Received by OSTI

JUN 04 1990

NUREG/CR-5436
SAIC-90/1130
Vol. 1

The Development and Evaluation of Programmatic Performance Indicators Associated with Maintenance at Nuclear Power Plants

Main Report

Prepared by J. Wreathall, J. Fragola, P. Appignani,
G. Burlile, Y. Shen

Science Applications International Corporation

Prepared for
U.S. Nuclear Regulatory Commission

DO NOT MICROFILM
COVER

DISTRIBUTION OF THIS DOCUMENT IS UNLIMITED

AVAILABILITY NOTICE

Availability of Reference Materials Cited in NRC Publications

Most documents cited in NRC publications will be available from one of the following sources:

1. The NRC Public Document Room, 2120 L Street, NW, Lower Level, Washington, DC 20555
2. The Superintendent of Documents, U.S. Government Printing Office, P.O. Box 37082, Washington, DC 20013-7082
3. The National Technical Information Service, Springfield, VA 22161

Although the listing that follows represents the majority of documents cited in NRC publications, it is not intended to be exhaustive.

Referenced documents available for inspection and copying for a fee from the NRC Public Document Room include NRC correspondence and internal NRC memoranda; NRC Office of Inspection and Enforcement bulletins, circulars, information notices, inspection and investigation notices; Licensee Event Reports; vendor reports and correspondence; Commission papers; and applicant and licensee documents and correspondence.

The following documents in the NUREG series are available for purchase from the GPO Sales Program: formal NRC staff and contractor reports, NRC-sponsored conference proceedings, and NRC booklets and brochures. Also available are Regulatory Guides, NRC regulations in the *Code of Federal Regulations*, and *Nuclear Regulatory Commission Issuances*.

Documents available from the National Technical Information Service include NUREG series reports and technical reports prepared by other federal agencies and reports prepared by the Atomic Energy Commission, forerunner agency to the Nuclear Regulatory Commission.

Documents available from public and special technical libraries include all open literature items, such as books, journal and periodical articles, and transactions. *Federal Register* notices, federal and state legislation, and congressional reports can usually be obtained from these libraries.

Documents such as theses, dissertations, foreign reports and translations, and non-NRC conference proceedings are available for purchase from the organization sponsoring the publication cited.

Single copies of NRC draft reports are available free, to the extent of supply, upon written request to the Office of Information Resources Management, Distribution Section, U.S. Nuclear Regulatory Commission, Washington, DC 20555.

Copies of industry codes and standards used in a substantive manner in the NRC regulatory process are maintained at the NRC Library, 7920 Norfolk Avenue, Bethesda, Maryland, and are available there for reference use by the public. Codes and standards are usually copyrighted and may be purchased from the originating organization or, if they are American National Standards, from the American National Standards Institute, 1430 Broadway, New York, NY 10018.

DISCLAIMER NOTICE

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, or any of their employees, makes any warranty, expressed or implied, or assumes any legal liability of responsibility for any third party's use, or the results of such use, of any information, apparatus, product or process disclosed in this report, or represents that its use by such third party would not infringe privately owned rights.

DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency Thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

DISCLAIMER

Portions of this document may be illegible in electronic image products. Images are produced from the best available original document.

The Development and Evaluation of Programmatic Performance Indicators Associated with Maintenance at Nuclear Power Plants

Main Report

Manuscript Completed: August 1989
Date Published: May 1990

Prepared by
J. Wreathall, J. Fragola, P. Appignani,
G. Burlile, Y. Shen

Science Applications International Corporation
2929 Kenny Road
Columbus, OH 43221

Prepared for
Office of Nuclear Regulatory Research
U.S. Nuclear Regulatory Commission
Washington, DC 20555
NRC FIN L1183

MASTER

DISTRIBUTION OF THIS DOCUMENT IS UNLIMITED

Abstract

This report summarizes the development and evaluation of programmatic performance indicators of maintenance. These indicators were selected by: (1) creating a formal framework of plant processes; (2) identifying features of plant behavior considered important to safety; (3) evaluating existing indicators against these features; and (4) performing statistical analyses for the selected indicators. The report recommends additional testing.

Table of Contents

Executive Summary	i
1. Introduction	1
1.1 Purpose of Project	2
1.2 Scope of Project	2
1.3 Outline a Work Performed	3
2. Literature Review	6
2.1 Relationship Between Maintenance and Safety	6
2.2 Scope of Maintenance	10
2.3 Maintenance and Human Performance	17
2.4 Roles of Programmatic Performance Indicators	18
2.5 Identification of Previously Suggested Candidate Indicators	25
2.6 Summary of the Results of the Literature Review	28
3. Frameworks of Examining Maintenance Programmatic PIs	29
3.1 Purpose of Frameworks	29
3.2 Conceptual Relationship of Maintenance to Safety	29
3.3 Maintenance Indicators Framework	30
4. Process for Identifying Candidate Performance PIs	35
4.1 Identification of Maintenance Indicators	36
4.2 Qualitative Evaluation of Maintenance Indicators	36
4.3 Data Collection	48
4.4 Validation Process	50
5. Results of Preliminary Evaluations	58
5.1 Detailed Results	58
5.2 Summary of Results	60
6. Further Analysis of Promising Indicators	62
6.1 Inadvertent ESF Actuations from Test & Maintenance Errors	62
6.2 Thermal Performance Indicator	88
7. Summary and Conclusion	111
8. Bibliography	113

List of Figures

Figure ES-1	iii
Figure 1.1, Overall Process of Analysis	4
Figure 2.1, Relationship Between Indicators and Safety	
Figure 2.2, Conceptual Relationship of Maintenance to Safety	9
Figure 2.3, Maintenance Inspection Tree	13
Figure 2.4, Process Model	16
Figure 2.5, Hierarchy of Performance Indicators	19
Figure 2.6, Cause Code Hierarchy	
Figure 2.7, Conceptual Framework for Causes of Events: Overview of Casual Chain	24
Figure 4.1, Validation Process Diagram	51
Figure 4.2, Preliminary Event Data	52
Figure 4.3, Smoothed Event Data	54
Figure 6.1, ESF Actuations, Plant A	68
Figure 6.2, ESF Actuations, Plant B	69
Figure 6.3, ESF Actuations, Plant E	73
Figure 6.4, ESF Actuations, Plant F	74
Figure 6.5, ESF Actuations, Plant D	78
Figure 6.6, ESF Actuations, Plant C	81
Figure 6.7, DPL Data, Plant S, 1985	94
Figure 6.8, DPL Data Plus Causes, Plant S, 1985	95
Figure 6.9, Power Density Spectrum, Plant S, 1985	96
Figure 6.10, Combined Frequency Plot, Plant S, 1986	99
Figure 6.11, Combined Frequency Plot, Plant S, 1987	100
Figure 6.12, DPL Plot, Plant X, 1985	104
Figure 6.13, DPL Plot, Plant I, 1986	105

List of Figures (Continued)

Figure 6.14, Power Density Spectrum, Plant X, 1985	106
Figure 6.15, Power Density Spectrum, Plant T, 1986	107
Figure 6.16, Combined Frequency Plot, Plant X, 1985	108
Figure 6.17, Combined Frequency Plot, Plant T, 1986	109

List of Tables

Table 2.1, Candidate Maintenance Indicators	26
Table 4.1, Candidate Maintenance Indicators (Composite List)	37
Table 4.2, Evaluation & Maintenance PIs	40
Table 6.1, ESF Events, Plant A	70
Table 6.2, ESF Events, Plant B	71
Table 6.3, ESF Events, Plant E	75
Table 6.4, ESF Events, Plant F	76
Table 6.5, ESF Events, Plant I	79
Table 6.6, ESF Events, Plant C	82
Table 6.7, Plant S, 1985	97
Table 6.8, Data for Cosine Plots of Yearly Frequencies	101

ACKNOWLEDGEMENT

The authors wish to acknowledge the assistance and support of Erasmia Lois, the NRC project officer in the development of the methods described in this program.

Additional assistance has been provided in the form of informal discussions with colleagues in many parts of SAIC, particularly Ernest Lofgren, Gary DeMoss, Martin Stutzke, and James Davis. Discussions with NRC personnel have provided significant inputs with Carl Johnson (RES), Thomas Novak and Mark Williams (AEOD) concerning the nature of indicators, and with John Jankovich (NRR), concerning NRC's maintenance inspection process and development of the maintenance inspection tree.

Numerous people have provided informal comments on this work. Of particular note were participants at an informal panel during the 1988 IEEE International Conference on Human Factors and Nuclear Power. These included: J. Rasmussen (Denmark), J. Senders (USA), E. Swaton (IAEA), and Shigeo Kojima (Japan). In addition, a meeting was held with senior maintenance and QC personnel of Pan American World Airways, Inc., to understand how commercial aviation maintenance uses indicators; the assistance of Ray Valeika and his staff is acknowledged.

Special thanks are extended to Mary Moore for the typing, assembly, and final production of this report; Zun-Yan Wang performed code analysis and generated many of the tables in Volume 2.

DEVELOPMENT OF PROGRAMMATIC PERFORMANCE INDICATORS ASSOCIATED WITH MAINTENANCE

EXECUTIVE SUMMARY

This report presents the technical approach, results and conclusions of a study to investigate programmatic performance indicators associated with maintenance at commercial nuclear power plants.

Management and control of any operational enterprise, from landing a spacecraft to developing investment portfolios, requires the development and tracking of indicators of the enterprise's performance throughout time. Properly developed indicators can serve as signals to management to allow them to take the appropriate control action. The need for a properly developed set of indicators of performance is no less required to track the safety performance of nuclear power generating station. Because of this, in 1986, US Nuclear Regulatory Commission (NRC) adopted a program to plan to use performance indicators as a source of information for NRC to use in "making timely regulatory decisions about the performance of nuclear power plants" (SECY-86-317). In that paper, the term "performance indicators" reflects a set of data that should have correlation with individual plant safety performance.

Two kinds of indicators were envisioned: (1) direct indicators of current plant performance (such as safety system failures) and (2) indirect, or programmatic indicators. Programmatic performance indicators focus specifically on the programmatic aspects of plant activities; these are such things as planning, prioritization, resource allocation, etc., and their impact on safety. This type of indicator is intended to monitor human performance issues rather than single hardware measures. However it was recognized that human performance may be most visibly revealed by hardware performance.

The first phase of this work identified seventy-eight indicators associated with maintenance as possible candidate indicators. Use of scoping frameworks and the development of an associated preference scheme led to the selection of a shortlist of nine maintenance indicators. Data collection activities were undertaken at eight plant sites; these visits included participation in joint NRC-contractor data collection visits and associated analysis, and were to obtain data on plant parameters related to the identified indicators. These data were used as inputs to a validation process, which evaluated the

ability of the indicators to anticipate changes in plant safety performance. This evaluation narrowed the number of indicators to two: thermal performance (gross heat rate), and the number of inadvertent ESF actuations due to test and maintenance errors. The first phase of the program concluded that a follow-up evaluation of these two indicators should be performed in order to monitor their behavior at a small number of plants.

