

**Thomas A Ferryman, Brett D. Matzke, John E. Wilson, Julia Sharp, Frank L. Greitzer**  
**Pacific Northwest National Laboratory**  
**Ed Hilferty**

**Naval Sea Systems Command, Naval Surface Warfare Center**

## **Engine Removal Projection Tool**

### **ABSTRACT**

This paper presents results of R&D associated with creating a new algorithm and software program to predict engine removals/down time for US Navy surface ships. The US Navy has over 3500 gas turbine engines used throughout the surface fleet for propulsion and generation of electrical power. Historical data are used to forecast the number of engine removals for the next ten years and determine engine down times between removals. To update and improve current prediction methods, we tested over 60 techniques on twenty years of Navy data from over 3100 engines and 120 ships. Among the techniques investigated were moving averages, empirical negative binomial, generalized linear models, Cox regression, and Kaplan Meier survival curves. Our approach was to apply the *best* algorithm based on its performance on real-world data, and to implement the selected algorithm in a new software program that allows the user to select a specific engine type, forecast time period, and op-tempo. Graphical displays and numerical tables present forecasts and uncertainty intervals. The technology developed for the project is applicable to other logistic forecasting challenges.

### **SPONSOR CLEARANCE**

This work was conducted under a contract from NAVSEA/NSWCCD to PNNL under a Related Services Agreement with the U.S. Department of Energy (Contract DE-AC06-76RLO 1830).

### **INTRODUCTION and BACKGROUND**

The US Navy has over 3500 gas turbine engines used throughout the surface fleet for propulsion and the generation of electrical power. The US

Navy, Naval Sea Systems Command (NAVSEA) Marine Gas Turbine Information System (MGTIS) Program archives engines' historical operating and removal data for use in logistics and lifecycle planning. One analysis conducted by the MGTIS Program is to project the number of engine removals for the next ten years and determine engine down times between removals. To support this effort, NAVSEA currently uses a software program, implemented in FORTRAN in the early 1970s. The program runs on a mainframe computer and is inconvenient to use. In addition, because it does not account for certain important factors (such as decommissioning), various manual manipulations and subjective judgments are needed to make the program fit the current needs of the MGTIS program. Finally, the underlying statistical methods employed by the FORTRAN program are not well documented or understood. Thus there was a need for a new program that is more convenient to use and that is based on a defensible scientific methodology.

This paper presents results of the research and development conducted to develop algorithms to forecast engine replacements and the implementation of a new user-friendly software program to enable the NAVSEA user to formulate forecasts that support their logistics needs. Requirements for this work included the following: The algorithms need to be defensible, well-documented and outperform the existing algorithms. The software needs to extend the forecasting algorithm work to present a variety of information based on the data analysis and forecast. It must present intuitive displays, be easy to run with minimal training, and run under Windows as a Web-based application.

## FORECAST ALGORITHM DEVELOPMENT

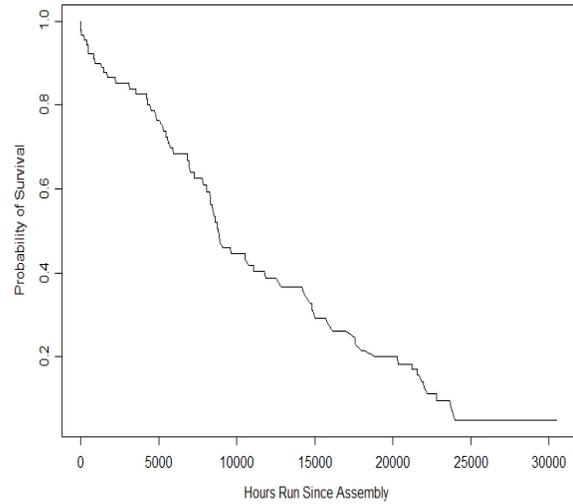
We tested over 60 techniques on almost 20 years of data collected from over 3100 gas turbine engine assemblies and 120 U.S. Navy ships. Among the techniques investigated as the forecast basis were moving averages, empirical negative binomial, general linear models, Cox regression, and Kaplan-Meier curves, most of which are documented in engineering, medical, and scientific research (Lawless 1982). We applied these techniques to the data and used test set validation to quantify the accuracy of each method and choose the best algorithm.

