

PC4CAST—A Tool for DSN Load Forecasting and Capacity Planning

S. J. Loyola
Systems Analysis Section

Effectively planning the use and evolution of the DSN is a complex problem involving many parameters. This article discusses the tool that models many of these complexities, yet requires simple structured inputs and provides concise easy-to-understand metrics to aid in the planning process. The tool, PC4CAST, is used for both load forecasting (predicting how well planned that DSN resources meet expected demand) and as a decision support tool in the capacity-planning process (determining the relative benefits of capacity expansion options). It is now in use in the TDA Planning Office, has been used in numerous studies, and is also being used by the JPL Multimission Operations System Office (MOSO) as an integral part of Resource Allocation Team activities. Experience using the tool has helped to identify additional requirements that will further improve the planning process, which can be met by future PC4CAST versions.

I. Introduction

The DSN is a set of resources in high demand that requires careful planning to ensure its success in providing telecommunications and radio astronomy services to its community of users. Design and construction of additional DSN antennas take considerable time and money. For the DSN to remain responsive to its users, future resource capacity and capability (e.g., transmitter/receiver frequency) to satisfy expected user demand must be determined well in advance. When there is mismatch between resources and user requests, analysis is required to determine what additional resources are required and/or how the user demand can be modified [1]. Because of funding and schedule constraints, long-range planning typically

involves a cost-benefit comparison of numerous capacity-expansion options.

Analysis to support this planning can be visualized along a spectrum. At one end is the amount of user demand that can be satisfied by a static view of DSN resources (i.e., the current set of antennas and any planned additions, upgrades, and removals). This approach is most applicable for evaluating the fit between the DSN evolution plan and the current set of mission requirements. At the other end is fixed user demand and the resource set (i.e., number, type, location, etc.) necessary to satisfy it, excluding issues of available funding and construction schedule constraints. The systems view, at the intersection of

these two views, considers both ends of the spectrum and provides trade-offs between user support and resource augmentation. For cost-effective planning of TDA resources, a planning tool must provide analysis appropriate to the problem.

The PC4CAST tool supports this spectrum of planning analysis in an integrated, consistent, accurate, and timely manner. The structure of analytical inputs (e.g., user requests, resource availability, and geometric constraints between user target and resources) is consistent for all types of analyses, while the type of output metrics may be chosen to fit the analysis being performed. The tool statistically models how user requirements would be accommodated without having to generate detailed schedules. PC4CAST is used both to forecast the DSN's ability to support any given mission set and to identify the DSN resources required for varying levels of mission support. A major capability of the tool is analysis of the marginal impact of new missions and new capacity.

Different approaches are required for planning the 34-m/70-m ground resources (which primarily support deep-space missions) than those required for 26-m and smaller resources (which typically support Earth-orbiting missions). The operational version of PC4CAST is designed to facilitate planning of the 34-m/70-m resources. A prototype PC4CAST is currently under development, which will model stochastic demand for 26-m and smaller resources. The balance of this article will focus on the operational capabilities of, and products from, PC4CAST for 34-m/70-m ground resources.

II. System Architecture

The PC4CAST tool consists of a spreadsheet-based user interface and a forecasting engine (see Fig. 1). Microsoft Excel version 4.0 is used for both input of problem-dependent parameters and graphical display of forecasting metrics. A forecasting engine, written in the C++ language, uses information entered into Excel, along with problem-independent information stored in databases, and provides the raw forecasting results. The Microsoft Windows environment seamlessly integrates Excel and the forecasting engine, so the separation of input/output functions from processing is invisible to the user and does not affect tool operation. This architecture takes advantage of the capabilities of standard commercial off-the-shelf products, allowing more development effort to be spent on design of the forecasting and capacity/capability planning algorithms.

Inputs and outputs are maintained in a type of Excel file called a workbook (a file which may contain multiple spreadsheets and charts) [2]. Each forecasting workbook represents one calendar year of a study. It contains spreadsheets for forecasting inputs, and both spreadsheets and charts for forecasting outputs. Inputs are entered by week, while results are output by week and month. The spreadsheet interface allows output formats to easily evolve with demand for different ways to view forecasting results. In addition, Excel users are able to design their own output formats to suit the needs of a particular study.

III. Inputs

Inputs to the PC4CAST tool can be grouped into three categories: user requirements, resource-capacity description, and view periods. User requirements are input in the same format that the JPL Multimission Operations Systems Office (MOSO) gathers and maintains them. This format is concise, yet of sufficient detail to allow accurate modeling of the user demand. Resource-capacity descriptions include each antenna's planned availability and any options for future capacity changes. View periods (times when a particular object is in view at a particular antenna) represent the intersection of the user and ground resource domains. The tool requires all mission view periods (typically one view period per station per day for deep-space objects) for the time period to be forecast.