Analyses were performed to compare qualitatively the relationship of unintended ESF actuations from test and maintenance errors with the narrative descriptions of maintenance-related program areas in NRC's Systematic Assessment of Licensee Performance (SALP). These qualitative analyses, performed for 10 plants over a 3 1/2 year period, have confirmed a strong link between the programmatic aspects of maintenance (management involvement, effective problem solving, timely scheduling, and so on) and the frequencies of test or maintenance-error-caused ESF actuations. In Figure ES-1, the frequencies (quarterly) of these events are plotted for 6 plants. Plants A & B, which are Units 1 & 2 at the same site, were identified by SALP as having a very ineffective program of maintenance, with poor root-cause analysis and no equipment-performance tracking programs. Plants E & F, again two units at one site, were rated poorly in the beginning and end of the period, with some improvements in the middle period. The ESF actuation frequencies follow this same pattern. Plant C was described as having strong management involvement and excellent control of maintenance activities. Correspondingly, Figure ES-1 shows no inadvertent ESF actuations for Plant C. Plant I had several periods of extended shutdown; three events occurred during one of the brief periods of operation. Plant I had previously a history of major weaknesses in maintenance but many were reported fixed during the lengthy outages.

For the remaining four plants, no events occurred in the period of interest.

The specific ESF-related indicator of interest here is to be distinguished from two other ESF-related indicators which are also under investigation by NRC. The first of these is the total frequency of ESF actuations for all causes. Current monitoring results suggest that this indicator is not effective and NRC is considering to drop it. However the ineffective nature of this indicator may be due to its all encompassing nature. It should be recognized that most inadvertent ESF actuations occur when the reactor is shutdown. At this time the reactor systems may often be found in unusual configurations.

Figure ES-1

In a state such as this, which may be significantly different from the normally stable operating plant state, it should not be surprising that the number of ESF actuations would be large. In this instance ESF actuation may be easily initiated via a variety of causes which would not be active during the normal operating plant state. For this reason the specific ESF-related indicator proposed here includes only inadvertent ESF actuations when the reactor is at power, and considers only those instances where test and maintenance errors were considered to be a contributing factor. Restricting the specific ESF-related indicator only to this subset is expected to tie it more directly to both maintenance and safety.

The second ESF-related NRC study pertains to the ongoing investigation into the patterns, causes, and corrective actions reported in all LERs. Since advertent ESF actuations are reportable under current LER rules they would be expected to be included in this investigation. However the ongoing analysis does not distinguish the types of events reported in the LERs. Omitted surveillances are included with scrams, safety system failures as well as ESF actuations. This aggregation may tend to weaken the signal of an individual indicator such as ESF actuations. In addition, these analyses do not provide any discussion of why these patterns may act as an indicator nor of their implications for maintenance programs.

The investigation of frequencies of the specific ESF actuation indicator in this study also suggests a potential relationship a plant's approach to taking corrective actions and its associated qualitative SALP assessments. The text given in the LERs reviewed for this study included references to "discipline" or "counselling" as corrective actions taken at plants A and B with little or no more detailed root-cause analyses. The use of discipline or counselling as a corrective measure can be considered to be evidence of plant management engaging "denial" actions---the avoidance of the responsibility for correcting the cause of or denying the existence of safety problems. This kind of behavior has been associated with what has been termed "pathological" organizations in the social science literature*. These organization types have tended to be associated with less safe operation.

* See, for example, Organizational and Inter-Organizational Thought, R. Westrum, World Bank Workshop on Risk Management & Safety Control, Washington, D.C., October 1988.

It is believed that for this ESF-related indicator, results of this work are sufficient to allow NRC to consider its adoption for trial use in a significant number of plants.

Because of data limitations and insensitivity of the indicator originally chosen as a measure of thermal performance, the specific measure was changed to Daily Power Loss (DPL). DPL is the daily average loss of generation compared with the maximum dependable capacity. For any plant that operates base-loaded, a reduction in output generation is usually an indicator of equipment failures within balance-of-plant systems, although other causes include technical specifications-caused shutdowns, other scrams, and fuel-conservation measures. NRC's definition of maintenance includes all functions to ensure and preserve the availability of plant systems. Because of the broad based nature of this definition almost all causes of plant outage (except refueling and human caused scrams) are related to "maintenance". Hence the measure has the potential to treat the effectiveness of maintenance in an integrated manner. In other words, the measure of DPL can be envisioned as a signal of the condition of the entire plant maintenance process.

The data that permit the calculation of this measure are already reported to NRC on a monthly basis through the docketed plant operational reports so that its evaluation is readily supported by an easily accessible data set. In addition, since it is a measure of the plant's overall output (as was gross heat rate) (being based on a direct measure of the megawatts electrical sent out daily) it is essentially free from interpretation or manipulation.

The investigation of the usefulness of the DPL signal presented by each plant as indicator of plant performance required the application of signal processing theory. This theory was used to develop techniques that allow the frequency and duration information content of the DPL signal output to be examined. In this way it was intended that plants might be discriminated according to the frequency of power losses and the duration of losses that result from equipment failure. This approach is consistent with the principle that an effectively maintained plant (using the NRC broad definition of maintenance) should exhibit very few power-loss events and these events only of short duration.

The work documented here indicates that the signal-processing technique appears to distinguish between plants having more frequent losses of power, and between plants having longer periods of loss. However, more work is needed to confirm this capability

and to develop a process for presenting results in form which is more compatible with the requirements of a plant performance indicator. In addition, further work is recommended to allow the relationship between the measure and the effectiveness of the plant maintenance, *per se* to be better understood.

The additional work recommended for the further investigation of the DPL signal should track DPL data for periods between major outages (especially refueling outages). Several alternative methods of analysis should be considered. These would include continuously monitoring performance on a 3- and 6- month window, and focusing on the frequency of small loss events. These would both have the potential to express a DPL measure as an indicator. Finally, additional evaluations against other indicators and qualitative maintenance criteria are required to determine the effectiveness of the indications given.

In addition to the specific analytical results mentioned the effort involved in project has led to the development of perspective among the project analysts which views a particular programmatic performance indicators as one of the symptoms of the existence of a extant human performance issue. In most cases it is the symptoms which provide indication that a safety problem may exist, but these symptoms are only manifestation of an underlying human performance issue. The terms "type" and "token" have been used in the human performance literature to distinguish between the underlying issue and its symptoms.⁺ Specifically "types" are only revealed through their corresponding "tokens". In the current instance of the ESF actuation indicator, a continuing high frequency of occurrence of this indicator is believed to be a token of the failure to identify and remove the root causes of failures (i.e. the type). However the corresponding safety impact of this type is primarily manifested in other tokens such as continuing high safety equipment failure rates which are inconsistent with in properly addressed failure causes. In other words while a selected programmatic performance indicator might not in itself be directly important to safety, it may be a token of some type that has other tokens and these other tokens are important to safety.

In this work, we consider the indicator-token simply provides a convenient mechanism to allow data to be easily gathered as an indirect indicator of the safety-measure-token

⁺ See, for example, "Human Error," J. Reason, Cambridge University Press, New York, (In Press)

which is less accessible. The aim is to select an indicator-token (here ESF actuations) so that it reasonably tracks the safety-token (safety equipment unavailability), where both emanate from the same type (ineffective root-cause analysis).

DEVELOPMENT OF PROGRAMMATIC PERFORMANCE INDICATORS ASSOCIATED WITH MAINTENANCE

1. Introduction

The US Nuclear Regulatory Commission (NRC) is sponsoring several research programs to develop nuclear power plant performance indicators. Performance indicators are intended to provide a quantitative means of monitoring the effectiveness of activities at commercial nuclear power plants as they influence safety. Several overall performance indicators have been proposed by the NRC and by the Institute for Nuclear Power Operations (INPO) based on plant performance trend analyses; such indicators primarily address equipment performance (i.e., they are based on scrams frequencies, system availabilities, component outage durations, etc.). Trends observed in these performance indicators along with other information (e.g., SALP ratings, plant inspection reports) are reviewed by senior NRC managers, to help focus their attention on areas of plant operation which are most important to safety.

Based on recommendations proposed by INPO, many plants have developed their own indicators for use by plant and corporate management (NUREG/CP-0099, 1988). Often these include trending overall performance as measured by the number of scrams and average plant availability, and trending indicators associated with individual plant functions such as maintenance work request (MWR) backlogs, the number of contractor staff on site, and the number of "red phone" calls (i.e., the number of operational events reported to the NRC). These indicators are often trended through time for the plant, and may be compared with industry averages where data are available.

During the overall NRC performance indicator program, an NRC Interoffice Task Group formed to address performance indicators, recognized that additional indicators (based on the performance of functional units within the plants, such as maintenance, training, operations, and quality assurance) could provide additional guidance to NRC managers in their decision-making, (SECY-86-317, 1986). These indicators, referred to as programmatic performance indicators (programmatic PIs), are intended to measure the effectiveness of plant programs that secure safety. As such they are not tied to any single type of observable event such as the number of scrams. They must reflect human

performance rather than equipment performance, although it is recognized that human performance is often revealed by equipment performance.

This report presents the principal technical accomplishments of a project sponsored by the NRC/RES to identify specific programmatic PIs related to nuclear plant maintenance. A maintenance-related programmatic PI has the potential for providing early indication of changes in plant maintenance activities (taken collectively) that can eventually affect plant safety. In this context, "safety" implies both the maintenance of a low frequency of plant upset conditions and the concomitant achievement of the functional requirements for ensuring an adequate state of readiness for responding to upset conditions.

1.1 Purpose of Project

The purpose of this project is three-fold:

- (i) To provide a context for describing the ways in which plant maintenance activities influence safety,
- (ii) To develop a structured process by which quantitative indicators of plant maintenance activities can be identified (or generated), and
- (iii) To characterize the ability of specific (maintenance-related) programmatic performance indicators to anticipate changes in the safety of nuclear power plants.

1.2 Scope of Project

The scope of this study has been to:

- (1) Perform a survey of relevant work by the NRC, the nuclear industry and other industries;
- (2) Identify and candidate programmatic performance indicators;
- (3) Develop alternative indicator validation techniques and criteria;

- (4) Collect data associated with the indicators and the validation methods; and
- (5) Evaluate and make suggestions concerning use of programmatic performance indicators.

This report presents the accomplishments of this scope of work.

In this study, maintenance is defined as "the aggregate of those functions required to preserve and promptly restore safety, reliability and availability of plant structures, systems and components." This definition has been taken from the NRC's previously published proposed rulemaking on maintenance (NUREG/CP-0099). It is noted that this definition is broader than the usually conceived scope of maintenance by including functions like: equipment trending, parts management, procedures, and communications.