The Kaplan-Meier estimate is defined as follows: suppose that there are  $n$  engines and that there are  $k$  ( $k \leq n$ ) distinct times  $t_1 < t_2 < \dots < t_k$  at which removals occur. The possibility of there being more than one removal at  $t_j$  is allowed, and we let  $d_j$  represent the number of removals at  $t_j$ . In addition to the lifetimes  $t_1, \dots, t_k$ , there are also censoring times  $L_i$  for engines whose lifetimes are not observed. The product-limit estimate of  $S(t)$  is defined as

$$\hat{S}(t) = \prod_{j:t_j < t} \frac{n_j - d_j}{n_j}$$

where  $n_j$  is the number of engines at risk at  $t_j$ , that is, the number of engines running and uncensored just prior to  $t_j$  (Lawless 1982).

Figure 1 shows an example where  $t_1, \dots, t_k$  are plotted versus  $\hat{S}(t)$ . This shows a step function where  $\hat{S}(t)$  decreases as the values of  $t_1, \dots, t_k$  increase. It is logical to assume that each engine has some non-zero chance of failing at any point in time when we are using this distribution for forecasting purposes, and in reality,  $S(t)$  should be continually decreasing as time increases. Therefore, we modified the method to smooth values of  $\hat{S}(t)$  between removal times.



**Figure 1. Example of Kaplan-Meier estimates**

To compute  $\hat{S}(t)$  for time values not equal to one of the values  $t_1, \dots, t_k$  we assumed the hazard rate between two consecutive values of  $t_1, \dots, t_k$  is a non-zero constant unique to times between those two points. For example, let's say there are two removal times from the original data,  $t_a$  and  $t_c$ , where the probability of survival for the Kaplan-Meier estimate is unchanged between the two points, and we wish to compute an estimated probability of survival for a time  $t_b$  that falls between times  $t_a$  and  $t_c$ .  $S(t)$  for  $t_b$  is computed as follows:

$$\hat{S}(t_b) = \hat{S}(t_a) * \left( \frac{\hat{S}(t_c)}{\hat{S}(t_a)} \right)^{(t_b - t_a)/(t_c - t_a)}$$

Another issue is how to compute forecast probabilities of survival when time values approach or exceed the highest removal time in the data. For this case, we assume that hazard rates are constant beyond a time where there are  $r$  values of  $t_1, \dots, t_k$  exceeding this time in the data. We would assume a constant hazard for the time value that achieves the mean Kaplan-Meier estimate of the  $t_{k-r}$  and  $t_{k-r-1}$  removal times where  $1 < r \leq k$ . The mean Kaplan-Meier estimate of the  $t_{k-r}$  and  $t_{k-r-1}$  removal times occurs at a time  $t_{radj}$  computed as follows:

$$t_{radj} = t_{k-1-r} + \frac{\ln\left(\frac{\hat{S}_{k-r} + \hat{S}_{k-r-1}}{2 * \hat{S}_{k-r-1}}\right)}{\ln\left(\frac{1}{\left(\frac{S_{k-r}}{S_{k-r-1}}\right)^{t_{k-r}-t_{k-r-1}}}\right)}$$

Let  $totalhrs_{radj}$  be the total number of hours and the total number of removals from the  $n$  engines that exceed  $t_{radj}$ , and  $totalrems_{radj}$  be the number of removals that occur at times greater than  $t_{radj}$ .  $S(t)$  is then computed as follows when  $t > t_{radj}$

$$\hat{S}(t | t > t_{radj}) = \hat{S}(t_{radj}) * \left(1 - \frac{totalrems_{radj}}{totalhrs_{radj}}\right)^{t-t_{radj}}$$

To compute projected removals for future months, projected operating hours for all engines are totaled. Assemblies in each ship class are assumed to operate at the average rate of the class's operating history. Using the average operating rate for all assemblies, assemblies are stepped through the ages they would reach each month. At each step, the installed population in each interval is subjected to the probability of being removed. Removed assemblies accumulated as expected removals are replaced by zero-timed assemblies for the next iteration<sup>1</sup>.

## EXPERIMENTAL RESULTS ASSESSING THE FORECAST ALGORITHMS' EFFECTIVENESS

The algorithms were evaluated on 12 different engine types using an iterative approach where we retrieved the first few years of data, applied the algorithm, and predicted the next year's removals. For the next iteration, we would drop the earliest year's data we used in the previous iteration, and add the next year's data. This was repeated until we had predicted the most recent year's removals. For each engine type, a residual analysis was performed on these results

<sup>1</sup> Methodology for MTBR and Removal Projections was based on notes provided by NAVSEA.

to evaluate how well each method performed. Many summary statistics and plots were used in our evaluation. Table 1 shows an example of a few of the summary statistics used in our evaluations.

