Proceedings of the 1984
Winter Simulation Conference
S. Sheppard, U. Pooch, D. Pegden (eds.)

B-1B AVIONICS/AUTOMATIC TEST EQUIPMENT: MAINTENANCE QUEUEING ANALYSIS

Lance M. Roark
HQ DLA-LO (DORO)
Defense General Supply Center
Richmond, VA 23297

James K. Feldman
School of Engineering
Air Force Institute of Technology
Wright Patterson AFB, OH 45433

James Bexfield
HQ USAF/SASB
The Pentagon
Washington, DC 20330

This paper describes a research effort on a queueing analysis of B-1B avionics maintenance. In this analysis, the queueing "customers" are the avionics components of the B-1B bomber, and the queueing "servers" are automatic test equipment stations. The purpose of the research effort was to develop a technique to determine B-1B test station quantities required to support B-1B avionics maintenance at base level. A detailed and complex simulation model was developed in the Q-GERT simulation language. The variance reduction technique of common random numbers was used in conjunction with a randomized complete block design for analysis of the simulation model outputs. In addition, considerable sensitivity analysis was performed with the simulation model. Finally, the paper describes two alternative techniques to determine test station quantities based on model output.

BACKGROUND AND PROBLEM STATEMENT

Background

Modern USAF aircraft have complex avionics components which require frequent unscheduled maintenance. As part of their maintenance concept, USAF aircraft are delivered to their bases with complex and expensive automatic text equipment (ATE). Avionics components, called line replaceable units (LRUs), are removed from the aircraft when a malfunction is detected and are taken to a nearby repair shop where the ATE is located. Maintenance technicians then use the ATE for fault detection, isolation, and ultimately repair of the avionics LRU. The ATE consists of specialized test stations. Each test station is devoted to a different grouping of the avionics LRUs. For example, one test station might be used for the repair of radio frequency (RF) LRUs, while another test station might be used for the repair of digital computer LRUs. Since there are many aircraft at each base, and since each aircraft has many complex LRUs, there may be numerous LRUs sent to the repair shop each day. This may generate a significant workload on each of the test stations.

The workload of avionics on ATE test stations is of particular interest to the B-1B System Program Office (SPO). The ATE which will be deployed with the B-1B bomber consists of four specialized test station types which will be used to support over 100 avionics LRUs. Unless the B-1B SPO has the ability to quantify and predict this workload, the SPO will be uncertain as to the best quantity of test stations of each type to procure for each B-1B base. Not buying sufficient test stations would degrade B-1B avionics readiness; buying too many test stations would be needlessly expensive. A technique to estimate the avionics workload on the ATE would not only be useful for decision-making about the best quantity of test stations to procure, but could also be used to justify ATE funding requirements. Due to the tremendous lead times involved in the procurement process, such decision-making must take place very early in the life of the B-1B program.

Statement of Problem

The B-1B SPO needs a technique, and ultimately a model, which can be used to analyze ATE test station requirements. The B-1B SPO will continually update its information, and therefore needs its own model and user's manual. This model would be used to examine the tradeoffs of cost versus avionics readiness. This model would also be useful for "what if" trade studies (for example, impact of different operational or maintenance concepts).

Objective of the Research

The overall objective of the research effort was to provide the B-1B SPO with a computerized model which provides the capability to assess the avionics maintenance workload on ATE test stations. Further, the research effort needed to develop criteria and procedures for selecting the best quantities of stations once the workload had been measured. This was accomplished by the sequential attainment of the following subobjectives:

1. It was essential to develop a detailed and accurate description (model) of B-1B avionics maintenance. As part of this description, it was necessary to consider all possible factors that could have a bearing on the avionics maintenance workload on ATE test stations. This description was obtained from a review of various B-1B logistics and maintenance planning documents and also from personal interviews with personnel from the B-1B SPO and HQ SAC. This conceptual description of avionics maintenance was to become the framework of all subsequent model development.

2. Since the research effort involved a practical application, it was necessary to obtain, collect, and review relevant data needed for the model development. B-1B operational data (such as number of aircraft per base, flying hours per aircraft per month, etc.) were obtained from the B-1B SPO. Reliability and maintainability estimates for each avionics LRU were obtained from the B-1B associate contractors. In addition, operational reliability and maintainability data were collected from actual experience with the F-16 fighter and the B-52 Offensive Avionics System (OAS) update. The data were examined by goodness-of-fit tests and other techniques to determine the most realistic probability distributions which were eventually incorporated during the model development.