1.3 Outline of Work Performed

The tasks completed in the present evaluation of maintenance-related programmatic performance indicators are summarized in Figure 1.1. These include:

- . A literature review to provide a basis for the development of a framework for the analysis, and to identify candidate indicators previously suggested;
- . Creation of structured frameworks to define the relationship of maintenance activities to plant safety and to assist in the identification and selection of candidate indicators;
- . Identification of candidate indicators and subsequent selection of a reduced set of indicators for further detailed analyses;
- . Collection of data from operating nuclear power plants;
- . Validation of candidate programmatic PIs via thorough statistical analyses of plant data; and
- . Engineering analyses of maintenance-related programmatic PIs that are shown (statistically) to characterize trends in overall plant performance.

Figure 1.1
Overall Process of Analysis

This report is structured to follow the flow of tasks shown in Figure 1.1. Chapter 2 summarizes the results of the literature review. Chapter 3 discusses the creation of frameworks and their roles in the selection of indicators. Chapter 4 describes the preliminary evaluation and shortlisting of indicators. Chapter 5 describes the evaluation of indicators by statistical evaluations against overall measures of safety, and Chapter 6 describes the more detailed analyses that seek to provide causal explanations of indicator-safety relationships.

This work identified and reviewed seventy-eight candidate indicators of maintenance performance. Use of the frameworks and the development of an associated preference scheme led to the selection of a shortlist of nine candidate maintenance indicators. Data were subsequently collected from eight plant sites to support detailed analysis of these nine indicators; the task included participation in joint NRC-contractor data collection visits and associated analysis. Analysis of these data suggested two maintenance indicators which were considered to have had a statistically significant potential for characterizing trends in maintenance effectiveness. These indicators are:

- . The number of inadvertent emergency safety feature (ESF) actuations due to test and maintenance errors, and
- . The plant gross heat rate.

These two indicators were subsequently evaluated (via engineering analysis of plant-specific operating experience) for their abilities to monitor substantive attributes of maintenance performance.

2. Literature Review

The literature review was performed to: (1) identify the potential relationship between maintenance and safety, (2) develop a working definition of maintenance, (3) identify the roles of performance indicators, and (4) identify potential indicators for analysis. The literature reviewed includes documentation from other NRC indicator studies, work by other organizations such as the Institute of Nuclear Power Operation (INPO) and the Electric Power Research Institute (EPRI), and individual plants. In addition, informal discussions were held with other organizations including the National Aeronautics and Space Administration (NASA), a civil aviation carrier (Pan American World Airways), and representatives of overseas commercial nuclear programs.

2.1 Relationship Between Maintenance and Safety

Maintenance has been shown in probabilistic risk assessments (PRAs) and, more importantly, by its involvement in safety-significant historical events to be an important factor in safety. In a PRA perspective, maintenance has three influences on risk. These are: (1) identifying and restoring failed safety equipment to an available state; (2) rendering standby equipment unavailable during testing or repair; and (3) rendering standby equipment unavailable because of inadequate repairs or testing. For example, in NUREG/CR-3541 (Vesely, *et al.*, 1983), the sensitivity of risk to test and maintenance activities at several power plants was analyzed. In that study, very large increases in core-damage frequencies results from hypothetically extreme deficiencies in test or maintenance for certain Emergency Core Cooling Systems (ECCS); for example, core-damage frequency increases by factors of 100 to 8,000 were calculated for the low pressure service water system at Oconee. Comparably large sensitivities were obtained for Calvert Cliffs and Grand Gulf.

In developing the NRC's plans for maintenance performance indicators, the agency's Interoffice Task Group on Performance Indicators developed a logic model portraying the relationship between safety and maintenance indicators (SECY-86-317, 1986). This logic model, shown in Figure 2.1, identifies safety in terms of three basic elements: low frequency of transient, high availability safety systems, and inherent design features and low potential for cognitive errors. Of these three, the first two are considered to be most directly influenced by maintenance. For example, a low frequency of scrams can result

Figure 2.1, Relationship Between Indicators and Safety
(From SECY-86-317)

from "well-maintained" (that is, having a low failure rate) balance-of-plant systems (feed-water and turbine-generator systems), and from a low frequency of maintenance errors that cause scrams. Similarly, as discussed in the PRA Procedures Guide (NUREG/CR-2300, 1983), a high availability of safety systems results from a high train availability and a low potential for common-cause failures. A high train availability results from components having low failure rates, low unavailabilities from maintenance activities, and low failure rates from failures to be restored after testing or maintenance. Common-cause failure mechanisms include maintenance-related causes, such as erroneous procedures, faulty replacement parts, and errors during concurrent testing.

Figure 2.2 is based on Figure 2.1, with modifications to express more explicitly the interfaces between maintenance and safety. The *scram frequency* is predominantly determined by the equipment failure rates for the balance-of-plant equipment, and by errors in testing, maintenance and operations. In turn, the balance-of-plant equipment failure rate is a result of maintenance effectiveness. The *availability of safety systems*, which are normally in a standby mode, results from: (1) the systems being tested for unrevealed failures and those failures being corrected; (2) from their being correctly aligned as part of a configuration-management process; (3) from preventive maintenance being performed that reduces component failure rates; and (4) from design modifications being made to improve the operability, maintainability or availability of systems. In expanding this logic model, the element concerned with *operator comprehension and design deficiencies* has been changed. The operators' capabilities are expressed in terms of the plant's equipment availabilities that may be used in incident management, such as instrumentation, controls, and other equipment not formally considered as safety systems. The element associated with design deficiencies has been modified to reflect the evolving nature of systems designs once a plant is in operation, and is now termed *design modifications*, part of the safety systems availability. It is recognized that there are both light and dark sides to design modifications. These can eliminate weaknesses in the original plant design; they may also introduce new failure mechanisms and unforeseen systems interactions.

Figure 2.2 expresses the links between maintenance and safety used as a basis for analysis in this study.

Figure 2.2, Conceptual Relationship of Maintenance to Safety

2.2 Scope of Maintenance

The NRC adopted a broad interpretation of maintenance in its proposed rulemaking on maintenance. In this activity, the NRC effectively identified maintenance to be all activities that influence the "health and well-being" of equipment. Specifically, maintenance is defined as:

"the aggregate of those functions required to preserve and promptly restore safety, reliability and availability of plant structures, systems, and components. Maintenance includes not only activities traditionally associated with identifying and correcting actual or potential degraded conditions, i.e., repair surveillance, diagnostic examinations, and preventive measure; but extends to all supporting functions for the conduct of test activities. These activities and functions are listed below:

- Technology in the areas of
 - Corrective maintenance,
 - Preventive maintenance,
 - Predictive maintenance,
 - Surveillance;
- Engineering support and plant modifications;
- Quality assurance and quality control;
- Equipment history and trending;
- Maintenance records;
- Management of parts, tools, and facilities;
- Procedures;
- Post-maintenance testing and return-to-service activities;
- Measures of overall program effectiveness;
- Maintenance management and organization in the areas of
 - Planning,
 - Scheduling,
 - Staffing,
 - Shift coverage,
 - Resource allocation;

- Control of contracted maintenance services;
- Radiological exposure control (ALARA);
- Personnel qualifications and training;
- Internal communications between the maintenance organization and plant operations and support groups;
- Communications between plant and corporate management and the maintenance organization;
- Maintenance recommendations or requirements of individual vendors."

This definition is based on earlier NRC and IAEA work, including that presented in the report of the NRC's Maintenance and Surveillance Program (NUREG-1212, 1986) and the International Atomic Energy Agency's Safety Guide on Maintenance of Nuclear Power Plants (IAEA, 1982).

Key factors in this definition are that maintenance is a *process* having the purpose of *preserving the safety, reliability, and availability* of plant structures, systems and components. This demands activities beyond those of the plant maintenance departments, including some activities associated with the engineering support, training, quality assurance, operations, purchasing department, and management.

In addition to this definition of maintenance, the NRC developed an analytical tool, called the *Maintenance Inspection Tree*, to provide a formal framework for the evaluation of maintenance programs during plant inspections; it is shown in Figure 2.3. This tree identifies three basic aspects of maintenance that are to be reviewed to establish whether a plant has established an effective plant maintenance process. These factors are: (1) management support for maintenance, (2) maintenance implementation, and (3) overall performance related to maintenance. Management support is assessed in terms of:

- Management commitment & involvement
 - application of industry initiatives
 - management vigor & example
- Management organization & administration
 - program coverage
 - policies, goals & objectives

- resources
- requirements
- performance measurements
- document controls
- formal decision processes

- Technical support
 - communication channels
 - engineering support
 - use of PRA information
 - roles of quality control
 - integration of radiation control
 - safety reviews of maintenance
 - integration of regulatory documents.

Maintenance implementation is comprised of four factors:

- Work control
 - work order controls
 - equipment records & histories
 - job planning
 - work prioritization
 - scheduling
 - backlog controls
 - procedures
 - post-maintenance testing
 - review of completed work documents
 - review of work in progress

- Maintenance organization
 - control of maintenance activities
 - control of contracted activities
 - deficiency identification & controls
 - maintenance trending
 - support interfaces

TREE INITIATORS

1. REASON FOR INITIATION
2. PLAN INITIATION
3. SCOPE OF INITIATION (LOCAL, PLANT, OR SYSTEM-WIDE)
4. INITIATION ACTION PLAN
5. INITIATION OF PLANT ACTIONS

MAINTENANCE INSPECTION TREE

Figure 2.3, Maintenance Inspection Tree

- Maintenance facilities, equipment & materials
 - facilities controls
 - materials controls
 - tools & equipment controls
 - controls of calibration & test equipment
 - plant condition inspections
- Personnel controls
 - staffing controls
 - personnel training
 - test & qualification process
 - current status assessment.

Overall plant performance is assessed on the basis of equipment reliability, availability and operability, and radiation exposure.

A similar broad interpretation has been taken by the nuclear industry. In the presentations made by representatives of NUMARC, EPRI and INPO to the NRC's Workshop on Rulemaking on Maintenance of Nuclear Power Plants (NUREG/CP-0099, also the Workshop Transcript, 1988), attention was focused on the broad and systematic perspective that must be taken in considering maintenance. For example, in developing INPO's Guidelines for the Conduct of Maintenance at Nuclear Power Plants (INPO-85-038, as reported in the Workshop Transcript) it was recognized that maintenance activities were much broader than the traditional mechanical, electrical and Instrumentation and Control (I&C) department repair functions. INPO recognized sixteen activities as comprising maintenance:

- Maintenance department organization & administration;
- Training & qualification of maintenance personnel;
- Maintenance facilities, equipment & tools;
- Types of maintenance;
- Maintenance procedures;
- Planning, scheduling & coordination of maintenance;
- Control of maintenance activities;
- Post-maintenance testing;
- Material receipt, inspection, handling, storage, retrieval, and issuance;

- Control & calibration of measuring and test equipment;
- Maintenance tools & equipment control;
- Station material condition inspection;
- Management involvement;
- Maintenance history;
- Analysis of maintenance problems.