Table 1. Summary of Test Results for 4 of the 60 Algorithms Tested

Algorithm	Mean Absolute Difference	Root MSE	Correlation
Kaplan-Meier	4.42	5.53	0.51
Cox PH	10.48	11.86	-0.22
GLM	6.14	8.59	-0.07
Negative Empirical Binomial	4.77	5.87	0.45

The Kaplan-Meier estimate method with the modifications we have described performed better than other methods when there were at least several removals in the data. A simple constant hazard rate based on removals per hour worked consistently well when there were only a few removals in the data used to compute the distribution. The method implemented uses the constant hazard for low numbers of removals, and the modified Kaplan-Meier estimate when a minimum number of removals are present in the test set data.

## ERP TOOL, USER-FRIENDLY SOFTWARE WORKSTATION

The algorithm chosen as most effective is a core part of the ERP Tool. Other key characteristics are important to allow it to function in the NAVSEA environment:

- Web-based, running on a windows operating system.
- Use .NET technology and the C# language.
- The ERP Tool and all MGTIS data remain on the NAVSEA server. Raw data will not be passed to the client workstation.
- The ERP Tool accesses the MGTIS ORACLE database on the NAVSEA server to acquire user-specified data (e.g., data for

the past three years) to be used by the ERP analysis. Input data include Ship ID, Engine ID, dates, operating hours, engine failure codes, selected component failure codes, and projected operating schedules.

- Outputs displays
  - Engine Hazard Rate and Reliability Summary table
  - Projected Engine Removals table
  - Charts of actual versus predicted removals by engine type and year in accordance with a user specified prediction period.
  - Charts of actual cumulative removals versus projected removals by engine type and month for a selected fiscal year.
  - Aggregated chart of engine removals by year including historical data and projected data (e.g., out to ten years).
  - Graph of removal rates for a selected time period.
  - Bar charts broken out by engine type of MTBF values by date
  - Engine MTBF values augmented for relevant components.
  - Component MTBF.
  - Projections based on user-selectable Peacetime Operational Tempo and/or Wartime Operational Tempo.

Many of these capabilities are illustrated below.

### Data Issues

Data for the prediction tool is extracted from MGTIS database, which is a relational database stored and managed by Microsoft's SQL Server 2000.

The data needed for the prediction tool is categorized as follows:

- Monthly Engine Operating data
- Engine Removal data
- Projected Engine Operating Hours
- Current Engine Location

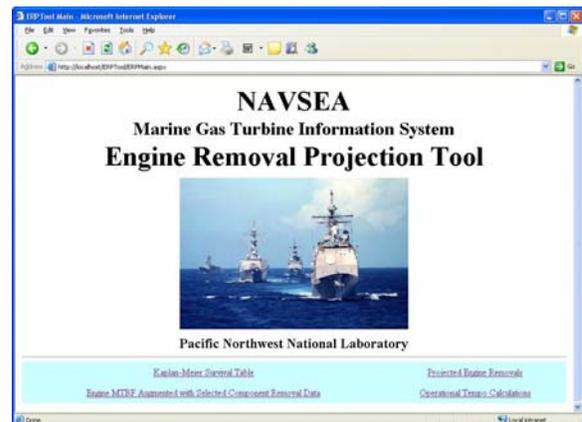
On start-up, the prediction tool gathers historical operating data into a single file to provide quick access for the various prediction functions. This includes organizing Monthly Engine Operating data to show “operating hours since removal”

and “operating hours since new” for each month. Removals are also associated with the monthly operating records.

The current engine location data is used to determine the inventory of engines to be used for engine removal predictions. This location information is combined with projected engine operating hours data (and optionally with wartime multiplier data) to determine the number of engine and operating hours for the prediction periods. Historical operating and removal data are used to establish removal rates. Those removal rates are used with the operating hours for prediction periods to predict engine removals.

### User Interface

The user interface is a series of “Web pages” generated by the ERP-Tool program, which is an ASP.NET program that resides on a server. The ERP-Tool program interacts with the MGTIS data through Microsoft SQL Server 2000. The Web pages are sent to the user through the Microsoft Internet Information Services (IIS). The Main Menu page is shown in Figure 2, below.