3. A very detailed and complex simulation model of B-1B avionics maintenance was developed. This model was developed in the Q-GERT simulation language. The model simulates the flow of avionics LRUs from LRU failure to LRU repair (and return to base supply). Once the model was developed, the next step was to design the simulation experiment. As part of this experimental design, variance reduction techniques were used. The model was also used for sensitivity analysis since much of the data inputs were preliminary contractor estimates.

4. The research effort also included two tradeoff studies on avionics maintenance. First, it was necessary to conduct a cost-benefit analysis on procurement of additional test stations versus procurement of additional LRU spares. This was part of the effort in determining the best quantities of stations to procure. Second, the impact of a hypothetical B-1B deployment for a major conventional war was considered.

B-1B AVIONICS MAINTENANCE

B-1B Program Description

The B-1B is a new multi-role bomber which will eventually replace the B-52 as the penetrating bomber element of the strategic TRIAD. The B-1B also has collateral missions as a conventional bomber or as a cruise missile launch platform. The B-1B could potentially be used for naval and theatre conventional warfare, or for theatre nuclear warfare. Current plans call for the production and deployment of a force of 100 B-1B aircraft assigned to four main operating bases (MOBs). The B-1B has an integrated avionics system totaling over 424 installed line replacable units (LRUs) of which there are approximately 212 repairable LRUs. Tentatively, 109 LRUs have been designated for base level repair on the B-1B automatic test equipment (ATE). Other repairable LRUs are designated for base level repair on other support equipment or for depot level repair. The B-1B avionics consist of offensive avionics, defensive avionics, and miscellaneous avionics associated with other systems.

Avionics Maintenance

Organizational level maintenance consists of those tasks normally performed on-aircraft (on the flight line) by SAC maintenance technicians. For the avionics LRUs, this maintenance consists of debriefing, fault isolation removal and replacement of the failed LRU, and clean-up tasks such as documentation. Once the technicians complete the maintenance, they then fill out all maintenance documentation and also take the failed LRU to a production control point of the Avionics Maintenance Squadron (AMS) at intermediate level. The AMS is responsible for the repair of the failed avionics LRUs.

Once scheduled, an AMS technician repairs the LRU on the ATE test station of the appropriate type. In most cases, the technician is able to successfully repair the LRU at intermediate level; such a maintenance action is designated as Repairable This Station (RTS). For an RTS maintenance action, the repair takes place in a complicated sequence of events. First, the technician must set up the LRU on the ATE test station. The LRU is physically connected to the station via an interface test adapter (ITA). The technician then runs a performance test until the fault in the LRU is found. The fault is usually in one of the printed circuit boards of the LRU. The printed circuit boards are known as shop replaceable units (SRUs). Once the bad SRU is identified, the technician goes to supply to get a good replacement SRU. A small number of high failure rate SRUs are kept in a forward supply point; other SRUs are obtained from base supply. In any event, the technician removes the bad SRU from the LRU and replaces it with the good SRU. The technician then runs one complete performance test to verify the success of the repair. The repaired LRU is then removed from the station and taken to base supply. The technician also fills out all required maintenance documentation. Repairs may not always be this simple, however. The performance test may not always be able to fault isolate to a single SRU. Rather, the test might only fault isolate to a group of SRUs, or may indicate ambiguous results. In such a case, the technician resorts to manual troubleshooting to precisely locate the fault. Another complication is that not all maintenance actions lead to actual repairs. One possibility is that the LRU has failed in a way which is beyond the capability of the AMS to repair. For example, if the failure has occurred in a chassis of the LRU, and not a removeable SRU, the LRU would then be sent to depot for repair. Such a maintenance action is called Not Repairable This Station (NRTS). The LRU is then sent to depot (for greater facilities and higher skill level technicians) for repair. Another possibility is that the AMS technician cannot find any problem with the LRU. This might occur if the LRU was unnecessarily removed, or if there is some incompatibility between the CITS and the ATE. This type of maintenance action is called Retest Okay (RTOK). The LRU is taken to base supply without any repair being required.

AMS technicians that repair LRUs on ATE can come from one of four branches. Each AMS technician has an Air Force Specialty Code (AFSC) which describes his particular specialty. There is a dedicated AFSC for offensive avionics LRUs, for defensive avionics LRUs, for communication and navigation LRUs, and for automatic flight control and instrument LRUs. There is no cross-utilization of personnel between branches.