Figure 2.4, developed in this project, portrays a process model of the functional activities in a nuclear plant that are considered to be part of maintenance based on the above definitions. The principal types of interactions and influences between these functions are also shown, based on discussions with the NRC and project staff with experience in power plant maintenance. These functions and their principal interactions are described in Appendix A of this report.

This model shows some important characteristics about the nature of maintenance at power plants. First, maintenance is a process that has several "feedback paths" from the results of maintenance (the performance of the physical plant systems and the overall plant outputs) to the people at the "front line" of the maintenance (technicians and supervisors in the maintenance group). These feedback paths permit deficiencies in maintenance effectiveness to be corrected, either within the maintenance group or involving training, engineering support, and so on. Effectiveness in feedback and self-learning are important characteristics whose values are related to the effectiveness of organizations (for example, see Peters, 1987, Rochlin, 1987). In addition, there are multiple influences on most of the functions identified in the process model shown in Figure 2.4. In other words, there are no single-path influences affecting the performance of any of the functions like the maintenance and the training groups. The existence of these multiple influences make it virtually impossible to monitor the "strengths of influence" on any one activity, such as the timeliness of the feedback of performance from plant behavior to the maintenance group and use the results of this monitoring as an indicator of overall plant performance. This is because to use this internal indicator as a measure of overall plant performance would require ensuring that related plant variables are not manipulated so as to produce a "good" value of this measured internal indicator. Alternatively all the interrelated paths could be monitored but in this case it is clear that multiple paths would require extensive monitoring of internal data at a plant.

Figure 2.4, Process Model

2.3 Maintenance and Human Performance

Maintenance is a process that is strongly influenced by human performance; almost all maintenance tasks involve "hands-on" human actions. Several studies have been performed to identify human-factors issues of importance to maintenance. These studies range from general reviews, such as the EPRI documents NP-1567, "Human Factors Review of Power Plant Maintainability"; AF-1041, "The Role of Personnel Errors in Power Plant Equipment Reliability"; and NP-4350, "Human Engineering Guidelines for Maintainability" to more specifically focused studies, such as NUREG/CR-3817, "Development, Use and Control of Maintenance Procedures in Nuclear Power Plants: Problems and Recommendations", and NUREG/CR-4600, "Human Factors Study Conducted in Conjunction with a Mini-Round Robin Assessment of Ultrasonic Technician Performance". Numerous military studies have been published by the Air Force, Army, and Navy.

However, all of these studies have dealt with relatively narrow human engineering issues such as accessibility, tool design, etc., or else have dealt with training or procedure-writing issues. While these issues are important to the improvement of maintenance-technician performance, little research has been performed to analyze the *programmatic* human activities associated with maintenance. Most of the studies that have been performed were performed in the context of Maintenance Management Information Systems (MMIS). Some studies that had this thrust are: EPRI NP-3416, "A Guide for Developing Preventive Maintenance Programs in Electric Power Plants" (in the electric utility industry), and such publications as "Production Oriented Maintenance Organization:..." (for the Air Force) and a number of papers at the National Bureau of Standards Maintenance Symposium (1983) and the ANS International Meeting on Nuclear Power Plant Maintenance (1986). These studies are not directly relevant to the issues of programmatic indicators because they focus on establishing feedback loops within the process itself. That is, these methods and techniques are certainly useful in tracking and trending maintenance performance from the point of view of taking immediate corrective actions within the administrative structure of the organization, but provide little or no help in evaluating maintenance performance from outside the organization.

Programmatic activities, on the other hand, include the evaluation of *training* of maintenance supervision personnel, the interface characteristics of organization

administration with maintenance management, effects of *communication* characteristics for the interface of maintenance and other plant personnel, and *integration* of maintenance into broader goals and objectives of the organization in question. These programmatic activities are visible from outside the organization and may be usable to gauge the maintenance health of the organization over a period of time.

2.4 Roles of Programmatic Performance Indicators

Programmatic indicators are intended to measure the effectiveness of plant programs that secure safety (Lois, *et al*; 1989). They are *program effectiveness* indicators, and, as such, are not directly tied to a single type of observable events, such as numbers of scrams. Moreover they must reflect human performance rather than equipment performance *per se*, although human performance is often revealed from equipment performance.

Programmatic performance indicators are one of several different kinds of indicators. NRC's Interoffice Task Group on Performance Indicators has developed conceptual diagrams showing the roles of different types of performance indicators, including programmatic performance indicators. Figure 2.5 depicts a model representation of different indicator types from the perspective of this licensee and INPO, and NRC. The figure has been taken from the report of the Interoffice Task Group on Performance Indicators (SECY-86-317, 1986), and shows the different domains of these indicators.

2.4.1. Overall Plant Safety Performance Indicators

There are indicators associated with overall plant safety performance; this is the set of indicators shown in Figure 2.5 going from "plant performance" to "NRC". The following indicators have been adopted by the NRC as such measures of plant safety:

No. of automatic scrams

- all power levels
- greater than 15% power
- at or less than 15% power

No. of safety system actuations

No. of significant events

No. of safety system failures

Figure 2.5, Hierarchy of Performance Indicators
(Taken from SECY-86-317)

Forced outage rate
No. of equipment forced outages
Collective radiation exposure.

The above set of indicators are defined in SECY-86-317 (1986). Guidance on their use was provided in a memo to the NRC staff by NRC's Executive Director of Operations (EDO Announcement No. 30, 2/5/88). This announcement identified this indicator set as one of several sources of data used by senior NRC management in decision-making regarding plant specific regulatory programs..."no regulatory action should be taken on the basis of performance indicator program results alone." They are "intended ... to monitor trends in overall performance for a given plant (and) should be viewed as a set". These indicators are used as a supplementary source of information to the NRC's Systematic Assessment of Licensee Performance (SALP) program, and the results of routine and special inspections by regional and headquarters staff. [In this study, these overall indicators are used as a set of reference measures for comparing the performance of programmatic indicators developed in this project.]

2.4.2 Risk-Based Performance Indicators

At a lower level of monitoring, the NRC is developing indicators associated with safety-systems availability (NUREG/CR-5323, 1989). These indicators, called risk-based indicators (because of their association with the PRA perspective which relates safety-system unavailability to risk), are calculated periodically from the unavailability of components in designated safety systems; unavailability data are obtained from equipment databases such as the Nuclear Plant Reliability Data System (NPRDS). In Figure 2.5, these indicators also flow from "plant performance" to "NRC", but are at a finer level of detail (i.e., at a lower level in the hierarchy in Figure 2.2) than the overall plant indicators, and are specifically intended to monitor maintenance performance. They are also intended to provide earlier indicators of safety trends than the overall indicators. This potential exists because safety system component unavailability events are conditioning events for other more direct measures such as the number of significant events and the number of safety system failures. Conditioning events or precursor events usually do not result in more significant events because of plant design or operating features such as redundancy. For this reason the frequency of increases in these unavailabilities is much greater than the more serious events to which they could potentially contribute. However, since an increase in the frequency of these frequency

conditioning events does signal a corresponding increase in the potential for the occurrence of the more serious events, these component unavailability events can provide early indication of safety trends. As an additional focus on the unavailabilities of particular equipment, they also provide some level of explanation of changes in overall safety measures. The risk-based indicators program is being performed concurrently with the work described in this report.

2.4.3 Failure Rate Trending

Two programs directly associated with maintenance performance indicators are under current trial application by the NRC. These are the trending of equipment failure rates derived from NPRDS data, and the analysis of cause and correction codes associated with Licensee Event Reports (LERs). The NPRDS trending study is most recently reported in AEOD/S804B, Application of the NPRDS for Maintenance Effectiveness Monitoring [Predecisional] (1989). In that report, AEOD used NPRDS data as a basis for monitoring maintenance effectiveness and for controlling failure rates; specifically this indicator system was intended to identify when failure rates associated with a specific system exceeded a predetermined level. The study reports the monitoring of 39 types of components in five systems (both primary and balance-of-plant systems) for 28 Boiling Water Reactor (BWR) plants over a period of 39 months. The components monitored were selected as being the "outage dominating equipment" (ODE); that is, those components that are the major sources of plant shutdowns. The use of failure rates as an indicator is based on the evidence that maintenance performance was reportedly associated with 77% of the failure events analyzed, and that the indicator correlated reasonably well with other information regarding maintenance problems derived from LERs.

These indicators can be considered as intermediate indicators of plant safety performance rather than either direct, or top-level, indicators of safety as is the case for programmatic indicators. As portrayed in Figure 2.2, the equipment reliability indicators are related to the mid levels of the diagram relating maintenance and safety. However; since this indicator was based purely on the occurrence frequency of failure events it provides no level of explanation for the issues affecting equipment reliability.

2.4.4 Cause-Code and Correction Code Trending

The concept of trending cause codes is reported in NUREG/CR-5241, Development of Programmatic Performance Indicators (1988). In that study, LERs reported over a period of three years (1984-86) were reviewed and the causes were classified into eight categories:

- Operator errors
- Other personnel errors
- Maintenance problems
- Design/installation problems
- Procedure errors
- Administrative control problems
- Equipment failures
- Unknown.

In the study the number of events, which were categorized by each of these cause codes, was correlated to the SALP scores, and the overall plant safety indicators for the corresponding plants. It was concluded on the basis of this condition that these cause-code scores could provide a basis for predicting future performance. On the basis of the study findings this cause-code score approach has been adopted by the NRC for trial use. The trial use application includes the additional consideration of the corrective actions reported in the LERs, such as training, counselling, procedure modifications, etc. The trial application of the cause-code score approach is continuing in parallel with the work described in this report.

Both the failure event and cause-code score trending indicators are consistent with the programmatic performance indicator level since they provide information about some causes of maintenance failures, such as procedure errors and administration errors. In the work reported in NUREG/CR-5241, a hierarchical representation of the causes of reportable events was constructed; this is shown in Figure 2.6. The logic shown in that figure is very similar to Figure 2.2 as it relates to safety systems availability. A broader framework showing the interaction of programmatic issues (training, staffing, administration, etc.) that lead to reportable events was also developed in the study reported in the reference NUREG. This framework is shown as Figure 2.7. However,

**Figure 2.6, Cause Code Hierarchy
(Taken from NUREG/CR-5241)**

Figure 2.7, Conceptual Framework for Causes of Events:
Overview of Casual Chain (Taken from NUREG/CR-5241)

the report does not present hypotheses as to how the presented indicators might be represented in terms of their relationship to any underlying causes.