User requirements are entered and stored in a spreadsheet within a forecasting workbook (Fig. 2). Requirements for all users of the resources must be input including those from spacecraft, ground-based scientific users such as Goldstone Solar System Radar (GSSR), High Resolution Microwave Survey (HRMS), and Radio Astronomy and Special Activities (RASA), and those for the DSN's own use (e.g., antenna calibration, maintenance). User requirements include the number of tracks or resource usages required each week and the following parameters: view period object (e.g., Cassini), user description (e.g., GSSR Mercury), usable resources, average and minimum durations for tracks in each week of this requirement, and pre- and post-calibration time, which includes antenna setup and tear-down overhead before and after each track. The set of resources that are able to satisfy the requirement is represented by the usable-resources parameter, described below.

The PC4CAST tool allows a requirement's usable resources to be stated in a variety of ways. Usable resources may be specified by antenna (e.g., DSS 14 or 14 for the 70-m antenna at Goldstone) or subnet (e.g., 70M, 34S,

34H). Antennas and/or subnets may be combined to represent antenna arrays, the simultaneous use of two or more antennas, or resource equivalents, when two or more resources can satisfy a requirement. For example, "70M, 34S/34H" represents a requirement that can use either a 70-m antenna or an array of 34S and 34H antennas.

Available resources are described by resource definitions, scenarios (evolution paths), and downtimes. Resource definitions simply define each resource by its code, location (i.e., which Deep Space Communications Complex), and subnet. This tool also provides network augmentation studies by maintaining multiple resource scenarios, each including the initial on-line and final off-line dates for each current and planned resource. The PC4CAST operator chooses the appropriate scenario based on the study to be performed. Resource downtimes represent extended periods of time when a particular resource is taken down. Regular antenna maintenance is treated as a user requirement, since this best models the scheduling environment.

IV. Outputs

The PC4CAST system has several different output metrics, each providing a different slice through detailed antenna-usage information. The calculation of *expected-usage profiles* is discussed in detail in Section V, but a quick description is necessary here. The software uses the input user requirements, resource scenario, and view periods to calculate a detailed view of the expected demand on each antenna throughout the study time period. Figure 3 is an example of the expected demand on one antenna for one week. The expected-usage profile provides the raw data from which many concise, useful metrics may be derived.

One primary metric used in forecasting studies is the amount of user community requirements that cannot be satisfied given the resource capacity defined by the input resource scenario and downtimes. In PC4CAST, this unserved user demand is labeled *lost time*. It is calculated as the hours that cannot be supported, and is usually output as a percentage of the total hours of user requirements on a per-subnet basis (Fig. 4). This is a convenient level of aggregation since each subnet usually has its own community of users. Lost-time information can be presented at varying levels of aggregation (e.g., the whole network) as required by the study being performed. Besides showing how the current resource implementation plan can support current user requirements, lost-time information may be used to compare the relative merits between two resource

scenarios or the impact of adding, removing, or modifying a user's requirements.

Whereas the lost-time metric is from the users' point of view, the *load-duration curve* presents user demand from the ground resources' perspective. This metric has been successfully used in the electric power utilities industry for management decision making on cost-effective capacity evolution [3]. It shows the percentage of time that a resource can expect various levels of demand (load). The load-duration curve for the 70-m subnet for 1995 (Fig. 5) reveals that demand exceeds capacity for half of the year. Moreover, there is demand exceeding 200 percent of capacity for roughly 8 percent of the year (about one month).

The load-duration curve displays the full range of user demand relevant to capacity planning over the time period of interest. The *y*-axis intercept shows the value of peak expected usage, and the shape of the curve displays the proportion of time that demand is above, at, or below capacity. When the user load is above capacity (greater than 1.0), opportunities exist for adding capacity from DSN and non-DSN sources and/or managing user loads. When user loads are under 1.0, the ground system has the potential for accommodating other missions, if there is a favorable view period, or non-view period-constrained missions such as HRMS. Opportunities for, and benefits of, load management can also be quantified. Postponing a user's period of high demand so that it does not coincide with another user's can have a leveling effect by shifting load from the left side of the curve to the right. Reducing a user's required minimum track duration can have a similar effect on the shape of the curve. Reducing pre- and post-calibration times will usually reduce the height of the curve at all points. The impact of these and other assumptions, such as resource scenarios and mission sets, may be examined by comparing the shape of a load-duration curve generated with these inputs to a baseline curve.

Another useful metric, which is shaped somewhat like the load-duration curve, is the *resources-versus-lost-time graph*. This graph shows how the percentage of unsatisfiable requests varies with changing resource capacity. Conversely, it can be viewed as showing what resource capacity is necessary to satisfy a certain percentage of user requirements. This metric should be used with some caution relative to nonintegral subnet values since the location of antenna sites impacts the ability to satisfy user loads. For example, Fig. 6 suggests that there would be 31 percent lost time if only one subnet were available and 3 percent if there were two 70-m subnets. But, the implication that 1-2/3 subnets (i.e., 5 antennas) would result in 8 percent lost time is highly dependent upon antenna location and

the declinations of the spacecraft involved. In cases like this, the metric provides direction for further analysis of additional antenna-placement strategies.