ATE Maintenance

Not only do LRUs fail and require maintenance, but so do the ATE test stations themselves. The test stations, being automatic test equipment, are complex electronic devices. When a station fails, a technician must be called to repair the station. The technician actually uses the station itself as part of the repair process. The technician must first isolate the station fault to a test replaceable unit (TRU). A TRU is the station equivalent of an LRU. The technician must then further isolate the fault to a bad SRU; the technician removes and replaces the failed SRU in a manner similar to LRU repair. Station maintenance is accomplished by technicians from the Precision Measurement Equipment Laboratory (PMEL) branch of the AMS.

DATA COLLECTION AND GATHERING

Reliability and Maintainability Estimates

LRU reliability estimates from each of the B-1B associate avionics contractors are regularly furnished to the B-1B SPO. The most recent estimates (as of August, 1983) formed the baseline for this research effort. Reliability in a logistics sense (as opposed to an engineering sense) is measured in Mean Time Between Demand (MTBD). Demand refers to a demand on supply (which occurs when an LRU is removed from the aircraft for corrective maintenance). This includes not only true (inherent) failures, but also includes induced failures and RTOKs. The time in MTBD is measured in aircraft flying hours (and not LRU operating hours). In addition to the contactor estimates, MTBD estimates were also obtained from HQ SAC maintenance and logistics personnel. The SAC estimates were typically, but not always, significantly more pessimistic than the contractor estimates. The SAC estimates were based on the current experience of like LRUs on the B-52, FB-111, and other systems.

LRU maintainability estimates were obtained from Rockwell International. However, modifications were developed based on discussions with HQ SAC maintenance and logistics personnel. For LRU repairs, the fault isolation to a single SRU by the performance test was assumed to be successful only 75% of the time. For the other 25%, the contractor estimated repair times were doubled to account for manual fault isolation. Other adjustments were made to account for repeated performance tests, NRTS maintenance actions, and the technician travel time to obtain SRU spares.

B-1B Operational Data

The building block of the B-1B fleet is the 16 aircraft squadron. A typical weekly flying schedule for a squadron was obtained from HQ SAC personnel. In any given week, only 8 aircraft will actually fly daily training missions. The other 8 aircraft will have alert obligations or will be undergoing scheduled (phase) inspections. The 8 aircraft that will fly missions will typically fly a total of 21 sorties in a week, leading to 107 flight hours in a week. A significant aspect of the flying schedule is that the sorties are not spread evenly during the days of the week. Typically, 3 sorties will be flown on Monday, 6 sorties on Tuesday and Wednesday, 4 sorties on Thursday, and 2 sorties will be flown on Friday. In addition, for any day of the week, the sorties are not spread evenly over the day. Sorties will typically be launched in the morning or in the late evening. What this means, from the point of view of a modeling strategy, is that the arrival rate of the avionics LRUs will not be constant over the maintenance day or during different days of the week.

In addition to B-1B flying hour data, information was also obtained on maintenance technician manning levels. These values represent maintenance manpower authorizations. As a first approximation, the maintenance policy will call for two 8-hour shifts of maintenance per work day (Monday through Friday). For modeling purposes, the authorizations were divided evenly between the two shifts. Moreover, only 60% of the technicians were assumed to be available for direct labor. The other 40% would account for illness, leave, or duties outside the scope of this research effort.

Goodness-of-Fit Analysis

The baseline inputs used in this research effort were reliability and maintainability estimates obtained from the B-1B associate contractors. However, the estimates represent mean values, and do not specify the distributional nature of the reliability and maintainability random variables. To select the best distributional assumptions for subsequent model development, operational data was obtained from the F-16 program and the B-52 OAS program for goodness-of-fit analysis.

The first factor considered was avionics reliability. Reliability in this context means Mean Time Between Demand (MTBD) and not the usual MTBF. It would be highly desirable to perform goodness-of-fit tests on the LRU inter-arrival times to obtain the best distributional approximation. However, actual operational maintenance data is not collected in this manner. Maintenance technicians record the time of the removal of the avionics LRU from the aircraft; it is not possible to reconstruct the actual inter-arrival times of the LRU failures. For this reason, it was only possible to test if the LRU arrival process could be approximated by a Poisson process. Another issue in testing the distribution of the LRU arrivals is the measure of time. Time could either be measured in flying hours or in calendar time. These two approaches would not be equivalent since the flying hours per base per month will not be perfectly constant; the variation in the flying hours would add to the variability of the LRU arrival process. Rather than derive a distribution for the flying hours and derive a distribution for the arrival process (in flying hours), it was simpler to fit a distribution to the LRU arrival process measured in calendar time. In the model discussed in the next section, flying hours per base per month is treated as a fixed constant and not as a random variable. The variability in the flying hours is reflected in the distribution of the arrival process. In any event, the key question was if the LRU arrival process could be approximated by a Poisson process.

