ACCURACY QUANTIFICATION FOR INTEGRATED DIAGNOSTICS

David L. Doel
GE Aircraft Engines

Abstract If rational decisions are to be reached regarding the role of integrated diagnostics in the design of a weapon system, it is necessary to provide a quantitative means for evaluating the costs and benefits of the diagnostics relative to the goals of the weapon system. This paper describes progress towards developing such a process made as part of the USAF GIMADS contract. The emphasis of this paper is on diagnostic accuracy, even though accuracy is only one of the quantifiable attributes associated with diagnostics. The paper suggests answers to two specific questions concerning the implementation of diagnostics for weapon systems:

1) How are diagnostic accuracy requirements derived from weapon system level metrics included in high level requirements documents?

2) How does the detail designer achieve the diagnostic accuracy requirements that are allocated to his weapon system component?

The techniques described in this paper are applicable to a much broader class of problems (beyond integrated diagnostics), and are, in particular, relevant to Integrated Product Development (or Concurrent Engineering).

Introduction On February 13, 1987, the Generic Integrated MAintenance DiagnosticS (GIMADS) program was launched to determine how the systems engineering process can be applied to integrated diagnostics. A key area for study in

1 The work described in this paper was sponsored by the USAF GIMADS Program Office under contract F33657-86-C-0193.

2 A Friday.
accuracy needs was not judged to be a routine application of the procedure. Therefore, a separate effort was initiated to develop a process for "diagnostic accuracy quantification". This effort sought to answer four questions:

1) How are diagnostic accuracy requirements derived from the weapon system level metrics included in high level requirements documents?

2) How does the detail designer achieve the diagnostic accuracy requirements that are allocated to his weapon system component?

3) Does the expression of diagnostic accuracy vary with weapon system design phase (e.g. from concept development through deployment) and, if so how?

4) How can diagnostic accuracy requirements be validated and verified.

Thus far, useful results for the first two questions have been obtained. Our studies suggest that the answer to the third question is largely self-evident, once the accuracy derivation and allocation processes are well understood. The final question still needs considerable work before any useful results can be presented.

**Derivation of Diagnostic Requirements from the SON**

The GIMADS studies show that the "design quantification" process begins with the establishment of one or more models that express Statement of Operational Need (SON) metrics in terms of design variables. Most engineers in the aircraft industry have some familiarity with aircraft or engine performance models that achieve this type of purpose. Many engineers are also aware of Life Cycle Cost models (or sub elements thereof) and of reliability models that similarly address other SON metrics. Full design quantification requires that all quantifiable parameters from the SON be modeled in terms of design variables. If diagnostics is to be used to carry out the design intent, then the diagnostic metrics must be included in the SON models. The existence of such models provides the means for the designer to perform trade studies that identify design solutions to achieve the SON requirements.

Diagnosis of weapon systems problems is accomplished through the use of a number of different techniques. Many problems may be detected and/or isolated through use of Built In Test (BIT). This technique is particularly popular for electronic systems, but it is less appropriate for propulsion systems or structures. In other situations, one or more sensors may be used to derive the health of a component. Such a derivation might be accomplished in on-board software or after the flight using ground based systems. Ground based systems often use trending to recognize the onset of a problem. Other problems are recognized during periodic inspections, either visually or through use of support equipment. Support equipment may also be used to fault isolate problems detected during flight. Another technique that is used for diagnosis consists of manual resolution of the problem using a maintenance manual or "Tech Order" (T.O.) which describes the procedure to be followed. If all other methods are unavailable (or perhaps for other reasons) the diagnostician may use selective parts replacement as a diagnostic technique.

The development of diagnostic accuracy requirements must be able to address all of these techniques. Diagnostic accuracy has just as much meaning when applied to manual techniques driven by a T.O. as it does when applied to Built In Test. Both techniques are subject to human frailties in their application as well as other classic error sources. The successful derivation of diagnostic accuracy requirements must be data driven because of the human elements in the process. It is tempting to include only those error sources that are well understood and readily modeled. This approach is sure
to fail, because the more difficult to model error sources are generally the most significant.

