini Board, under an international partnership agreement.

REFERENCES

[1] Gemini Science Requirements 3.0, 1996.
[2] Leibundgut, B., 1997. SPIE, 2871, 755.
[3] Peron, M. & Grosbol, P., 1997. SPIE, 2871, 762.
[4] Puxley, P. J. & Boroson, T., 1997. SPIE, 2871, 744.
[5] Walker, S., Gillies, K., Puxley, P.J. & Wampler, S. These proceedings.
[6] Simons, D. & Jensen, J., 1997. Gemini Technical Note TN-PS-G0050.
[7] Boroson, T., Davies, J. & Robson, I., 1996. "New Observing Modes for the Next Century", PASP, 87.
[8] Silva, D. & De Young, D., 1996. NOAO Newsletter No. 45.
[9] Johnston, M. D. & Miller, G. E., 1994. In "Intelligent Scheduling", ed. M. Fox & M. Zweben, (San Francisco: Morgan-Kaufmann), 391.
[10] Puxley, P.J., 1997. "Use of Gemini for Short, Quick Response or Time Critical Observations", Gemini Technical Note TN-PS-G0051.