Operational maintenance data on 19 major F-16 avionics LRUs were obtained from the F-16 SPO Centralized Data System (CDS). All of these 19 LRUs are repaired on the F-16 ATE. These LRUs were examined on an aggregate basis, and also two LRUs were examined on an individual basis. A chi-square goodness-of-fit was used to test the null hypothesis that the LRU arrival process was a Poisson process. Failing to reject the null hypothesis does not prove that the LRU arrival process is truly Poisson; it merely suggests that it should be a reasonable approximation. In examining the F-16 reliability figures, two concerns become quickly evident. First, in general, the F-16 avionics reliability had significantly improved over time since the initial F-16 deployment. To overcome this concern, the data used in the chi-square goodness-of-fit test reflected only the last six months of F-16 experience, thereby capturing a mature, steady-state situation. It should be pointed out, therefore, that the model described in the next section should only be used to estimate the workload during mature, steady-state experience. Second, the F-16 reliability varied considerably from one base to another. This is in part due to configuration differences, and in part due to differences in the quality of maintenance documentation. To estimate the variability at a single base, the data used in the test was generated by a single base, Nellis Air Force Base. For the data for a mature six month period from a single F-16 base, it was concluded that the Poisson process would be a reasonable approximation to the LRU arrival process. This was true for the arrivals of the two individual LRUs tested, and it was true for the aggregate arrivals of the 19 LRUs considered. Another point that needs to be mentioned is that the time interval selected (for each data point in the test) was one week. This would then, of course, smooth over any possible differences between days of the week or different times of the day.

The second factor considered was avionics repair time (or other task time) at intermediate level. Again, operational data was obtained from the F-16 program. For reasons discussed earlier, the data was obtained for a mature six month period from a single base. It was not possible to get separate data on set-up times, test times, repair times, or tear-down times; maintenance documentation is not that detailed. It was only possible to get task times for the overall maintenance action. It was possible, however, to separately analyze repairs, RTOKs, and NRTS actions. The mean and variance are quite different for different types of maintenance actions.

The task time was first collected for a single F-16 LRU. 52 observations were obtained for repairs (RTS actions), and 19 observations were obtained for RTOKs. There were only 3 NRTS actions during the six month period which is not sufficient to do a meaningful test. For the repairs and the RTOKs, the following distributions were tested for the possible distribution of the maintenance task times: exponential, Erlang, and lognormal. An Erlang was used instead of a (general) gamma because an Erlang is easier to simulate. It would have been necessary to resort to a (general) gamma only if all Erlang distributions were poor approximations. The results for the RTS maintenance actions were that the Erlang distribution was a good approximation; the exponential and lognormal distribution were not. Similar results were obtained for the RTOK maintenance actions. At this point two assumptions were made. First, the Erlang distribution, with the right k value, could be used to approximate the maintenance task times. Second, since there was not sufficient data to perform tests on the NRTS maintenance actions, it was assumed that the distribution of a NRTS action would be the same as for a RTOK action. The next step was to look at a second F-16 LRU (the inertial navigation unit) and also a B-52 OAS LRU (the signal data converter). The purpose was to make point estimates of the coefficient of variation (the mean divided by the standard deviation) to see if there were any patterns in Erlang k-values. At this point in the research effort, it was assumed that the RTS actions could be approximated by an Erlang distribution with a k-value of 2, and that RTOK and NRTS actions could be approximated by an Erlang distribution with k=5.

SIMULATION MODEL

Model Overview

A detailed Q-GERT simulation program was developed which can be used to measure the workload of avionics maintenance on intermediate level automatic test equipment (ATE). The model simulates the flow of avionics line replaceable units (LRUs) from LRU failure to LRU repair and return to base supply. The model also simulates the failures and maintenance of the ATE test stations themselves.

The model begins with a simulation of avionics LRUs which fail on B-1B aircraft while in flight. Sorties are assumed to be launched early in the morning and at night, so the arrival rate of avionics LRUs is not constant throughout the day. The arrival rate is also different for different days of the week due to different flying schedules. As an LRU fails, the model assigns various attributes (characteristics) to that LRU which describes its subsequent repair. These attributes include the test station type requirement, the intermediate level technician requirement, the type of maintenance action required, and the hours that will be required for the LRU repair.