The *trade space graph* displays the information contained in multiple resources-versus-lost-time graphs as the number of resources required for a few selected lost-time levels over multiple years. Figure 7 shows the number of 34-m subnets that is required to provide four different levels of lost time (0, 5, 10, and 20 percent). Since the absolute lost-time value is dependent on many factors, including how much filtering has been performed on the user requirements and how many missions are in prime phase, no one level of lost time can always be considered acceptable. Nevertheless, this graph does help to define the trade space where the cost of additional resources must be weighed against the cost of losing scientific data due to unsatisfied user requirements.

The PC4CAST system's library of metrics can be used to analyze problems in both load forecasting and capacity planning. The current status of user support can be measured with the lost-time metric. Load-duration curves can be used to assess the marginal benefits of adding resource capacity and/or capability. Cost-effective capacity evolution paths may be identified using resource-versus-lost-time and trade space graphs. Together, these metrics provide an internally consistent package to aid DSN management decision making.

V. Algorithms

The PC4CAST system uses a statistical approach that provides quick run time and models how user requirements would be scheduled, without actually generating schedules. Figure 8 is a process flow chart that gives a graphic presentation of the PC4CAST forecasting engine algorithm. This algorithm can be divided into two main phases. The first is the calculation of expected-usage profiles from forecasting inputs (user requirements, resource-capacity descriptions, and view periods). The second is the derivation of higher level metrics from expected-usage information. Presently, each week of a study year is calculated separately. Forecasting metrics for one year of inputs are calculated in around five minutes, running on a 486-based computer. The system design allows for relatively easy modification to other time granularity, should the need arise in the future.

A. Calculation of Expected-Usage Profiles

The calculation of expected-usage profiles is a three-stage process that incorporates user requirements, resource

capacity, and view period information. As stated earlier, the goal of this process is to determine the expected level of demand for all time periods within the study time frame. The first stage is to calculate an *expected-usage value* for each requirement. This value represents a time-weighted distribution of the requirements over the view period length. Second, this expected usage value is used with the view periods and the required tracking overhead to generate individual expected-usage profiles. Each of these profiles represents the demand from that requirement for each point in time for each antenna. Third, all individual expected-usage profiles for each antenna are summed, resulting in one expected-usage profile for each antenna. These expected-usage profiles are the product of the first phase of the algorithm. Each of these three stages will be now discussed in detail.

The following describes the calculation of the expected-usage value for one requirement in one week; each requirement and each week are processed in a similar manner. First, the total amount of requested time is calculated from the requirement's average duration, pre- and post-calibration times, and the number of tracks requested in the week, as follows:

$$\begin{aligned} \text{requested time} &= \text{number of tracks} \\ &\quad \times (\text{average duration} + \text{precal.} + \text{postcal.}) \end{aligned}$$

Calibration times are included because they represent demand on the resources just as the actual track does. The next step is to find all *usable view periods* for the resources specified in the usable resource field that are long enough to support this requirement. In other words, the individual view period duration must not be less than the requirement's minimum duration. Then, *request slots* are defined as usable view periods with pre- and post-calibration appended to them. Request slots are defined this way because the time before and after each view period represents time when the calibrations could take place and still have the track within the view period. The *request slot time* is then calculated as the sum of all request slot durations.

$$\text{request slot time} = \sum_{\text{all request slots}} \text{duration (request slot)}$$

Finally, the expected-usage value is calculated as

$$\text{expected usage} = \frac{\text{requested time}}{\text{request slot time}}$$

The expected-usage value therefore represents the percentage expectation that when the requirement can use an antenna (constrained by the minimum duration constraint), it would use it or be scheduled to use it.

If the expected usage is greater than 1.0, the requirement cannot be supported by the resources it specified. This is caused by physical constraints such as positions of antennas and users, required view period mask, and overly restrictive requirements; it is not due to competition with other users. The system informs the analyst of any *infeasible requests* at the end of the attempted run. It is then the analyst's responsibility to modify the requirement until it is feasible. This can be accomplished by one or more of the following means: adding to the usable resources field; or reducing the minimum duration, average duration, precalibration, postcalibration, or number of tracks. Forecasting metrics will not be generated until all requirements are feasible.

The next stage is to generate individual expected-usage profiles for each requirement on each antenna requested. These are defined as stepwise constant functions, which can be visualized as the expected-usage value assigned to each request slot.

The final stage involves the creation of an expected-usage profile for each antenna from all individual expected-usage profiles for that antenna. Each individual profile is summed to a total expected demand for each antenna over the study period. Figure 9 graphically shows the generation of an expected-usage profile for a hypothetical mission set. This sample profile is used in the next section to help describe how PC4CAST high-level metrics are derived from expected-usage profiles.

B. High-Level Metric Derivation

Once the expected-usage profiles are calculated for all requested antennas, many high-level metrics may be derived from them. The derivation of the most frequently used metrics—lost time, load-duration curves, resource-versus-lost-time graphs, and the trade space graph—are discussed below.

Within PC4CAST, lost time is defined as the area of the expected-usage profile that is above one expected-usage unit (Fig. 10). Since the antennas that are being modeled can support only one user at a time, any expected usage over 1.0 represents time that cannot be supported. This is insupportable user demand or lost time. While lost time is calculated by antenna, the values for all antennas in a subnet are usually summed to give the lost time for the subnet. This subnet lost time is then charted as a percentage

of the requested time on that subnet. Lost time is calculated on a per-week basis, but may be aggregated to any longer time periods, such as monthly or quarterly. Parenthetically, it should be noted that subnet requested time is simply the sum of the areas of expected-usage profiles for each antenna in the subnet. The lost-time metric captures the essential information contained in the expected-usage profiles and provides a concise measure of the impact of the input resource scenario upon user requirements.