It is also important to note that diagnostic accuracy is linked to other diagnostic design variables. For example, if diagnostic coverage is reduced, there will be a greater proportion of problems to be solved by arbitrary change of modules. For problems that are extremely rare or those which can readily be traced to a particular module, this may be acceptable. However, in general, as diagnostic coverage is reduced, diagnostic accuracy is apt to decrease as well. Another factor to be considered is the time available to perform the diagnosis. As this analysis time is reduced, the mechanic is forced to make a decision with less information. In the extreme, the mechanic is forced to change a module based on instinct rather than reasoned isolation. This again will decrease diagnostic accuracy.

Figure 1 shows a view of the process that is proposed for the derivation of diagnostic requirements. The source of the highest level requirements may be a SON, a SORD, a Statement of Work (SOW) or some contractual agreement between a contractor and a subcontractor. In the figure, this highest level requirements source is illustrated as being a SON. Many of the requirements that are expressed in the SON will be quantifiable. Some typical examples are shown in the figure. A more complete list of potential SON parameters is provided in table 1. For each quantifiable parameter identified in the SON (or other source document), a model should be generated which is capable of predicting the parameter using design variables as input.

**Table 1  Weapon System Level Performance Metrics**

<table>
<thead>
<tr>
<th>Metric</th>
<th>Unit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Life Cycle Cost</td>
<td></td>
</tr>
<tr>
<td>Development Cost</td>
<td></td>
</tr>
<tr>
<td>Production Cost</td>
<td></td>
</tr>
<tr>
<td>Support Cost</td>
<td></td>
</tr>
<tr>
<td>Mission Success Probability</td>
<td></td>
</tr>
<tr>
<td>Availability</td>
<td></td>
</tr>
<tr>
<td>Turn Rate</td>
<td></td>
</tr>
<tr>
<td>Maintenance Man Hours per Flight Hour</td>
<td></td>
</tr>
<tr>
<td>Maintenance Concept (2 or 3 Level)</td>
<td></td>
</tr>
<tr>
<td>Mean Time to Repair</td>
<td></td>
</tr>
<tr>
<td>RefTest OK's</td>
<td></td>
</tr>
<tr>
<td>Bench Check Serviceable's</td>
<td></td>
</tr>
<tr>
<td>CanNot Duplicate's</td>
<td></td>
</tr>
<tr>
<td>Crew Size</td>
<td></td>
</tr>
<tr>
<td>Skill Level</td>
<td></td>
</tr>
<tr>
<td>Ambiguity Group Limit</td>
<td></td>
</tr>
<tr>
<td>In Commission Rate</td>
<td></td>
</tr>
<tr>
<td>Mean Time to Diagnose</td>
<td></td>
</tr>
<tr>
<td>Abort Rate</td>
<td></td>
</tr>
<tr>
<td>Safety - Mishap Rate</td>
<td></td>
</tr>
</tbody>
</table>

In some instances, an element of a weapon system cannot be permitted to fail because it is critical to flight safety. Various techniques are available to address this problem. The most common is to adopt a conservative policy towards component replacement that calls for removal and replacement long before
Performance models are regularly generated to achieve this purpose. For example, a propulsion system cycle model is capable of predicting thrust, fuel flow and other performance parameters as a function of propulsion system design variables. These propulsion system models are linked to airplane models to predict aircraft performance, fuel usage, etc. with additional input of airplane design variables.

A life cycle cost model is another type of model that is used to predict high level figures of merit from design parameters. These models are substantially different from the performance models described above; yet, they serve the same type of role in the design process. Operational effectiveness models provide another example of models used to relate design variables to weapon system level figures of merit.