Once an avionics LRU fails, it requires maintenance from the Organizational Maintenance Squadron (OMS) technicians. This maintenance consists of debriefing, removal of the failed LRU from the aircraft, and transportation of the LRU to the avionics shop. The model treats OMS technicians as constrained resources, and therefore failed LRUs must wait (or queue) for an available technician until maintenance may be performed.

Once at the avionics shop, the LRUs queue for intermediate level maintenance at the Avionics Maintenance Squadron (AMS). The LRUs require both a technician and a test station before maintenance can commence. Each of the station types and each of the AMS technician types are treated as constrained resources in the model.

Not only do avionics LRUs fail and require maintenance, but the ATE test stations themselves fail and require repair. The model generates ATE failures, and for each failure, assigns attributes which describe the subsequent repair. The model assumes that ATE repair takes precedence over LRU repair since an ATE test station must be in good working order in order to be used during the LRU repair.

Once the failed LRUs (or in an abstract sense, the failed ATE test stations) arrive at the shop, they must wait in queue until resources are available so that the intermediate level (avionics shop) maintenance can be performed. First, the failed LRU (or station) must wait until a station of the right type is available. The model allows for four station types with a user input quantity of stations of each type. The four types of stations for the B-1B are the Digital station, the Digital Analog Video station, the Radio Frequency station, and the Radar Electronic Warfare station. Each LRU has a known station type requirement and that station type requirement is set as an attribute value. The LRU is then routed to the correct station queue by conditional branching. Similar branching occurs for station maintenance. Second, the LRU (or station) must also wait until a technician of the right type is available. The model allows for the five AMS maintenance technician types. This includes the four types for LRU maintenance, and also the PMEL technician type for ATE station maintenance. Each LRU also has a known technician type requirement and that technician type requirement is set as an attribute value. The LRU is then routed to the correct technician queue by conditional branching. All station maintenance is routed to the PMEL technician queue. For example, if the failed LRU is from the B-1B ECM system, then the failed LRU will require (say) a Radar/EW test station and a defensive avionics technician.

Once both (station and technician) resources are available, the technician performs the first part of the maintenance action. For ATE maintenance and for LRU repair, this includes the time for setup and fault isolation. For LRU and station repairs, a SRU spare is required before any further maintenance may continue on the LRU or station. There is a 90% chance that the correct replacement SRU spare will be available at base level. If the spare part is not available for an LRU repair, the LRU is taken off of the test station and placed in awaiting parts (AWAP) status. The station and technician resources are then freed to do other work. If the spare part is not available for station maintenance, then only the technician resource is freed to do other maintenance. The station stays in down-for-parts status until the correct spare part can be ordered and shipped. For both LRUs and stations, the order and ship time for SRU spares was assumed to be a constant 8 days. On the other hand, if the correct spare part is available at base level, then the technician obtains the SRU spare and proceeds with the maintenance. The time to obtain the SRU spare is determined by probabilistic branching. There is a 25% chance that the SRU spare will be available at a forward supply point. This is assumed to take a constant time of 15 minutes. There is a 75% chance that the SRU spare will be obtained from base supply. This is assumed to take a constant time of one hour. Of course, for LRU NRTS and RTOK actions, no SRU spare is required, and there is a separate branch in the model for these cases. The maintenance technician then performs the second part of the maintenance action. For station and LRU repair, this includes the time to remove and replace the failed SRU and the time for teardown. Once the second part of the maintenance is complete, the station and technician are then freed to do other work. After maintenance is complete, LRUs are then taken to base supply.

Numerous statistic nodes have been included in the model to measure the performance of the maintenance queueing system. Interval statistics measure the base repair cycle time which is the time from LRU failure to return to base supply. Statistics are collected separately for each station type (for example, the Digital station). LRU maintenance statistics are also kept separate from ATE maintenance statistics.

Finally, the model keeps track of all changes associated with the shift changes. Maintenance technicians and stations are available from 0800 until 2400 each working day (i.e., two maintenance shifts per day). Resource alter nodes are used to control the beginning and end of each shift. Since Q-GERT alter nodes are nonpreemptive, however, this means that a maintenance task must be completed before the technician is allowed to go home. This is why the maintenance time was divided into two parts. It is possible to have a technician complete the first part of a maintenance action, go home at the end of a shift, and complete the second part of the maintenance action the following day. Dividing the maintenance time into two parts therefore minimizes the actual amount of "overtime" which occurs in the simulation model. Actual experience with the model indicates that the amount of "overtime" varies from zero to two hours per day (both shifts combined) depending on the station workload. This was considered reasonable and realistic since the SAC maintenance policy does allow for occasional "graveyard" maintenance to respond to maintenance workload requirements. The resource alter nodes also model scheduled maintenance on the ATE test stations. The stations are assumed to require a daily confidence test at the beginning of the first maintenance shift. This is assumed to take a constant time of 15 minutes for each station.