**Figure 2. ERP Tool Main Menu Page**

The Main menu page provides links to other pages. Each page represents a particular function. The following is a tree of the various functions and menus in the ERP-Tool program:

- Main Menu
  - o Kaplan-Meier Survival Table

- Engine MTBF Augmented with Selected Component Removal Data
- Historical Operating Hour and Installation Summary
- Operational Tempo Calculations
- Projected Engine Removals Menu (see Figure 3, below).
  - Projected Engine Removal Table
  - Projected Engine Removal Graph
  - Cumulative Projected Engine Removal Graph (Multiple Years)
  - Cumulative Projected Engine Removal Graph (Single Year)

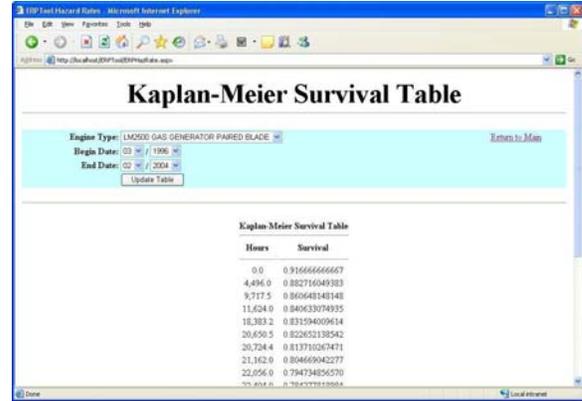


**Figure 3. Engine Removal Projection Menu**

A synopsis of each function follows.

***Kaplan-Meier Survival Table***

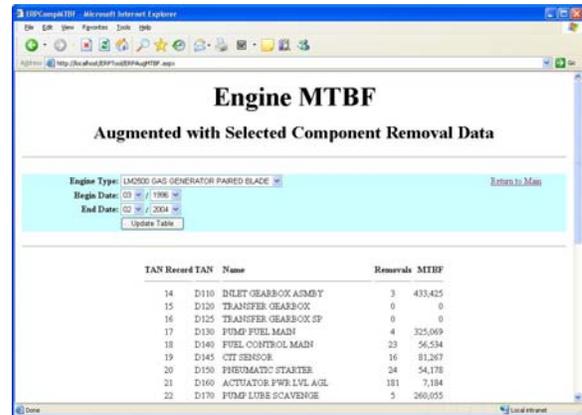
This provides the table of survival values versus operating hours that are used to predict removals (Figure 4).



**Figure 4. Kaplan-Meier Survival Table**

***Engine MTBF Augmented with Selected Component Removal Data***

This provides a list of components for a selected engine and the mean time between failures (MTBF) for each of those components during the chosen operating period (see Figure 5). Additionally, it provides the overall MTBF for the engines which includes component failures.



**Figure 5. Engine MTBF**

***Historical Operating Hour and Installation Summary***

This provides a list of operating hours and number of engines that operated during the chosen period, broken down into hour intervals (chosen by the user). It also includes the number of removals that occurred in the various intervals.

## Operational Tempo Calculations

This provides a means for the user to change the multipliers to be used for projected monthly operating hours during wartime conditions (see Figure 6). There is a different multiplier used for each engine type on each class of ship where the engine is used. Each user has a separate set of multipliers that persists during the current web session (cleared when the user exits the Web browser). The underlying peacetime operation hour projections as well as the default wartime multipliers are stored in tables in the database, and are changed outside the ERP Tool program.

Class	Desc	Engine	Peacetime Multiplier	Wartime Calculated
1 DD	501-E17	COMPRESSOR	200.00 (1.33)	264.00
1 DD	501-E17	TURBINE	200.00 (1.33)	264.00
1 DD	LM2500	GAS GENERATOR PAIRED BLADE	130.00 (1.33)	172.90
1 DD	LM2500	POWER TURBINE	130.00 (1.33)	172.90
10 MCM	T-1302S	SOLAR GAS GENERATOR	10.00 (1.00)	10.00
3 FFG	LM2500	GAS GENERATOR PAIRED BLADE	115.00 (1.75)	201.25
3 FFG	LM2500	POWER TURBINE	115.00 (1.75)	201.25
4 CO	501-E17	COMPRESSOR	195.00 (1.50)	292.50
4 CO	501-E17	TURBINE	195.00 (1.50)	292.50
4 CO	LM2500	GAS GENERATOR PAIRED BLADE	125.00 (1.50)	187.50
4 CO	LM2500	GAS GENERATOR SINGLE SHAFT	125.00 (1.50)	187.50
4 CO	LM2500	POWER TURBINE	125.00 (1.50)	187.50

Figure 6. Operational Tempo Calculations

## Projected Engine Removal Table

This predicts the number of engine removals for a selected period based upon the removals that occurred in a selected historical period (see Figure 7). If sufficient removals occurred in the historical period, a Kaplan-Meier prediction method is used; if not, then a simple rate prediction method is used. The number of predicted removals along with an upper 90% confidence limit, mean time between removals (MTBR), and average removal age are all displayed in tabular format.