Load-duration curves include all expected usage-level information, but aggregate the time-of-day information to a level more useful for capacity planning. An easy way to visualize a load-duration curve is as an expected-usage profile with the highest peaks sorted to the left and the lowest usage levels to the right (Fig. 11). The load-duration curve is generated for each antenna, but may be aggregated to the subnet level.

The resources-versus-lost-time graph displays lost time for a range of hypothetical resource capacities. The lost-time metric is usually calculated based on the fact that an antenna can support only one user simultaneously. But, for this graph, it is calculated with an assumed resource capacity ranging from zero to the highest expected-usage level (Fig. 12). As with the load-duration curve, these graphs can be used at the antenna or subnet level. This results in a graph that contains some immediately obvious data points. For instance, a resource capacity of zero results in 100 percent lost time; likewise, capacity equal to the highest expected usage results in no lost time (but very low capacity utilization). Also, the lost-time value for a capacity of one resource corresponds to the usual lost-time metric. The value provided by the resources-versus-lost-time graph is the identification of how addition or removal of ground resources will impact the users.

The trade space graph (Fig. 7) is derived from the resources-versus-lost-time graph for a few selected lost-time values. Each line on the graph corresponds to the resources required to provide for a lost time less than or equal to the selected level. The required-resources value is calculated by reading, on the resources-versus-lost-time graph, the number of subnets required for the selected lost-time level. This value is then rounded up to the nearest whole antenna (e.g., 1.5 subnets becomes 1-2/3 subnets, or 5 antennas).

VI. Future Potential

To date, PC4CAST has proven to be helpful in load forecasting and capacity-planning analysis. Experience

with the tool has identified requirements for further algorithmic enhancements which would provide improved modeling of the DSN system, more accurate metrics, greater insight into the distribution of lost time, and measures of augmentation cost effectiveness. The definition of a user requirement's usable resources could be extended to include required antenna size and additional equipment parameters such as transmitter and receiver frequencies. Lost time would be redefined to include not only requested time that is unsupported because of competition, but also requests made infeasible by an antenna being down. Also, definition and implementation of a priority scheme would allow the lost-time and load-duration curve metrics to be disaggregated to the level of prioritized mission sets or even individual users. Resource augmentation costs and the value of user support would also be included to better support the decision-making process. These features

would provide for improved analysis of the marginal benefits of resource evolution paths.

Earth-orbiting missions have not been modeled in this implementation of PC4CAST. Accurate calculation of the required view period knowledge for all users and time frames of interest is usually not a problem for the deep-space missions that normally request time on the 70- and 34-m antennas. The Earth orbiters that typically request the 26-m and smaller antennas have far less predictable view periods. The Systems Analysis Section has been developing probabilistic methods to model such users where view periods cannot be calculated very far into the future. This work has followed the general PC4CAST paradigm so the two approaches could be readily integrated to provide a comprehensive DSN load-forecasting and capacity-planning tool.

Acknowledgments

The initial PC4CAST prototype was funded by MOSO. The author thanks Gene Burke, Robertson Stevens, and J. R. Hall for providing PC4CAST metric development motivation and feedback at various stages; Fred McLaughlin for years of nearly continuous operational testing and feedback; David Werntz for his efforts in the development of the PC4CAST algorithms; Sil Zendejas for design and implementation of the underlying database structures; and Chet Borden for review of this project and the notion to apply load-duration curves to this domain.

References

- [1] E. C. Posner, "Expected Antenna Utilization and Overload," *The Telecommunications and Data Acquisition Progress Report 42-107*, vol. July-September 1991, Jet Propulsion Laboratory, Pasadena, California, pp. 104-112, November 15, 1991.
- [2] *Microsoft Excel User's Guide (Version 4.0)*, Microsoft Corporation, Redmond, Washington, 1992.
- [3] Decision Focus, Incorporated, *Costs and Benefits of Over/Under Capacity in Electric Power System Planning*, Report no. EA-927, Electric Power Research Institute Palo Alto, California, October 1978.