Design of Simulation Experiment

The purpose of the simulation experiment was to determine if varying test station quantities can influence LRU base repair cycle time. Therefore, base repair cycle time was the dependent variable of interest, and test station quantity was the factor to be varied. Treatments considered were quantities of one, two, or three test stations (of each station type) at each base. To reduce variance in the simulation experiment, the method of common random numbers was used. This means that, for each block, the base repair cycle time was measured against the three treatments with the same number of LRU failures, the same type of maintenance actions, and the same LRU repair times. This was achieved through the use of separate random number streams. This, of course, requires blocking to be used in the experimental design since the observations within a block were now related. In this experimental design, there are three treatments and ten blocks. This experiment has 18 degrees of freedom in the error term. A crude rule of thumb is that the degrees of freedom for the error term should be at least 10. This experiment was conducted a total of 12 times (one for each of four station types and one for each of three aircraft quantities).

For each of the 12 experiments, the factor of test station quantity was found to have a statistically significant influence (with $\alpha = 0.05$) on LRU base repair cycle time. Since the variance for different treatments was not constant, the (normal based) ANOVA was not appropriate to analyze the experimental results. For this reason, the experimental results were analyzed using Friedman's test, which is the nonparametric equivalent of one way ANOVA with blocking (complete randomized block design). Multiple pairwise comparisons in each of the 12 experiments were also almost always statistically significant. In fact, 32 out of 36 pairwise comparisons were found to be statistically significant.

Statistical significance, however, may not be the critical issue in determining test station quantities. These quantities should be selected to minimize the overall cost of avionics support. This cost should include both the cost of service (cost of test stations) and the cost of waiting (cost of avionics LRU spares). Test station quantities should be selected to achieve the optimum balance between cost of service and cost of waiting and achieve the lowest overall cost. This tradeoff will be analyzed later in the paper.

Sensitivity Analysis

Considerable sensitivity analysis was conducted on the various model inputs. This was essential since the baseline analysis was based on preliminary contractor estimates. Major elements investigated include LRU reliability, LRU test times, ATE reliability, ATE repair times, and other factors currently used in the model. It appears that the test station quantities may vary considerably with only modest and quite credible changes to most of the major elements. For this reason, it is too early in the B-1B program to precisely determine ATE test station quantities. Rather, it is only possible to determine a reasonable range of quantities.

Verification and Validation

Verification means ensuring that the model behaves exactly as it is intended. To assist in the verification process, the simulation model was the synthesis of 5 smaller models which were developed earlier. Each of these smaller models corresponds to a major portion of the final simulation model. Specifically, a smaller model was developed for (1) B-1B flying hour profile and LRU failure generation, (2) resource allocation (both test stations and technicians), (3) ATE failure generation, (4) detailed test and repair procedures, and (5) FORTRAN user functions and subroutines used in the model. Each of these smaller models was programmed with very detailed output and each was sufficiently simple to allow manual verification.

Validation is a process to ensure that the model realistically portrays the real world. The most important element of the continual process of validation was the coordination of all major ground rules and assumptions with personnel from HQ SAC that have had actual experience with avionics maintenance in the B-52 and FB-111 programs. These individuals provided a significant amount of feedback and constructive criticism. In addition, an effort is now being conducted to run this model on the B-52 OAS and its ATE. Model output will then be compared to the actual operational experience.

TRADEOFF STUDIES

Determining Station Quantity Requirements

A statistically significant improvement in base repair cycle time may have little advantage from a logistics support cost or weapon system availability point of view as discussed earlier. The purpose of this section, therefore, is to develop criteria and techniques for determining the best test station quantities by station type for each base.

The first point to be considered is that the test stations must be able to accommodate the mean avionics workload. In other words, at a minimum, the test station utilizations must always be smaller than 100%. If this were not the case, the queues would grow indefinitely, and the LRU base repair cycle time would become infinite. The minimum number of stations which achieve station utilization under 100% could therefore be regarded as an absolute floor for the test station quantities. In some cases, however, this approach may not be sufficient. It is possible that in certain situations the base repair cycle time, although finite, may nevertheless be "excessive" is some sense. If a rule or technique could be developed which could indicate when a predicted base repair cycle time was "excessive", it would then be appropriate to select the minimum number of test stations such that no base repair cycle time was "excessive".