The underlying purpose for generating and maintaining these models is to support trade studies that evaluate various design solutions for satisfying the statement of need. Thus the various elements of the design being considered as possible solutions must be included in the models. This work proposes the addition of diagnostics so that it can be considered with other schemes for achieving the design goals. Clearly some weapon system goals can be accomplished in more than one way. For example, availability can be achieved either by high reliability, or at a lower reliability level by excellent diagnostics coupled with rapid maintenance. One or the other of these solutions may recommend itself on a cost basis. This is the type of trade study that can be performed to evaluate alternate design solutions.

If trade studies are to be accomplished involving diagnostics, then diagnostic design parameters must be included in the SON requirements models. Reference 3 provides an illustration of the addition of diagnostics to a mission model. At this stage of the design process, it is proposed that diagnostic accuracy be stated in terms of "false faults" and "misses". A later step will convert these measures into more fundamental diagnostic design parameters (sensor accuracies, etc.). The models may be based either on observed empirical relationships between the design variables and the weapon system metrics or on a physical understanding of the relationship. Clearly, the physically derived relationship is preferable; however, it cannot always be obtained. Once the higher level trade studies have been completed and values for false fault rate, miss rate, etc. have been derived, allocation to lower levels of the design is accomplished via use of more detailed models. The false fault rate (expressed as number per flight or in some equivalent manner) for the weapon system is the sum of the false fault rates for all of its elements. Thus, the use of more detailed models permits a partition of the false fault rate among the various weapon system components in a manner that the high level metrics are still achieved. In some cases, trade studies that evaluate various approaches to meeting SON figures of merit will be repeated at lower levels. These lower level trade studies must be retraced to the top level to assure achievement of SON metrics.

The allocation process will continue down to the smallest elements of the weapon system, where it should be possible to evaluate the allocated requirements against past experience for that type of element. This reference to prior experience may be modified to reflect the insertion of new technology. At this lowest level, it should be possible to confirm that the allocated requirements can be achieved, or failing this to consider design changes that may lead to their achievement. In any event, the results of the lower level analyses should be incorporated into the system of allocation models so that the details may be accumulated to monitor achievement of the SON figures of merit.

It is clear from the description of this process, that the results of this process
will only be as reliable as the data that are the basis of the estimates for the model. Thus, it is important that the participants in the design process have access to relevant data from earlier designs so that the model estimates will, as far as possible, be based on sound empirical evidence.

![Diagram](image)

**Figure 2** Diagnostic "Miss" Illustrated for 1-Sided Test

**Design Implementation of Diagnostic Requirements** Once requirements for the diagnostic accuracy have been established, it is the job of the diagnostic designer to select measurement devices to achieve the desired accuracy specifications. The requirements given to the designer are apt to specify figures of merit such as hit rate, miss rate or false fault rate. The designer is more likely to be able to select sensors that provide a specified accuracy or repeatability. Thus there is a need to translate the requirements to a form that can be used to specify product. Our study demonstrated that this translation involves the development of an additional mathematical model.

Perhaps the simplest case to be considered is one where a component has a single performance limit which is amenable to direct measurement. In this case, the component is considered to be serviceable if the performance measurement lies below the limit, and the component is judged to be faulted whenever the measured performance exceeds the limiting value. If the sensor used to determine the performance of the component were perfect (no error), there would be no false faults or misses. Thus, in this case, the occurrence of false faults and misses results from the sensor error. (Note that for more complex examples, there can be other sources for diagnostic problems such as faulty assumptions in the component performance model.)

Figure 2 exhibits the occurrence of a miss for an imperfect sensor. In figure 2, the sensor is assumed to be without bias; hence, the peak of the sensor error distribution is taken to be the true value of the performance parameter for the component. In figure 2, this true value lies above the limit; thus, the component should be judged to be faulted. However, as indicated by the shaded area in the figure, there is a nonzero probability that the component will be judged to be serviceable even though the performance is in the failed regime. The figure graphically illustrates the probability of a miss given that the performance is located at the precise value shown. To determine the composite miss rate, it is necessary to consider every possible value for the component performance as suggested by the component variation curve in the figure. Note that only those cases where the true performance is in the faulted region can lead to a miss.