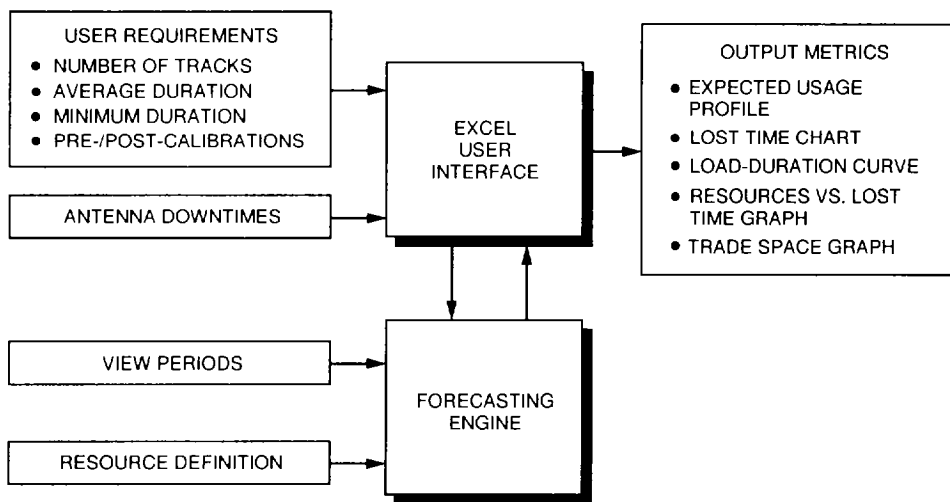


Fig. 1. Overview of the PC4CAST forecasting system.

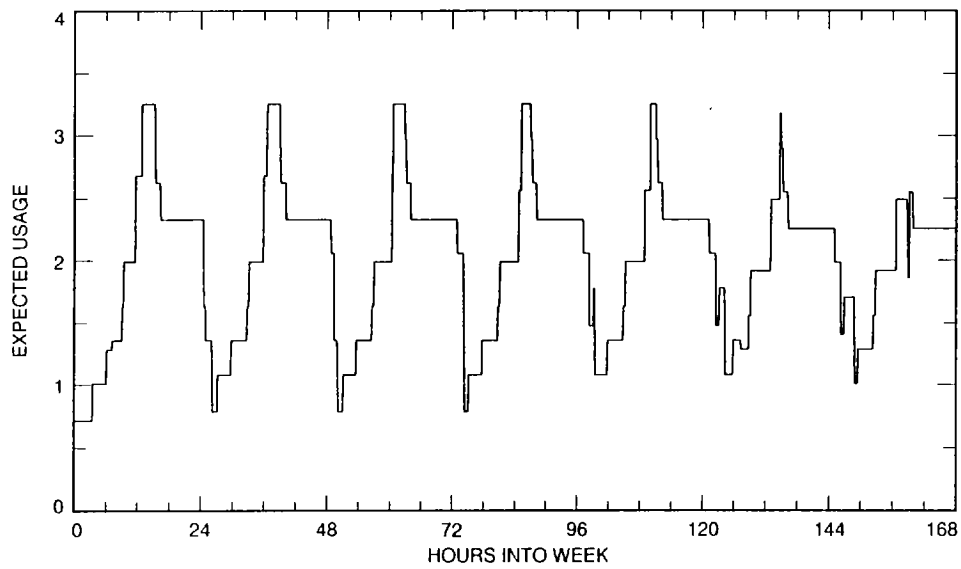


Fig. 3. DSS 12 expected-usage profile for week 18 in 1995.

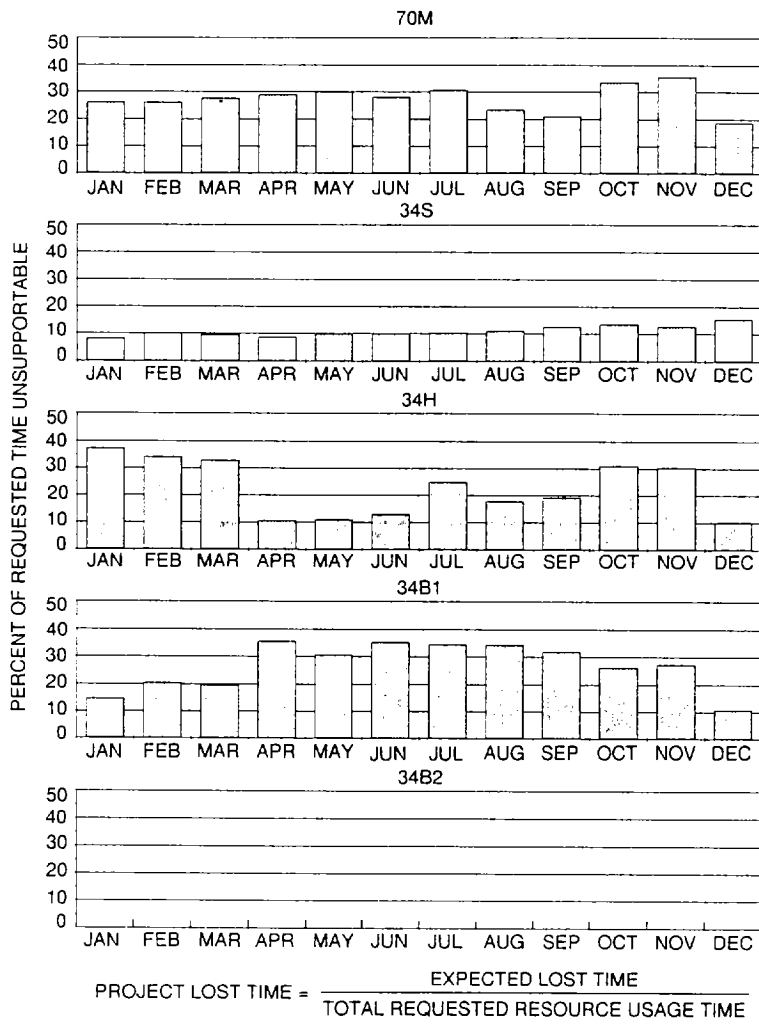


Fig. 4. Monthly lost-time chart for 1997.