One approach would be to compare the predicted base repair cycle time to some established standard. Such a standard should not be an arbitrary number, but should be selected to achieve a necessary level of support. Such support is in the form of the LRU base repair pipeline. For example, suppose that at a given base, two avionics processors fail per day. In addition, suppose that the planned or desired base repair cycle time is five days. This means that the base would require an LRU pipeline of ten spare processors at base level. Of course, the actual spares level computation would be more complex than this. First, it would have to account for the small percentage of time that the avionics LRUs are NRTSed (sent to depot) for repair. Second, it would have to make some distributional assumptions about the variability of the LRU arrivals per day and about the variability of the base repair cycle time. The spares level computation would then add a safety stock level to the average (expected) pipeline quantity. However, the principle remains the same. The approach is to postulate that LRU spares will be procured based on the assumption of a planned base repair cycle time. This planned base repair cycle time could then become the standard to judge whether the test station quantities were sufficiently large. Early in a program when little or no operational data is available, LRU spares are typically procured based on a base repair cycle time of four calendar days. It would then be a simple matter to compare the simulation model output to the standard of four days to judge if a given base repair cycle time were "excessive". The model output is based on five working days per week and excludes weekends. Thus, the four calendar day standard must be converted to 2.857 working days (by multiplying by five-sevenths) or 68.57 hours. Given that four calendar days is to be the standard, there is still some room for judgement as to how to compare the simulation output to the standard. One way would be to perform a one-sided statistical test of hypothesis. Suppose, for example, that the simulation model were very close to the standard of four days. Then, because of the statistical noise present in the simulation model output, it would be impossible to determine if the actual (expected) base repair cycle time were in fact below the standard. For a one-sided statistical test, suppose that the null hypothesis to be tested is that the base repair cycle time is less than or equal to four days, and that the alternative hypothesis is that base repair cycle time is greater than four days. This approach implicitly assumes that it is much worse to buy an unnecessary test station (make a type I error) then it is to buy an insufficient quantity of test stations (make a type II error). Similarly, the role of the null and of the alternative hypotheses can be reversed, and the judgement of the relative importance of the errors would therefore change as well. For cases where the station utilization is under 100%, however, it is not at all obvious as to which type of error is the more serious. If the two types of errors were to be weighed equally, then the one-sided test of hypothesis would not be the correct approach. Assuming that the simulation model output is normally (or at least symmetrically) distributed, then the way to weigh the two types of errors equally is to simply compare the sample mean of the simulation output to the standard. This was the approach taken for the remainder of this research effort.

Baseline Case

The simulation model was run for the baseline case (four day repair cycle time) for each of the bases and for each of the test station types. The mean of 10 replications was compared to the four calendar day base repair cycle time standard; minimum test station quantities were selected so that the standard could be achieved. The baseline case had an operational requirement for 24 test stations in total.

Logistics Support Cost Tradeoff

The first approach in determining test station quantities was to buy a sufficient quantity of test stations such that the predicted base repair cycle time was always under the standard of four calendar days. An alternative approach is to compare the costs of LRU spares (pipeline and safety stock spares) to the cost of test stations. Specifically, an additional test station would be procured as long as the savings in LRU spares (due to the reduced base repair cycle time) were greater than the cost of the additional station. In order for this approach to be valid, it is necessary to assume that the actual procurement of LRU spares will be based on the actual base repair cycle time, and not some standard factor.

The cost of LRU spares was calculated by use of the

Air Force Acquisition Logistics Division (AFALD) Logistics Support Cost (LSC) model. The Mean Time Between Demand (MTBD) for each LRU, the unit cost of the LRU, and the LRU base repair cycle time were the critical inputs. The MTBD estimates were the contractor estimates that were described earlier. The unit costs of each LRU, measured in FY 81 dollars, were also obtained from contractor estimates. The base repair cycle time was obtained from the output of the simulation model for each of the 12 situations. The LSC model, which calculates the LRU spares (pipeline and safety stock combined) quantity, was then used to compute the total LRU spares dollar investment required for each station type and station quantity at each base. Other LRU logistics support costs, such as maintenance manhour costs or inventory management costs, did not vary with changes in test station quantities and were not considered in this analysis (since they are constants). ATE test stations were assumed to cost 1.5 million dollars (in FY 81 dollars) for each of the four station types. In addition, it was assumed that the marginal cost for ATE SRU spares associated with a second or third station was zero. Most SRUs, due to their relatively high reliability, have a spare level (pipeline plus safety stock) of only one per base. Thus, adding a second or third station of the same type would usually cause no increase in the spare level (of one). Finally, recurring station maintenance costs (also in FY 81 dollars) were estimated as 10% of the unit cost of the station (or $150,000) per year. However, since this represents outlays over a 20 year period, the recurring station maintenance costs were discounted at a real rate of return of 10% per year. The "present value" of $150,000 per year over a 20 year period is 1.3 million dollars, so the total cost per station is estimated at 2.8 million dollars. The outlays for spares and test stations were treated as "front end" expenditures and therefore not discounted.