Using the model of figure 2, the probability of a miss can be expressed as:

\[ P_m = \int_{-\infty}^{\infty} \left( \int_{x_1}^{\infty} f_c(x) \right) f_s(y) \, dy \, dx \]

where

- \( P_m = \text{Probability of a "miss"} \)
- \( x = \text{Value of measurement used to judge go/no-go status of a component} \)
\[ f_c = \text{Probability density function for component variation} \]
\[ f_s = \text{Probability density function for sensor error} \]
\[ x_1 = \text{Limiting value of "x" for the component} \]

The double integral can be solved either analytically or numerically (depending upon the assumptions regarding the probability distributions. In general, the miss probability can be expressed in terms of the sensor error and component variability standard deviations, and the component performance margin.

\[ Pf = \int_{-\infty}^{x_1} f_c(x) \left( \int_{-\infty}^{\infty} f_s(y) dy \right) dx \]

where

\[ Pf = \text{Probability of a "false fault"} \]

The detailed mathematical analysis for this simple model is given in Reference 3.

Note that the addition of a two sided limit, or of measurement bias (perhaps with its own probability distribution) does not add to the inherent difficulty of the analysis, although it does increase the numeric complexity. Figures 2 and 3 suggest the use of the Gaussian (or Normal) probability distribution. Often, the assumption of a Gaussian distribution proves to be a poor approximation in practice. In these cases, the appropriate distribution may be substituted for the Gaussian distribution and the resulting integrals may still be evaluated (probably via numerical techniques).

Figure 3 illustrates the analysis of a false fault for the same model. A false fault can only occur when the true performance is acceptable (lies within limits). The mean of the sensor distribution in figure 3 is indeed within acceptable performance bounds. However, there is some probability of a false fault for this example as suggested by the shaded area. Once again, it is necessary to consider all possible levels of component performance, together with their relative likelihood, in order to compute the probability of a false fault. The equation that represents the false fault probability for figure 3 is:
The design engineer should not be asked to perform these statistical analyses for misses and false faults. Instead, design rules or curves such as those shown in Figures 4 and 5 should be supplied. The actual curves shown in Figures 4 and 5 are derived using the simple model of Figures 2 and 3. More sophisticated models will lead to different curves; however, the design approach is similar. Figures 4 and 5 may be used to identify the combination of component margin, component variability and sensor error that yields sufficiently small values for false fault rate and miss rate. Often, the component margin and variability will be more difficult to alter than the sensor error; hence, sensor error is most likely to be derived from the analysis. In some cases it may be necessary to increase component margin in order to achieve the desired false fault and miss rates.

Note that if the component is very reliable, the assumption that it never fails may satisfy the requirements for miss rate. In this circumstance, diagnostics may be discarded in favor of achieving weapon system goals through reliability alone.

It should also be noted that the relative frequencies of false faults and misses may be adjusted by altering the limit (with no change to component margin, component variability or sensor error). Adjustment of the reject limit, without change to the system has the effect of trading false faults for misses. This strategy might be selected when one of these diagnostic accuracy measures is met with significant margin while the other is being missed by a small amount.

After reviewing the analysis described above, the designer of Technical Orders or manual diagnostic techniques may be tempted to despair at the thought of meeting quantitative measures for diagnostic requirements. This despair is justified, given the current state of data available to reach quantitative decisions. However, it should be possible to develop workable strategies for improving manual repair techniques. A successful strategy might involve the following steps.

1) Categorize the problem areas associated with manual diagnostic techniques including specific root causes that frequently lead to false faults or misses.

2) Explore whether the use of support equipment or built in test might cost effectively improve the diagnostic capability.

3) If not, try to determine the specific problems that result in misses or false faults. This might be achieved through interviews of maintenance personnel following known incidents or calling on expert evaluation of these incidents.