2.5 Identification of Previously Suggested Candidate Indicators

Prior to the initiation of this study several efforts directed at the development of candidate indicators had been underway both at the NRC and elsewhere in the nuclear industry. Part of the initial effort of this study was therefore directed at identifying these pre-existing sets of candidate indicators.

Three sources were principally used to identify candidate indicators. These were: (1) the evaluation of indicators performed by Pacific Northwest Laboratories (PNL) reported in NUREG/CR-5241; (2) documentation associated with INPO's performance indicators programs; and (3) discussions held with NRC's maintenance assessment staff. In addition, discussions at the NRC Workshop on the proposed maintenance rulemaking (summarized in NUREG/CP-0099, and reported in the published transcript) were reviewed for references to potential indicators. Theses and other sources are listed in Section 8.

The literature search conducted was augmented by discussions with NRC staff, NASA, a commercial aviation carrier (Pan Am), and utility personnel. In addition, an informal discussion was held with representatives of foreign nuclear power programs during the 1988 IEEE International Conference on Human Factors and Nuclear Safety.

A total of 75 candidate indicators were identified as a result of this review; these are listed in Table 2.1. All had previously been identified as part of either NRC or the industry-sponsored programs. For example, as part of these industry studies, INPO has proposed a series of maintenance-related indicators, including:

- Equivalent availability factor

- Forced outage rate

- Scram frequency

- Safety system actuations

- Gross heat rate

- Collective radiation exposure.

Table 2.1

Candidate Maintenance Indicators (from Literature Search)

Annunciator Alarms Continuously On
 #Lifted leads
 #Maintenance Work Request (MWRs) on Safety-Related Equipment
 #of Components Tagged out for Maintenance > 3 Months
 #of Maintenance Personnel Contaminations
 #of Missed Surveillances on Equipment
 #of Repeat Maintenance Items
 #of Temporary Modifications over 3 Month Delay (%)
 #Realignment Errors during Maintenance
 #Temporary Modifications
 #Wrong Unit/Wrong Train Events
 %Corrective MWRs Older than 3 Months
 %LERs due to Maintenance
 %Preventive MWRs Completed on Safety Equipment
 Accumulated Duration of Limited Conditions of Operation (LCO) Conditions
 Backlog of Engineering Change Notices (ECNs) Related to Equipment
 Performance
 Backlog of Maintenance Procedure Revisions
 Component in LCO Condition
 Corrective Maintenance Backlog > 3 Months
 ESF Actuations due to Maintenance & Testing
 Fraction Labor Hours on Surveillance
 Fraction MWRs Reviewed by Quality Control (QC)
 Fraction of Components Under Condition Monitoring
 Gross Heat Rate (Thermal Performance)
 Industrial Safety Lost-Time Accident Rate
 Maintenance Backlog
 Maintenance Overtime
 Maintenance Rework
 Maintenance Staff Radiation Exposure
 Maintenance Staff Size
 Mean Repair Time
 Mean Time Between Forced Outages from Equipment Failures
 Mean Time Between Repairs (Most Frequently Repaired Items)
 Mean Time to Return to Service
 Number & Duration of BOP Equipment Out of Service
 Part 21 Reports
 Preventive Maintenance Items Overdue
 Rate of Adoption of Industry Upgrades
 Rate of Calibration Errors
 Rate of Deferred Periodic Tests
 Rate of Downtime due to Degraded Failures
 Rate of Downtime due to Failures
 Rate of Faults Detected by Actual Demand
 Rate of Faults Detected by Periodic Testing
 Rate of LCOS
 Rate of Maintenance Errors
 Rate of Maintenance Requested Training Programs
 Rate of Maintenance Staff on Vendor Courses
 Rate of Maintenance Staff Retraining

Rate of Manhours in Maintenance
Rate of Misalignments
Rate of MWRs
Rate of MWRs completed
Rate of Out-Of-Service Tags
Rate of Pending Modification Requests
Rate of Time Spare Parts Unavailable
Ratio QC/Maintenance Staff
Ratio, #Hours to Repair Degraded Components/Total Maintenance Hours
Ratio, # Repairs while Degraded/#Repairs Failed + Degraded
Ratio, Deficiencies Discovered in Surveillance/Total Discovered
Ratio, Failures During Post Maintenance Test/# of PM Tests
Ratio, Highest Priority MWRs/Total MWRs
Ratio, Mean Repair Time/Time to Failure or Degrade
Ratio, Preventive Maintenance/Total maintenance
Ratio, Utility/Contractor Staff
Repair Duration w.r.t. Allowable Outage Times (AOT) by Tech. Specs
Safety System Function Trend
Safety System Performance Indicator
Scrams due to Maintenance & Testing
Turnover Rate/Vacancies
Wall Thinning/Pitting
Wrong Parts Events

During the informal meeting held with representatives from several overseas nuclear industry programs (during the 1988 IEEE International Conference on Human Factors and Nuclear Safety) an attempt was made to determine whether other countries were using indirect performance indicators in their industries. Although no directly relevant applications were identified in this meeting the systematic approach adopted in this study was generally supported by the participants.

2.6 Summary of the Results of the Literature Review

Maintenance is considered to be one of the plant activities most influential on safety. Its effects are pervasive, influencing the frequencies of scrams, the availability of safety systems and the ability of plant personnel to respond to incidents. Maintenance activities involve personnel in many departments in a plant when maintenance is broadly interpreted as the function responsible for the health and well-being of the equipment; both NRC and the industry groups have effectively adopted such a broad interpretation.

The NRC is currently developing several types of indicators associated with different levels of hardware/people interactions. Indicators have been developed and implemented for monitoring overall performance (number of scrams, numbers of significant events, etc.) for some time. Indicators at lower levels, such as equipment reliability indicators and LER causes, have previously been developed and are being applied by NRC on a trial basis. Programmatic indicators are being developed in several studies, including this one, for future trial application.

3. Frameworks for Examining Maintenance Programmatic PIs

The development of the set of frameworks has been an important part of the work reported here. These frameworks are utilized in an attempt to make explicit the roles of maintenance activities as they relate to safety, and the essential characteristics of maintenance programs. These characteristics form a significant influence on the shortlisting of indicators.

The frameworks were developed, in part, to describe concepts already developed by NRC, SAIC, and others about the attributes of effective maintenance and their influence on safety as has been summarized in the literature review. The frameworks also provide a specialized organization of these concepts for this project.

3.1 Purpose of Frameworks

Reference to the inventory of candidate indicators given in Table 2.1 shows that the survey activities related to the identification of indicators produced a large number and wide variety of candidates. It soon became apparent to the study analysts that some organized approach was required to review this set to determine which of these previously identified indicators showed the greatest promise as maintenance programmatic PIs, and to determine what additional indicators, if any, were required to be added to the set. It was the purpose of the frameworks developed to evaluate the existing indicators set and to identify additional ones in a systematic fashion. The set of frameworks therefore were developed to provide a common and uniform basis for this evaluation process which was to develop a shortlist of indicators.

3.2 Conceptual Relationship of Maintenance to Safety

The effect of maintenance on safety is both direct and indirect. One direct effect occurs because the performance of maintenance requires safety equipment to be taken out of service. This temporarily increases the unavailability of safety systems. This effect is one which is modeled explicitly in probabilistic risk assessments (PRA). However, PRAs have shown that, for the most part, this contribution to risk is small. A larger direct effect is that, subsequent to the repair, failed safety equipment may fail to be restored for accident prevention or mitigation by human error. In PRAs, this is found to be a more significant effect.

The indirect effects are more positive from a safety perspective. They include reducing the frequencies of scrams caused by failures in the balance-of-plant, and they include reducing the failure rates of safety equipment.

These indirect effects are achieved by controlling or improving the "health" of equipment over time---that is, by performing preventive maintenance, by appropriate replacement of aging components, and by performing effective testing. Utilization of these maintenance-related techniques in proper balance allows the availability of all systems (safety, steam supply, and balance-of-plant) to be maximized.

3.3 Maintenance Indicators Framework

One model developed based on the work described in Section 2 was the maintenance process model. This model was used to describe the interactions and extent of maintenance activities. In addition, two classification schemes for the evaluation of indicators were developed, associated with the impact of maintenance on safety, and with the human performance associated with maintenance programs.

3.3.1 *Process Model*

One perspective which was thought to be useful in developing indicators of a plant's safety performance was to develop a model of the process by which maintenance related activities affect plant performance, and hence safety. The approach here is to develop the information flow paths and associated information types between activities as potential sources of information to be used as indications of the plant's safety performance, much in the same way that a chemical process flow sheet or flow diagram is developed for a chemical system, or as a signal flow diagram might be developed for an electronic system, or as a health monitoring model might be developed for the internal vital signs of a human system. If such a model of a plant's maintenance process information flow could be constructed then it might be possible to intercept the information flow at key points (these points to be determined by this signal strength and the directness of the measured parameter's relation to safety) and so provide indication of the plant safety performance. It is argued that the measured information flow at these "probe points" would act in much the same way that circuit probe points provide health indications of electronic circuits.

The process model developed for this study was shown in pictorial form in Figure 2.3. This process model is a portrayal of the interactions and influences between the various functional units in a nuclear power plant. A brief description of the functional units and the interactions between these units shown in the process model are described in Appendix A.

The purpose in developing this framework, as has been mentioned, was to examine potential internal flowpaths and influences within an operating plant facility. As can be seen, there are many interactions between maintenance and other groups; somewhat more important are those between maintenance and operations, training, engineering support, and management. These are considered the most important because:

- (1) poor interactions between maintenance and operations can directly lead to equipment failures and scrams because of errors in selecting equipment for maintenance (a frequency cause of inadvertent scrams);
- (2) training can lead to systematic changes in knowledge and philosophy of maintenance activities within the plant. Weaknesses or strengths in this area will have an extensive influence on the quality of maintenance;
- (3) relations between maintenance and engineering support have a strong influence on the effectiveness of root-cause identification and the implementation changes to eliminate them; and
- (4) management decisions strongly affect the resources and priorities of all activities within the maintenance function.

From an analytical perspective, however, some of the important characteristics of this power-plant system are very different from an analogous electronic system. In the first place it is mostly a closed-loop system (that is, its previous performance strongly affects its present behavior) involving the use of plant-performance data analyses. Second, it is a complex system involving many elements that are interactively coupled. Finally, it is dynamic (that is, behavior in each element is not stationary) as in the case of plant reorganizations or responses to management changes. The extent to which the system can behave in a closed-loop manner is judged to be increased significantly by the

industry's recent adoption of the INPO systematic approach to maintenance programs, which encourages self-evaluations and peer reviews of maintenance programs and the use of root-cause analyses.