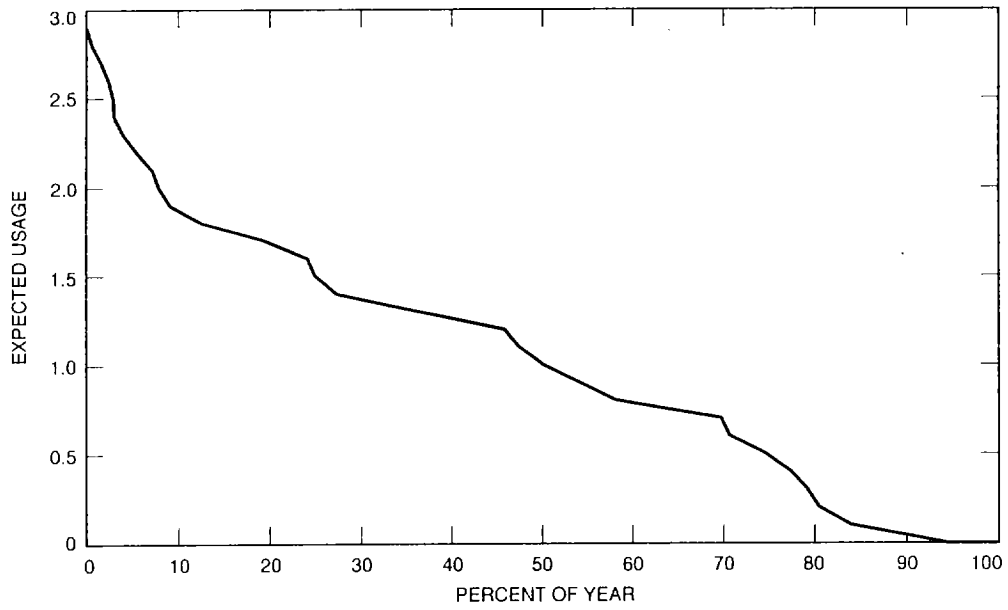


Fig. 5. Load-duration curve for the 70-m subnet in 1995.

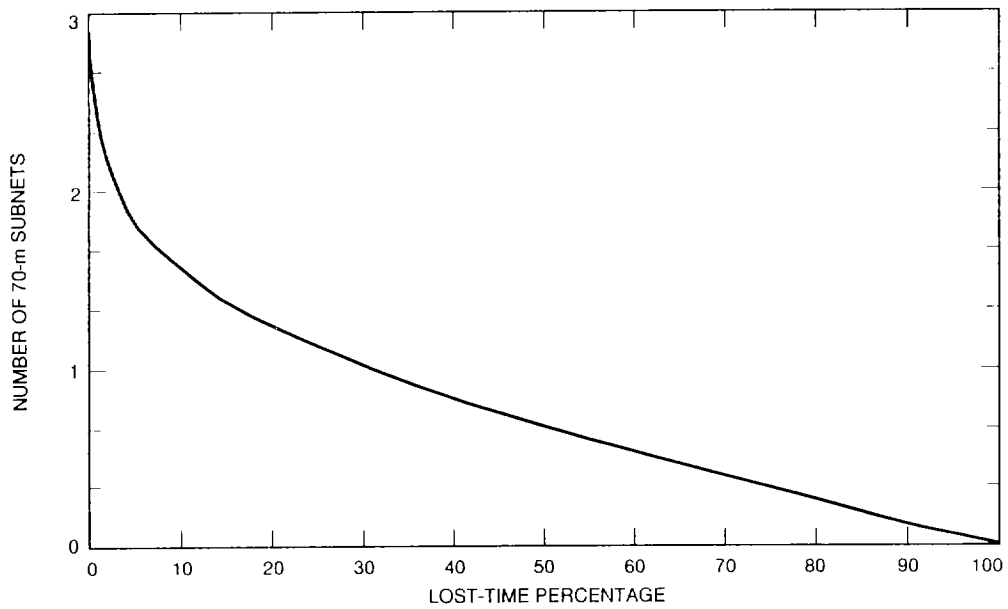


Fig. 6. Resources-versus-lost-time graph for the 70-m subnet in 1995.