4) Once problems in the procedures and/or documentation have been isolated, rework the faulted materials in order to eliminate sources of confusion or to add steps that correct the problems. If the problem appears to be one associated with the experience level of the personnel, expert systems might prove to be a solution to the difficulty.
5) If possible, verify the modified procedures or documentation by testing in actual service conditions.

In many instances, the detection or isolation of a fault is not accomplished via use of a single sensor. For example, jet engine component performance is normally deduced indirectly through use of pressure, temperature and other sensors. In this case, one or more algorithms may be used to carry out the analysis. The algorithms will generally incorporate assumptions (e.g. that some potential fault mode is so rare as to be neglected) that have some associated uncertainty. When making use of one of the multi sensor algorithms, it is important to perform an appropriate study to properly reflect all of the error sources and their impact on the diagnostic accuracy. The potential error sources include the errors of the individual sensors, and also the potential errors in the assumptions.

It is also important to consider all sources of error for a specific sensor. A temperature sensor may do an excellent job of measuring the temperature at its particular location. Often, the temperature that is needed for the diagnosis is an average over some larger area, or at a location other than the precise location of the sensor. In this case, the error associated with the difference between the temperature being sensed and the desired sensor input must be included in the analysis. Frequently, this “displacement” error is the most significant element of the diagnostic error.

**Design Phase Dependence** Major military weapon systems are usually designed in phases, beginning with a Concept Exploration Phase and preceding through Dem/Val, Full Scale Development, Production and Deployment. As these phases succeed one another, the information available to carry out the design becomes increasingly more detailed. At some point in the design evolution, hardware and software are available for testing and evaluation. Prior to this time, analyses of the design must be based on analogy from earlier similar designs.

To a very great degree, the information available defines the type of analyses that can be carried out. At the earliest stages, there is very limited data available concerning the implementation of the weapon system requirements. The form of this data is more likely to be a functional description of the system than a physical description. Even with this limited information, one can begin to assess the relative criticality of the various functions to the weapon system requirements. Each of the design disciplines is attempting a comparable appraisal in order to identify where effort is likely to be needed. In the diagnostics arena, the designer should be concerned with difficulties associated with previous designs for achieving the specific function, the availability of new technology for achieving the design, and lessons learned that may dictate design approaches. It is at this stage that the designer can begin to identify issues that will require the greatest attention during the design.

As the design proceeds, details of the design will be developed in the various design communities so that the physical elements of the design can begin to be addressed. The criticality assessments from the earlier phases should still be valid. New data will include the FMECA (Failure Modes, Effects and Criticality Analysis) results. The availability of this information allows the diagnostic designer to transition from a top down, functional point of view to a bottom up, physical point of view. Specific failure modes can be addressed in the diagnostic design and failure rate statistics can be used to improve the diagnostic allocation. The diagnostic designer should be especially looking for insertion of new technology into the weapon system design so that associated diagnostic problems can be identified and addressed.
As design hardware becomes available and testing is started, it is possible to accumulate experience on the performance of the diagnostic design. It is important to realize that these results are not statistically significant; however, they can be useful for detecting problems in the diagnostic design. It is extremely important to future design efforts that data be gathered on the effectiveness of the diagnostic design. This begins during the latter stages of the development cycle and continues through the production and deployment phases.

**Verification of Diagnostic Design** The specification of the diagnostic design is only the first step to ensure that the needed diagnostic capability has been achieved in the weapon system design. Carefully designed testing and/or analysis is necessary to prove that the elements of the design sum up to the required capability.

Normally, the highest level design requirements do not directly address diagnostics. The diagnostic requirements, are instead, derived from these highest level requirements. Thus, it would appear to be most desirable to eliminate verification of diagnostics from the process in favor of verification of the overall weapon system figures of merit. Unfortunately, this is not practical due to the difficulty of obtaining satisfactory verification of the high level measures. To obtain adequate assurance that the weapon system achieves its goals, it is necessary to verify that the design elements perform according to their design requirements, and to validate the model used to derive these requirements from the high level weapon system requirements.