For these reasons although the model was originally developed with the intent of identifying "probe points" (i.e., points within the system at which monitoring should be undertaken), once the strongly interactive nature of the "feedback" loops were recognized, it was clear that indicators from within the process were unlikely to bear any relationship to overall performance in a practical way. This is because there were too many alternative ways of successfully arranging priorities and influences within an organization to be captured in an indicator program; that is, there were too many "degrees of freedom". This is a more analytical statement of the point made by industry representatives at the NRC's workshop on the maintenance rulemaking, namely that there are many different ways of implementing both effective and ineffective maintenance programs.

In consequence, the process model was used alternatively as a way of classifying indicators as "output-" or "process-" related. Output indicators were indicators tied to observable events ultimately resulting (usually) from equipment performance; these include equipment failure rates, the number of misalignment errors, and wrong parts events. Process indicators were associated with how maintenance is carried out within the organization. Examples of the latter include the backlog of preventive maintenance items, the fraction of work reviewed by QA, and the amount of overtime worked. Process indicators are measures from within the maintenance function that require internal assessments to provide data; their relationships to safety are less clear than those of output indicators.

3.3.2 Frequency/Duration/Quality Classification

The issues relating maintenance to safety discussed in Section 2.1 were used to consider the influences of maintenance on safety based on Figure 2.2. These are:

- (1) maintenance can detect safety equipment failures through testing or inspection. Once detected, corrective maintenance restores its operability;

- (2) maintenance can ensure operability of safety systems by checks and tests to ensure correct alignment of systems;
- (3) maintenance can improve the reliability of safety systems through the use of preventive maintenance to eliminate wearout or other time or cycle-related failure modes;
- (4) errors in maintenance or testing can lead to reactor scrams that result in demands on safety systems; and
- (5) equipment failures may occur in balance-of-plant systems that result in reactor scrams that, with preventive maintenance, could have been prevented.

In these cases, three maintenance-related variables demonstrate the impact on safety. These are: how often maintenance is performed, for how long maintenance is performed, and how well maintenance is performed. Each of these impacts the availability of safety equipment and the scram frequency. Therefore, a classification scheme based on these three variables should be relevant to assessing the impact of maintenance on safety. Using this classification scheme each of the candidate indicators were classified based on their association with the frequency, the duration, and the quality of maintenance.

Each of these factors reflect possibly different underlying activities and approaches to maintenance that can have an effect on safety. For example, more *frequent* corrective maintenance can result from a failure to correct the root causes of breakdowns, from deficiencies in the repair methods used, or from a failure to control aging mechanisms in an effective manner.

Increased *durations* of maintenance may be caused by understaffing or a lack of training and preparedness (such as a shortage of spare parts); alternatively, extended durations for certain types of equipment may reflect a management decision to await the next planned outage to perform upgrading or refits.

The *quality* of maintenance reflects the number of errors that result in an adverse influence on the plant, such as an inadvertent scram or ESF actuation. It may represent a

variety of underlying causes, including communication problems, incorrect training, and a lack of management monitoring.

3.3.3 Human-Factors Issues

The literature review described in Section 2.3 led to the creation of a classification of human factors issues found important in carrying out maintenance effectively at the personnel level. The classification elements are: (1) workload/fatigue, (2) communications, (3) training, and (4) supervision/administration.

4. Process for Identifying Candidate Performance PIs

The process utilized for identifying candidate indicators was directed at the identification of potential indicators and the screening of those identified so as to indicate those worthy of more detailed and in-depth analyses. The process used included three steps: (1) the identification of possible indicators; (2) an initial screening using a qualitative preference scheme; and (3) a quantitative screening analysis using statistical tests. As part of the quantitative screen process, data were gathered from selected plant sites (jointly with other contractors and NRC personnel) to support the statistical analyses. In addition, criteria set to be used for these analyses was developed from the framework results discussed in the previous section. These sets were structural into preference scheme to allow for the development of the final short list of indicators.

This section summarizes these steps. The results of the statistical analyses are described more fully in Section 5.

4.1 Identification of Maintenance Indicators

As was mentioned in Section 2.4, a total of 75 possible indicators were initially identified, from existing literature concerning performance indicators, and from reviews of maintenance technology descriptions. These were supplemented with indicators identified within the project from the perspectives represented in the frameworks.

4.1.1 Literature Review

The first task was a literature search of material related to existing maintenance performance indicators, described in Section 2.5.

4.1.2 Project-developed Indicators

Additional indicators were developed within the project to supplement those identified from the literature review. These were principally indicators associated with interfaces between maintenance and other functional units within a nuclear power plant, and were identified with reference to the process-model framework described in Section 3.3.1. These indicators were specifically associated with the relationship between maintenance and QA, training, and operations.

4.1.3 Aggregation of Indicators

The result of this indicator identification is the composite list of 78 candidate indicators provided in Table 4.1.

4.2 Qualitative Evaluation of Maintenance Indicators

As can be seen from Table 4.1 the resulting set of maintenance indicators is formidable. While it would be potentially possible for the NRC to track all of these-data gathering and processing effort could be prohibitive. Further, conflicting indicators would inevitably result from such an extensive and interrelated set of indicators. Therefore a more appropriate approach is to systematically review the comprehensive set developed to select a subset which is considered to be most representative of the effect of maintenance on safety. This was the approach followed in the study. In the present instance, the systematic selection was performed by developing a preference scheme, and applying the scheme to each indicator, in turn utilizing the insights derived from the frameworks described in Section 3 to highlight their individual characteristics. The application of the approach, which is described in the following sections, reduced the 78 indicators listed in the table to a subset of 9 which were considered for further statistical analysis utilizing the data collected at selected plant sites.

4.2.1 Evaluation Process

In the first step of the evaluation process, each of the 78 indicators was reviewed to determine its ability to reflect the attributes and characteristics of maintenance as defined in the frameworks. The indicators characteristics were: (1) does the indicator reflect the *outputs* (equipment availability, scrams due to maintenance, etc.) or the *process* (maintenance backlog, staff turnover, etc.) of maintenance, or both; (2) the *frequency*, *duration* and *quality* of maintenance; (3) the human-factors issues associated with *workload/fatigue*, *communications*, *training*, and *administration*. Subsequent to this review consideration was also given to the availability of relevant data to support the development of each indicator.

Table 4.1, Candidate Maintenance Indicators (Composite List)

Annunciator Alarms Continuously On
 #Lifted leads
 #Maintenance Work Request (MWRs) on Safety-Related Equipment
 #of Components Tagged out for Maintenance > 3 Months
 #of Maintenance Personnel Contaminations
 #of Missed Surveillances on Equipment
 #of MWRs Written by Maintenance Personnel
 #of Repeat Maintenance Items
 #of Temporary Modifications over 3 Month Delay (%)
 #Realignment Errors during Maintenance
 #Temporary Modifications
 #Wrong Unit/Wrong Train Events
 %Corrective MWRs Older than 3 Months
 %LERs due to Maintenance
 %Preventive MWRs Completed on Safety Equipment
 Accumulated Duration of Limited Conditions of Operation (LCO) Conditions
 Backlog of Engineering Change Notices (ECNs) Related to Equipment
 Performance
 Backlog of Maintenance Procedure Revisions
 Component in LCO Condition
 Corrective Maintenance Backlog > 3 Months
 ESF Actuations due to Maintenance & Testing
 Fraction Labor Hours on Surveillance
 Fraction MWRs Reviewed by Quality Control (QC)
 Fraction of Components Under Condition Monitoring
 Gross Heat Rate (Thermal Performance)
 Industrial Safety Lost-Time Accident Rate
 Maintenance Backlog
 Maintenance Overtime
 Maintenance Rework
 Maintenance Staff Radiation Exposure
 Maintenance Staff Size
 Mean Age of Maintenance Procedure Revisions
 Mean Repair Time
 Mean Time Between Forced Outages from Equipment Failures
 Mean Time Between Repairs (Most Frequently Repaired Items)
 Mean Time to Return to Service
 Number & Duration of BOP Equipment Out of Service
 Part 21 Reports
 Preventive Maintenance Items Overdue
 Rate of Adoption of Industry Upgrades
 Rate of Calibration Errors
 Rate of Deferred Periodic Tests
 Rate of Downtime due to Degraded Failures
 Rate of Downtime due to Failures
 Rate of Faults Detected by Actual Demand
 Rate of Faults Detected by Periodic Testing
 Rate of LCOS
 Rate of Maintenance Errors
 Rate of Maintenance Requested Training Programs

Rate of Maintenance Staff on Vendor Courses
Rate of Maintenance Staff Retraining
Rate of Manhours in Maintenance
Rate of Misalignments
Rate of MWRs
Rate of MWRs completed
Rate of Out-Of-Service Tags
Rate of Pending Modification Requests
Rate of Root Cause Evaluations due to Maintenance
Rate of Time Spare Parts Unavailable
Ratio QC/Maintenance Staff
Ratio, #Hours to Repair Degraded Components/Total Maintenance Hours
Ratio, #Repairs while Degraded/#Repairs Failed + Degraded
Ratio, Deficiencies Discovered in Surveillance/Total Discovered
Ratio, Failures During Post Maintenance Test/# of PM Tests
Ratio, Highest Priority MWRs/Total MWRs
Ratio, Mean Repair Time/Time to Failure or Degrade
Ratio, Preventive Maintenance/Total maintenance
Ratio, Utility/Contractor Staff
Repair Duration w.r.t. Allowable Outage Times (AOT) by Tech. Specs
Safety System Function Trend
Safety System Performance Indicator
Scrams due to Maintenance & Testing
Turnover Rate/Vacancies
Wall Thinning/Pitting
Wrong Parts Events

The product of this evaluation is given in a matrix representation in Table 4.2. In Table 4.2 shows, for each indicator, its relative strength on reflecting the attributes discussed above. In the table, a strong representation is shown by a "bold" entry.

Of the 78 candidate indicators, 16 are considered strongly associated with the maintenance process, and 11 are strongly associated with maintenance outputs; 10 indicators strongly reflect both the outputs of maintenance. Additionally, of the 78 indicators, 12 are considered to reflect strongly the frequency of maintenance, 12 the duration of maintenance, and 19 the quality of maintenance; 14 strongly reflect workload, 6 strongly reflect communications, 13 strongly reflect training, and 21 strongly reflect administration and supervision.