Projection Period	Engine Operating Hours	Installed Engines	Expected Removals	Upper 90% Bound	Avg. Expected Removal Age	Projected MTBR (Hrs)	Stable MTBR (Hrs)
200510 - 200609	416,209	156	10	14	42,731	42,400	34,213
200610 - 200709	416,209	156	8	12	48,942	51,766	34,213
200710 - 200809	416,209	156	8	12	50,099	53,109	34,213

Figure 7. Projected Engine Removals Table

## Projected Engine Removal Graph

This function operates basically the same as the Projected Engine Removal Table, except that the projected number of removals and upper 90% confidence are plotted on a graph (see Figure 8). If part or all of the prediction period occurs in the past where historical data is available, actual removals are also plotted on the graph. Other predicted variables like MTBR and average removal age are not plotted.

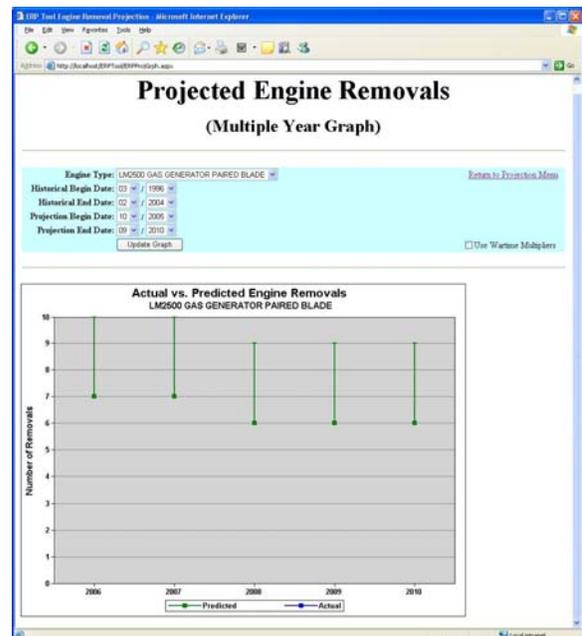
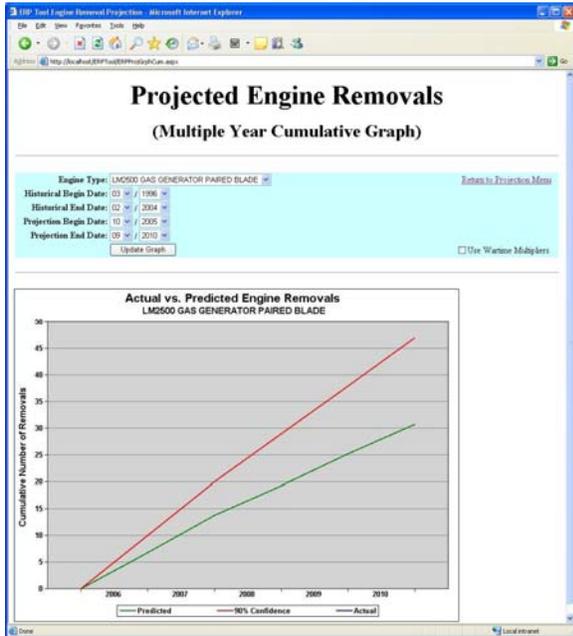


Figure 8. Projected Engine Removals Graph

### ***Cumulative Projected Engine Removal Graph (Multiple Years)***

This function operates basically the same as the Projected Engine Removal Graph, except that the plotted removals are accumulated from one time period to the next. This is shown in Figure 9.



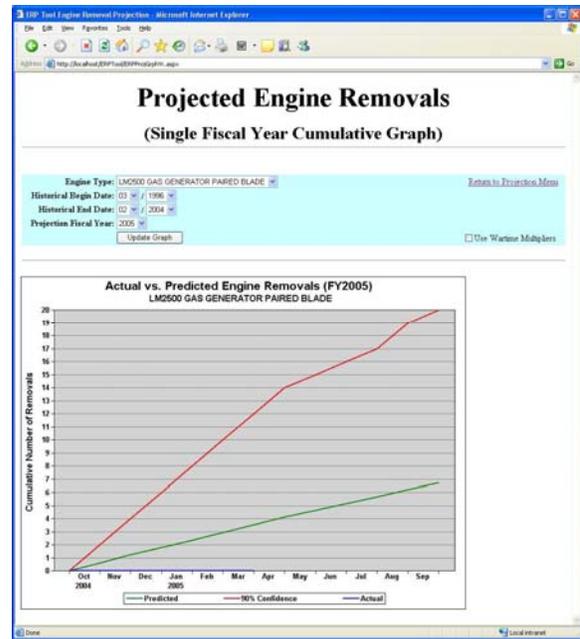
**Figure 9. Cumulative Projected Engine Removals—Multiple Years**