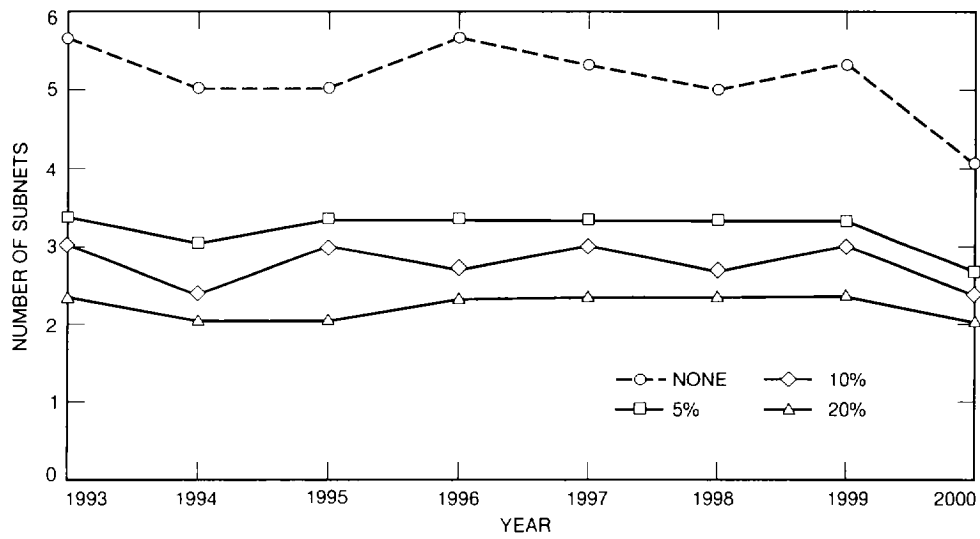


Fig. 7. Trade space graph for 34-m subnets for 1993 to 2000.

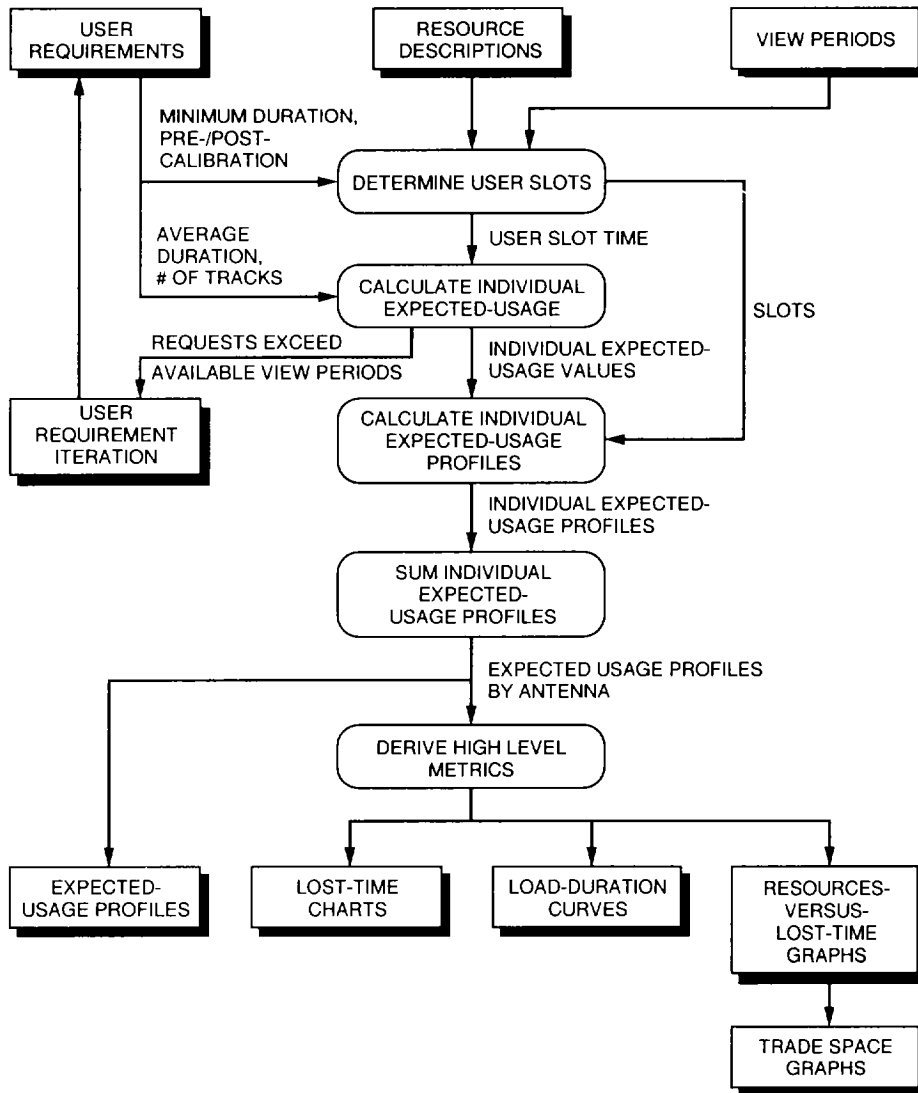


Fig. 8. PC4CAST forecasting engine process flow chart.