4.2.2 Preference Schemes

As has been stated, a preference scheme was developed and applied to reduce the 78 indicators to a set more suitable for analysis. The following was the basis for the preference scheme that was used:

1. Typically the outputs can be more directly seen as having an impact on safety such as with detected failures of safety systems; however, these are often infrequent occurrences. The process-related indicators are more tenuously associated with safety, especially since there are no normative models of organizational performance that apply directly to maintenance; however, data associated with these may be available more frequently, and therefore these may provide early indications. Therefore both output-and process-related indicators will be selected, but with a preference to output indicators. This is because output indicators have a more direct tie to safety; and they should be less easily manipulated since they often represent observable events. On the other hand process indicators could be manipulated via some type of "accounting" process.
2. The most important single attribute in the set of frequency/duration/quality was considered to be quality, again because of its direct tie to safety. Quality problems manifest themselves by incidents such as selection of wrong equipment for maintenance, unintentional ESF actuations, and so on, that can lead to plant-wide reductions in standards

Table 4.2, Evaluation & Maintenance PIs

P I				Maintenance			Human Factors				
	Data	Output/ Process	MIT Item(s)	Freq	Duration	Quality	Workload	Commun	Training	Admin	Other
# Annunciator Alarms Continuously On	o	p	7.5			x	x			x	
# Lifted leads	o	p	5.10,7.5	x						x	
# MWRs on Safety-Related Equipment	x	p	5.1	X			x				
# of Components Tagged out for Maintenance	x	p	5.10,7.5	X	X		x				
# of Components Tagged out for Maintenance > 3 Months	x	p	5.10,7.5	X	X	x	x			x	
# of Maintenance Personnel Contaminations	x	p	1.0,4.5	x	x	x				x	ALARA
# of Missed Surveillances on Equipment	x	p	5.5,3.4	x		X	x			X	
# of MWRs Written by Maintenance Personnel	x	p	3.3				X	x			
# of Repeat Maintenance Items	x	P	6.4	x		X	x		X	x	
# of Temporary Modifications over 3 Month Delay (%)	x	p	5.6,5.4	x	X	x	x			x	
# Realignment Errors during Maintenance	x	O,P	6.1,5.10	x		X	x	X	X		
# Temporary Modifications	x	p	5.10,7.0?	X	X		x			x	
# Wrong Unit/Wrong Train Events	X	O,P	1.0	x		X		X	X	x	
% Corrective MWRs Older than 3 Months	x	p	5.6				x			x	
% LERs due to Maintenance	X	O,P	1.0	x	x	X	x	x	x	x	
% Preventive MWRs Completed on Safety Equipment	x	p	5.1			x	X			x	
Accumulated Duration of LCO Conditions	x	p	1.0	x	X						
Backlog of ECNs Related to Equipment Performance	x	P	5.6			x	x	x		X	
Backlog of Maintenance Procedure Revisions	o	p	3.6,5.7			X	x	X	x	X	
Component in LCO Condition	x	p	1.0	X	X						
Corrective Maintenance Backlog > 3 Months	x	p	5.6			x	x			x	
ESF Actuations due to Maintenance & Testing	X	O	1.0	x		X	x	X	X	X	
Fraction Labor Hours on Surveillance	x	p	3.4,5.5				x			x	
Fraction MWRs Reviewed by QC	x	p	4.4			x		x		x	

Table 4.2, Evaluation & Maintenance PIs

P I				Maintenance			Human Factors				
	Data	Output/ Process	MIT Item(s)	Freq	Duration	Quality	Workload	Commun	Training	Admin	Other
Fraction of Components under Condition Monitoring	x	P	3,4,4,2			X					Tech . Support
Heat Rate Loss (Thermal Performance)	X	O,P	1.0			X					
Industrial Safety Lost-Time Accident Rate	x	p	8.4	x	x	X	x	x	x	x	
Maintenance Backlog	X	p	5,6,2,2?	x	x	x	X			x	
Maintenance Overtime	x	p	3.3	x	x	x	X			x	
Maintenance Rework	x	O,P	1.0,6.4	x		X	x		X	X	ALARA Resources
Maintenance Staff Radiation Exposure	x	p	4,5,1.0		x	x				x	
Maintenance Staff Size	x	p	8.1				x			x	
Mean Age of Maintenance Procedure Revisions	x	p	3.6			x	x	x	x	x	
Mean Repair Time	x	p	5.3		x		x		x	x	
Mean Time Between Forced Outages from Equipment Failures	x	O,P	1.0			x			x	x	Catchall
Mean Time Between Repairs (Most Frequently Repaired Items)	x	p	5,2,6,3?	x		X			x	x	
Mean Time To Return to Service	x	p	5.3		X		x			x	
Number & Duration of BOP Equipment Out of Service	x	o,p	5,4,5,6	X	X		X			X	
Part 21 Reports	X	O,P	1.0,2.1			x					
Preventive Maintenance Items Overdue	x	p	5,6,2,2?			x	X			X	Resources
Rate of Adoption of Industry Upgrades	x	p	2.1							x	
Rate of Calibration Errors	x	o,p	6,3,6,4			X	x		X	x	
Rate of Deferred Periodic Tests	x	p	3,4,5,5				x			x	
Rate of Downtimes due to Degraded Failures	x	o,p	1.0	X	X		x				
Rate of Downtimes due to Failures	x	o,p	1.0	x	x		x				
Rate of Faults Detected by Actual Demand	x	p	1.0			x	x				
Rate of Faults Detected by Periodic Testing	x	p	1.0			x	x				
Rate of LCOs	x	p	1.0	x	x	X	x			x	
Rate of Maintenance Errors	x	O,P	6.3	x	x	X	x		X	X	

Table 4.2, Evaluation & Maintenance PIs

P I				Maintenance			Human Factors				
	Data	Output/ Process	MIT Item(s)	Freq	Duration	Quality	Workload	Commun	Training	Admin	Other
Rate of Maintenance Requested Training Programs	o	p	8.2			x	x		X	X	
Rate of Maintenance Staff on Vendor Courses	o	p	8.2			x	x		X	X	
Rate of Maintenance Staff Retraining	x	p	8.2			x	x		X	X	
Rate of Manhours in Maintenance	x	p	3.3	X	X		X			x	
Rate of Misalignments	x	O,P	6.1,7.5	X		x	X	X	X	X	
Rate of MWRs	x	p	5.5	X			X			x	
Rate of MWRs Completed	x	p	5.5	X	x		X			x	
Rate of Out-Of-Service Tags	x	p	5.5,7.5	X	x		X			X	
Rate of Pending Modification Requests	x	p	5.5	x	x		x	X		X	
Rate of Root Cause Evaluations due to Maintenance	x	P	3.5,6.4	x		X	x		X	X	
Rate of Spare Parts Unavailable	x	P	3.3,7.2	x		x	x			X	Planning
Ratio QC/Maintenance Staff	x	p	3.3,4.4							x	
Ratio, # Hours to Repair Degraded Components/Total Maintenance Hours	x	p	5.2								
Ratio, # Repairs while Degraded/# Repairs Failed + Degraded	x	p	5.2			x					
Ratio, Deficiencies Discovered in Surveillance/Total Discovered	x	p	3.4								
Ratio, Failures During Post Maintenance Test/# of PM Tests	x	p	6.4			X	x		x	x	
Ratio, Highest Priority MWRs/Total MWRs	x	p	5.4,5.2			x	x			x	
Ratio, Mean Repair Time/Time to Failure or Degrade	x	p	1.0	x	X		X				
Ratio, MWRs Received/MWRs Completed	x	p	5.6				X			X	
Ratio, Preventive Maintenance/Total Maintenance	x	p	5.6,3.4			x	X			X	
Ratio, Utility/Contractor Staff	x	p	6.2,8.1							x	
Repair Duration with respect to AOT by Tech. Specs	x	p	1.0		X		x			x	
Safety System Function Trend	X	P	1.0	x	x						

Table 4.2, Evaluation & Maintenance PIs

P I				Maintenance			Human Factors				
	Data	Output/ Process	MIT Item(s)	Freq	Duration	Quality	Workload	Commun	Training	Admin	Other
Safety System Performance Indicator	x	p	1.0	x	x						
Scrams due to Maintenance & Testing	X	p	1.0	x		X	x		X	X	
Turnover Rate/Vacancies	x	p	8.1,8.4				x			X	
Wall Thinning/Pitting	o	p	1.0			x					
Wrong Parts Events	X	O,P	3.3,7.2	x		X	x		x	X	
<div> <div>KEY</div> <div> <p><u>Data</u></p> <p>X = Data readily available</p> <p>x = Data available</p> <p>o = No data available</p> <p><u>Output/Process</u></p> <p>O = Strong output ind.</p> <p>o = Output indicator</p> <p>P = Strong process ind</p> <p>p = Process indicator</p> <p><u>MIT Item</u></p> <p># = Entry # in MIT</p> <p><u>Other Entries</u></p> <p>X = Strong Indicator</p> <p>x = Possible Indicator</p> </div> </div>											

of safety. Quality problems manifest themselves by incidents such as selection of wrong equipment for maintenance, unintentional ESF actuations, and so on, that can lead to plant-wide reductions in standards of safety. The second preference was for the frequency of maintenance because patterns of repeated repair activities are indicative of a lack of systematic elimination of equipment problems, such as repairing symptoms rather than causes of failure. In contrast, lengthy durations of maintenance may be caused by many reasons, which may be either good (extension root-cause analysis, planned major overhaul) or bad (low morale, shortage of parts, work overload).

3. From a human factors perspective, the most important influences on system-wide performance are communications (particularly between operations and maintenance and within the maintenance group itself), and supervision/administration. Problems in these areas can extend their impact to many individuals and systems in a short timescale.

In summary, the following preferences were used to select indicators:

1. Output indicators were preferred over process internal indicators because of their closer tie to safety. However some process indicators were included as a comparison, because process indicators may act as an earlier signal;
2. The most important single attribute in the set of frequency/duration/quality was quality, again because of its direct tie to safety. The second preference was for the frequency of maintenance because patterns of repeated repair activities indicate a lack of systematic elimination of equipment problems.
3. From a human factors perspective, the most important influences on overall performance was considered to be communications and supervision/administration. This is because problems in these areas can extend their impact to many individuals and systems in a short time-scale.

4.2.3 Shortlisted Maintenance Indicators

Based on the preference scheme described in the previous section, the shortlisted indicators identified were:

- Number of repeat maintenance items
- Number of realignment errors
- Number of wrong unit/wrong train events
- Number of inadvertent ESF actuations
- Backlog of ECNs related to performance
- Mean time between repairs
- Scrams due to test & maintenance
- Wrong parts events
- Thermal performance

A brief description of the interpretation of these indicators, and a rationale for their particular section, is given below.

4.2.4 Synopsis of Indicators

The following sections briefly explain, for each of the nine selected indicators, why it was selected and how it is interpreted in terms of plant data:

Number of repeat maintenance items

This indicator is a strong indicator of the quality of maintenance through its reflection of the need to repeatedly visit a component or group of components. It is differentiated from the "maintenance rework" indicator in that it does not imply that the same cause of failure requires repair; this is important in terms of the ability to identify similarity in causes from the available data. Repeated repairs are also a reflection of the frequency of maintenance and of deficiencies in training.

It is interpreted as the number of components which are revisited by maintenance staff for corrective repair after replacement in service in a given quarter.