Even if this were not the case, it would still be desirable to acquire performance data for the diagnostic elements of a weapon system. This data is needed, as suggested above, during the design process in order to project the capability of future diagnostic systems. The quality of the judgements reached in the design process is only as good as the data on which they are based.

The verification process is particularly difficult for the diagnostic elements of a weapon system. The principal reason for this difficulty is the relative scarcity of diagnostic events. Well designed weapon system components fail rarely and thus seldom call for use of the diagnostic system. Hence, very extensive testing is required to be able to obtain a statistically significant evaluation of the diagnostic system. In many instances, the total experience of a fielded weapon system would not be adequate to evaluate the diagnostic system at a statistically significant level. Even special testing (with faults deliberately induced) may be precluded because of the prohibitive expense of such exercises. This difficulty places an extra burden on the analytical approaches for diagnostic design verification.

At the time that this work is being carried out, it is probably not possible to perform an adequate verification of the diagnostic capability of a new weapon system. This is because there is a lack of historic data available upon which to base the validation of SON models, especially as they address diagnostic issues. A principal objective of the GIMADS activity is to develop a process that may be used to carry out such a verification and to identify the data necessary to support that process so that it will be practical to verify the diagnostic design for future weapon systems. Of course, the future ability to perform such verifications will only be possible if the government and/or weapon system manufacturers begin to acquire and organize the needed data based on experience with current weapon systems.

**Conclusions** This paper has described a rigorous process for the derivation of diagnostic accuracy requirements from the Statement of Operational Need (SON) for a proposed weapon system. The process consists of the establishment of models.
for each of the figures of merit that are prescribed for the weapon system that relates these figures of merit to the design variables (including diagnostic accuracy). The models allow the design engineer to identify the degree of attainment of the weapon system goals in a quantitative manner, and to perform trade studies to weigh various design approaches for achieving the SON requirements. The design parameters that are proposed for diagnostic accuracy are "misses" (or the converse parameter "hits") and "false faults".

Clearly, the approach that is described is not limited to diagnostic accuracy considerations. The same technique may be used to derive requirements for any design variable. In fact, this technique is used routinely to derive propulsion system and aircraft performance designs, and is also used to evaluate life cycle cost and operational effectiveness of proposed designs. The ideas presented here merely extend this philosophy to include all design variables. This method offers a viable approach for integrated product design (or concurrent engineering).

The paper goes on to describe a method for expressing the diagnostic accuracy requirements in terms that are more fundamental in the diagnostic design. The method proposed involves a further modeling process to express false faults and misses in terms of sensor accuracy, component margin and component variability. These detailed design models must be developed for the elements of the design, and may involve different variables in different settings.

It will not, in general, be practical to perform all of the indicated analyses for all elements of a weapon system's design. The amount of labor to achieve a complete evaluation would be prohibitive. The designer should choose those details of the design that weigh most heavily on the weapon system goals, as identified in the SON, and be sure that these are properly analyzed. Many elements of the design can be accepted based on prior successful experience. The key areas for analysis are those that involve new technology or those that will play a large role in mission achievement.

Currently, it is not likely to be practical to fully validate the diagnostic accuracy design in a new weapon system. In fact, it is not clear that this is an achievable task within reasonable cost constraints (due to the infrequency of some diagnostic events). If progress in this area is to be achieved, a well thought out plan for acquiring and using diagnostic event data must be developed.

Acknowledgements The work described in this paper reflects many valuable discussions with various GIMADS program participants. Many valuable ideas have originated from Fred Stevenson and Paul Abendroth of the USAF GIMADS Program Office, from Bob Cook of the USAF LSA Office, from Dick Clothier, Mike Golas, Del Doss, Grant Sorensen, J. R. Wildridge and Chris Dollin of General Dynamics and from George Neumann of Giordano Associates. The work was sponsored by the USAF GIMADS Program Office under contract F33657-86-C-0193.

References