The logistics support cost tradeoff (between LRU spares and ATE stations) was estimated for each of the four station types and for each base. Interestingly, this second approach (logistics support cost tradeoff) yields results which were very close to the results of the first approach (compare base repair cycle time to four day standard). The first approach results in an overall quantity of 24 stations for the B-1B fleet, while the second approach results in an overall quantity of 26 stations for the B-1B fleet. Either approach, however, is based on very preliminary data and must be regarded as tentative.

Deployment for a Major Conventional War

All analysis discussed so far has been restricted to peacetime maintenance and support. The peacetime scenario calls for aircraft on daily alert and for aircrew training missions. In times of crisis, however, some B-1Bs might be dispersed to a satellite base. Only organizational level (and not intermediate level) maintenance would be performed at these bases, thus no additional test stations would be required to support such a dispersal. There is also no plan to use the B-1B ATE to support operations during a sustained nuclear war. Again, only organizational maintenance would be performed to support such operations. After all, the B-1B ATE is not hardened for protection against electromagnetic pulse (EMP) events. Another possible wartime mission would be a major conventional war. Although there is no formally documented requirement for such a mission, the potential conventional role for the B-1B is largely undefined. The analysis that follows is entirely hypothetical and does not constitute any formal SAC plan for actual usage of the B-1B.

The B-1B fleet consists of 5 squadrons of 16 aircraft each and 1 squadron of 10 aircraft for combat crew training. In this analysis, it was assumed that a flight of 6 aircraft would be taken from each of the first 5 squadrons and that 10 aircraft would remain behind for purposes of strategic alert. It was also assumed that the training squadron would also remain at its normal location. Thus, a total of 30 aircraft would be deployed while 60 aircraft would remain at the normal base. It was also assumed that the 30 deployed aircraft would be sent to two sites overseas; one site would receive 18 aircraft, and the other would receive 12 aircraft.

At the deployment sites, it was assumed that the aircraft would fly three times as many flying hours as in peacetime, and that the aircraft would fly the same amount every day for seven days per week. It was also assumed that the maintenance shifts would be expanded from two 8-hour shifts per day to two 12-hour shifts per day, and that the maintenance shifts would be expanded from five days per week to seven days per week. It was also assumed that the reliability and maintainability characteristics of the LRUs and ATE test stations would be the same in wartime as in peacetime. The model was adjusted to account for the expanded maintenance shifts and used to determine the minimum test station quantities required to support a four calendar day base repair cycle time. The total requirement to support the hypothetical deployment was found to be 31 test stations.

SUMMARY

The most significant product of this research effort was a detailed simulation model developed in the Q-GERT simulation language. The B-1B SPO now has its own version of this model which can be used to revise estimates of required quantities of ATE stations. These estimates can be updated on a periodic basis as better reliability and maintainability estimates become available. The second major product of this research effort was the development of techniques to determine the best choice of test station quantities. To accomplish this, the simulation model was expanded to include organizational (flight-line) maintenance and various administrative delays so that the model output is the (complete) LRU base repair cycle time. This is the time from LRU failure to repair and return to base supply. Two approaches were then developed in determining the best choice of station quantities. The first approach was to buy sufficient quantities such that the LRU base repair cycle time was shorter than some established standard. This standard would be the planned base repair cycle time used to determine the avionics LRU spare (pipeline and safety stock) quantities. The second approach was to perform a cost-benefit analysis on procurement of additional test stations versus procurement of additional LRU spares. This approach compared the costs of additional test stations (and the benefits of shorter base repair cycle times) to the costs of additional LRU spares (and the benefits of fewer test stations). The two approaches yield nearly identical results. Finally, a modified version was developed which could be used to determine test station requirements in a deployment for a conventional war.