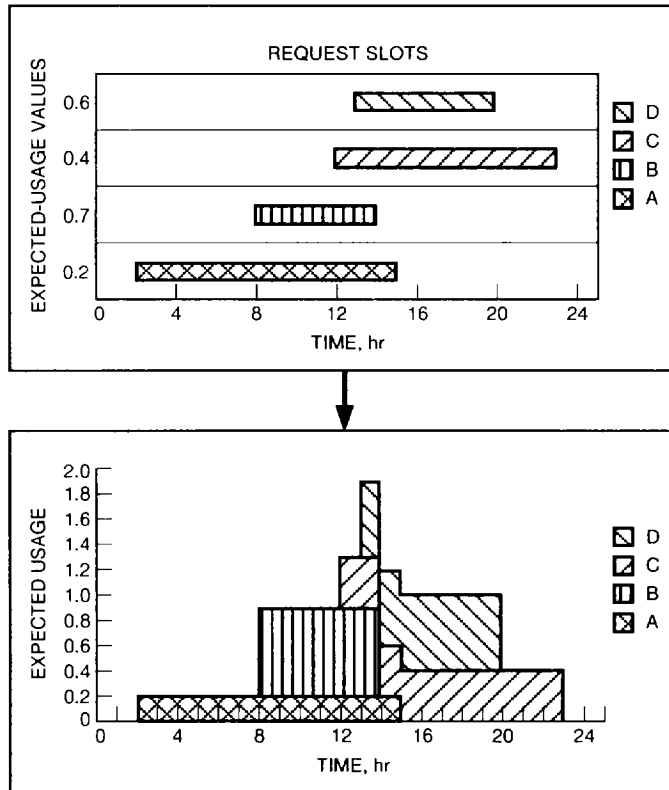


Fig. 9. Generation of an expected-usage profile using expected usage values and request slots for a hypothetical mission set (A, B, C, and D).

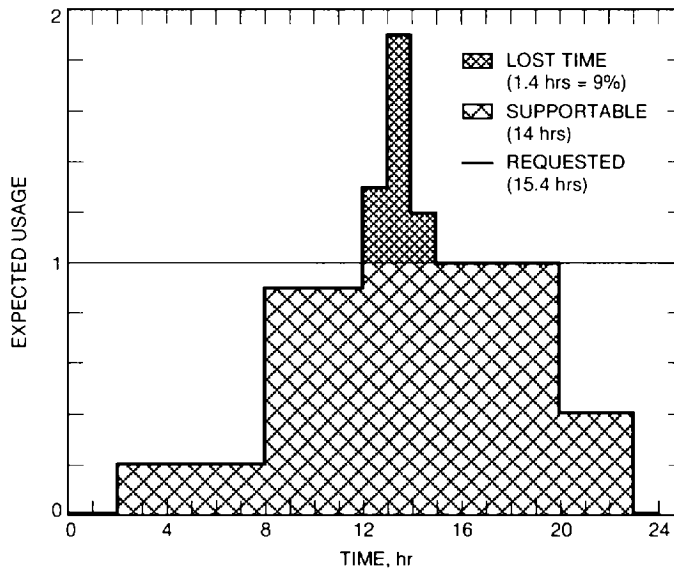


Fig. 10. Calculation of lost time from a hypothetical expected-usage profile.

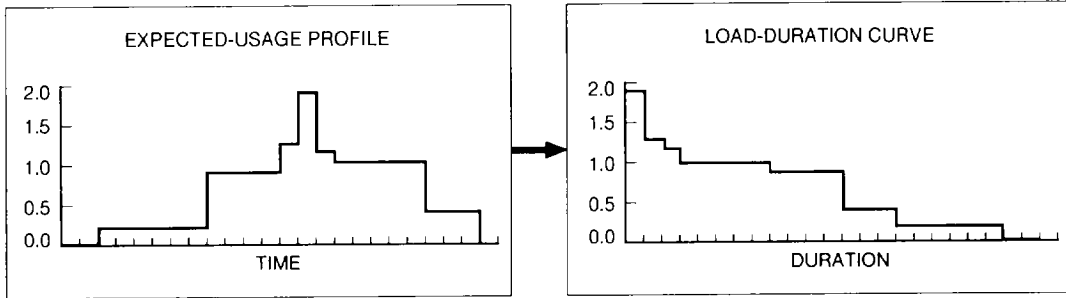


Fig. 11. Derivation of a load-duration curve from a hypothetical expected-usage profile.

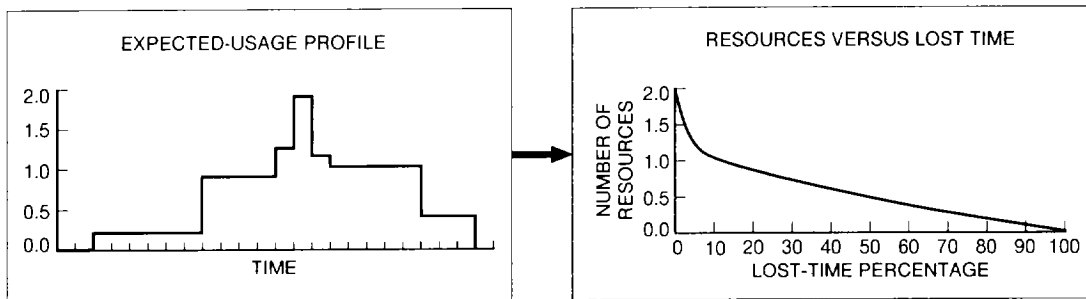


Fig. 12. Derivation of a resources-versus-lost time chart from a hypothetical expected-usage profile.