### ***Cumulative Projected Engine Removal Graph (Single Year)***

This function operates basically the same as the Cumulative Projected Engine Removal Graph (Multiple Years), except that the prediction period is limited to a single fiscal year (see Figure 10).

## **CONCLUSIONS**

This project demanded the merger of good mathematical statistical science and computer science to create an effective tool to support NAVSEA’s logistics needs. The initial implementation of the ERP Tool is on the collection of 3100 Gas Turbine Engines.



**Figure 10. Cumulative Projected Engine Removals—Single Year**

The ERP Tool algorithm results in approximately a 10% improvement over the current practice. This will help guard against purchasing in advance of demand and, in addition, will help avoid the more serious situation—a shortfall. The software uses the *best* algorithm in combination with user-friendly interfaces and intuitively understandable displays. The user can select a specific engine type, forecast time period, and op-tempo. Graphical displays and numerical tables present forecasts and uncertainty intervals.

The technology developed for the project is applicable to other logistic forecasting challenges any where improved forecasting and/or user-friendly software tools could help the logistics tasks

## **REFERENCE**

J.F. Lawless, Statistical Models and Methods for Lifetime Data, John Wiley & Sons, New York, 1982.

## **ACKNOWLEDGEMENTS**



**Thomas Ferryman** is a Battelle Chief Scientist at the Pacific Northwest National Laboratory with over 30 years of experience in system engineering and mathematics/statistics. He has developed prognostic tools for use on gas turbine engines. He leads the technical development

of aviation safety data analysis tools for NASA (numeric, categorical and/or text data). Prior to coming to Battelle, Dr. Ferryman was Chief Systems Engineer for Lockheed leading a major weapon system modification (AC130H Gunship) and Senior Operations Research Analyst for the U.S. Navy.



**Brett Matzke** is a Battelle Research Scientist at the Pacific Northwest National Laboratory. His research interests include statistical modeling, multivariate analysis, experimental design, and automation of analyses. He has previously served as a Senior Product Line Manager of

Quantitative Analysis for Reader's Digest, and a Senior Statistical Consultant for CAMO, Inc. He has an M.S. degree in Statistics.



**John E. Wilson** is a Battelle Research Scientist at the Pacific Northwest National Laboratory. His research interests include desktop software development, graphic user interface development, and process automation and visualization. He is the lead programmer of the

Visual Sample Plan software, a sample planning software package, used by thousands of users worldwide. He previously served as Computer Services Manager and software engineer for the Grand Junction Office of Oak Ridge National Laboratory for over 10 years.



**Julia Sharp** is a doctoral student in Statistics at Montana State University in Bozeman, Montana. Her research interests include statistical modeling, experimental design, and biostatistics. Ms. Sharp recently held a summer internship

position at Pacific Northwest National Laboratory. She has an M.S. degree in Statistics from Montana State University and a B.S. in Mathematics from the University of Evansville in Evansville, Indiana.



**Frank L. Greitzer** is a Battelle Staff Scientist at the Pacific Northwest National Laboratory with 30 years of experience in user-centered design and development of advanced technology software systems,

including performance support/training systems and information technology systems to augment human performance. Currently managing the R&D effort in support of the work reported here, Dr. Greitzer has also managed seminal research developing an onboard, proof-of-concept prototype system (REDI-PRO Real-time Engine Diagnostics-PROgnostics) to diagnose and predict faults in Army gas turbine engines.



**Ed Hilferty** has been an Engineering Statistician with the Naval Surface Warfare Center-Ship System Engineering Station, Philadelphia PA for 15 years. He has held various positions including project manager and senior programmer for the

Navy's Marine Gas Turbine Reliability and Maintainability OPHR program and currently is also the project manager in the development, implementation, and management of the new Marine Gas Turbine Information System (MGTIS) and Autolog. He is the station's senior statistician with responsibility for generating and reviewing much of metrics for the Propulsion and Power Systems Department. Ed holds an M.S. degree in Applied Statistics.