Number of realignment errors

The number of realignment errors is the number of time trains of safety systems are restored to service following test or maintenance with an inappropriate configuration (usually of valves or circuit breakers) such that the train would not be available on demand. This would be a reportable event as an LER.

It strongly reflects the output of maintenance (systems unavailability) because of maintenance errors. It is considered similarly to reflect strongly the quality of maintenance, and problems in communications and training.

Number of wrong unit/wrong train events

This indicator is very similar to the number of realignment errors described above, in that it strongly reflects the quality of maintenance, and problems in communications and training. It is an output indicator because of its influence on system unavailability. It has been reported as being strongly influenced by poor communications between operators and maintenance staff (NUREG-1192).

It is again counted by the number of reported events in a given quarter via LERs.

Number of inadvertent ESF actuations due to test & maintenance

This is considered self-explanatory. Inadvertent ESF actuations are considered important because of their direct association with errors, such as accidental selection of the wrong cabinet or relay for maintenance tasks; it is a direct and unambiguous output of maintenance (it usually causes a scram); it reflects strongly the quality of maintenance; and it is strongly influenced by poor communications and training.

Backlog of ECNs related to performance

This is a process-oriented indicator that strongly reflects a lack of management and administrative support to maintenance, particularly because Engineering Change Notices (ECNs) are usually the effective path to the elimination of root causes of failure. (An ECN is an internal document that describes and reports

plant modifications; it acts as the coordination document between Engineering Support, Maintenance, Operations, and other units in the plant.) A willingness to accept an increasing backlog of ECNs indicates a willingness to deal with symptoms rather than remove causes. It is accepted that this indicator in isolation is not a complete measure of effectiveness, but may provide a part of the picture. It is calculated on a continuously tracked basis as reported within the plant.

Mean time between repairs

Mean time between repairs is again a measure of the need to repeatedly visit individual components, whether or not it is to repair the same failure modes. It is functionally similar to the number of repeat maintenance items, but can be tracked as a continuous variable.

It is calculated as the mean of the times between pieces of equipment being returned to service following maintenance, to the next time the equipment is reported unavailable. It is intended to monitor only the most frequently repaired equipment in this way.

Scrams due to test & maintenance

This is directly comparable with the number of ESF actuations. It broadly covers mistakes in the maintenance interface with equipment and with operations. It is strongly reflective of the quality of maintenance and of communications and administration of maintenance. It is based on LER reported events.

Wrong parts events

Wrong parts events are reportable events where equipment fails because incorrect materials or parts were used in a repair. This is again a strong reflection of the quality of maintenance, and of the supervision of maintenance. It is an output indicator reflecting maintenance-caused equipment failures.

Thermal Performance

Thermal performance is a catch-all kind of indicator that reflects the overall care that has been taken of the facility, and particularly the preservation of assets by the maintenance facility. If the plant is not preserved, then the economic efficiency will decline accordingly. For this analysis, gross heat rate (the ratio of the thermal to electrical power) was selected for its ready availability of data.

4.3 Data Collection

Visits were made as part of this project to 8 plant units to collect data. These are identified as plants A-H in this report.

These visits were made as part of a task to collect maintenance-related data for several indicator programs.

4.3.1 Identification of Data Needs

Data needs were identified in conjunction with another data-collection program and were aimed at particular indicators' parameters. These were identified by a joint NRC/contractor team that was created to prepare for the data-collection visits.

4.3.2 Data Sources

The principal data sources were the maintenance-related work records at the 8 plants, supplemented by data reported in reportable programs such as Licensee Event Reports. These were summarized according to the data needs of the various indicators. In addition, some of the indicators required information from LERs and from operational data summarized in NUREG-0020, Licensee Operating Reactors: Status Summary Report (published monthly). Particularly, data associated with indicators of events (inadvertent ESF actuations, inadvertent scrams, wrong train/wrong unit events, wrong parts events, and realignment errors) were derived from LERs. Data associated with gross heat rate were developed from NUREG-0200.

In the case of event data, this search was performed manually by reviewing descriptions of events in NRC's computerized LER database. The following groundrules were adopted for this review.

1. The event must have occurred on or after January 1, 1985, the timescale associated with other data sources.
2. The reactor must have been at power at the time of the event; i.e., the plant was not shut down for maintenance or refueling. (This eliminates many ESF actuations that appear to occur during equipment modifications, testing, etc.)
3. There was an apparently significant involvement of operations (training indicators) or maintenance staff (maintenance indicators) in the ESF actuation events. This may have been direct, such an error during a test that causes shorts between terminals, or indirect, as in missing procedures.

LERs were used as an important data source for this study. While LERs have value as a data source, their value is limited because of the limited descriptions of events, and because events are reported only when there is a conscious realization by the utility that the event matches the current LER reporting requirements. It is our experience that some event types, such as breaches of Technical Specifications, can be under-reported by up to a factor of 5. These under-reportings are largely a result of a failure to recognize the equipment interactions. While for safety-significant events, such as ESF actuations, the under-reporting is judged to be less in some cases. For example, consider the case of wrong train/unit events or realignment errors. These errors may occur and be detected before a system is restored to service. In this case such an event may not be reported. The occurrence and self-detection of the error is not necessarily an indication of either good or bad performance. So from the perspective of overall performance indicators this omission from the LERs was not considered to be significant.

However, a second and more significant limitation in this project was the limited accessibility of plant data. This problem was particularly acute for those plant which do not store data in an easily retrievable form. In the case of these plants it must be recognized that separate and extended visits would be required to support better and more

detailed interpretation of the available data should related indicators prove worthy of analysis.

4.3.3 Data Analysis

The data obtained from the plant and from the sources listed were stored in computer-readable files after inconsistencies and ambiguities were resolved. This resolution activity as well as the computer input activity was also conducted as a joint effort among AEOD staff, and the staff of this project.

4.4 Validation Process

The purpose of the analysis performed was to investigate the ability of the indicators to anticipate changes in safety performance of the plants; in other words, to what extent did changes in the indicators lead to subsequent changes in safety, and with what confidence could this relationship be established. The validation process which was carried out is shown in a process-diagram form in Figure 4.1. The performance of this analysis required the development of independent references-measures of safety performance for each plant. These measures, listed as "Alternative Validation Measures" in Figure 4.1, were combined with models which provide estimates of overall plant safety based upon the measure values used to establish an indicator validation scheme. The modeling approach used was similar to, but an extension of the methodology used in the Accident Sequence Precursor (ASP) Study. The ASP approach was extended to include trips, Loss of Feedwater events, as well as significant human error events.

The validation process involved estimating the time scales and degrees of association between the indicators and a particular measure of safety. The estimate was performed using standard methods approaches from systems modeling and statistics. However care was taken in this evaluation to avoid the assumptions of linear, normal, or static distributions.

4.4.1 Factors Important to Performance PI Validation

Any validation study includes constraints which limit the level of validation obtainable. In this particular study the constraints included the quantities of data available, the need to allow for lead/lag relationships, and the presumption of a relatively "loose"

Figure 4.1, Validation Process Diagram

relationship between indicators and measures of safety. These constraints certainly limit the precision of the validation process. However, since an implicit purpose of the programmatic indicators is to alert NRC to perform more close monitoring by other indicators, a high degree of precision is not necessary. For example, false triggering would not have an immediate consequences. However, a prolonged period of false alerts would certainly weaken the acceptability of the indicators. Based on these factors, a degree of association of 80% was selected as providing an adequate degree of confidence in the existence of a relationship between the indicator and safety. Therefore, this criterion was used to select the indicator to be evaluated in later phases of the program.

As has been mentioned, validating the effectiveness of indicators in providing a safety signal requires the selection of some measure or measures of safety. Two approaches were taken in this project to establish such a measure. First, an overall measure of safety, based on plant risk, was constructed; and second, intermediate measures of safety currently used by NRC were used as reference measures of safety. The purpose of this two-fold approach was to permit more in-depth analyses to be performed than those which would be based on just selecting one measure. Additionally the use of two measures provided redundancy in the search for validation.

4.4.2 Alternative Criteria for Validation

Several alternative measures were considered as a basis for the objective measures of safety; these include the previously used SALP numerical ratings, "significant event" occurrences, risk-oriented measures of scrams and system unavailabilities, and measures of unit and systems availability.

The preferred attributes to be considered for the selected measure were that it should be based on a variety of historical events (in the spirit of the Commission's wish to use retrospective analyses for validation), that it should be quantifiable, and that it should be strongly associated with risk as an expression of safety significance. None of the measure reviewed uniquely met these preferences. As a result, it was decided to use two approaches; these were to develop for a trial basis, a metric of safety performance that reflected the historical risk perspective, and also to use a series of indicators representing intermediate elements of safety, such as the number of scrams, system unavailabilities, etc.

The metric of safety was developed to reflect the above concerns. As was mentioned, it was based on methods developed as part of the NRC's Accident Sequence Precursor Study (ASP), but methodology was extended for this application. The ASP program provides a methodology for calculating the conditional probability of core damage given the occurrence of equipment failures and initiating events. These analyses were performed using PRA models (fault trees and event trees) for each class of plant. This methodology is summarized in NUREG/CR-4674, Precursor to Potential Severe Core Damage Accidents: Status Report. A version of this report is issued each year for events occurring the past year.

For this program, a review of the License Event Reports for each plant in the database was performed and relevant events were identified. Relevant events were events involving scrams, loss of feedwater events, loss of offsite power events, and significant human-factors events. This range of events was selected for their potential impact on risk as had been determined from the previous PRA studies known to the analysts; this range of events is broader than the range of events normally considered in ASP analyses. Further, it should be noted that no loss of coolant accidents or anticipated transients without scram occurred at these plants in the period of concern.

Using this extended ASP methodology, the conditional core damage probabilities were calculated for each event. These probabilities are calculated taking account of equipment failures described in the event report. The values were then plotted as a scatter plot according to the data of occurrence of each event. The results of these analyses are shown in Figure 4.2 and 4.3 for one of the reference plants. In these charts, the "Y" axis is the probability of core damage (expressed as a *rank* in ASP terminology) and the date of occurrence is the "X" axis. In Figure 4.3, "Day 0" is January 1, 1984.

A smoothing technique was developed in order to provide a distribution through these scatter-plot points, a. This method consisted of calculating a "decay" of the risk value over time as an expression of the decreasing impact on plant and human behavior. A time constant was selected to reduce the exponent of the risk impact of the event by about a factor of 2 over four months. This was judged to imply that the impact of the event is small after one year.

In order to reflect the influence of recurring or otherwise frequent events, a cumulative distribution of these individual events was calculated and plotted. These form a "safety

Figure 4.2, Preliminary Event Data

Figure 4.3, Smoothed Event Data

trajectory" as a measure of safety performance; that is, a track of the safety measure overtime. An example of such a smoothed trajectory is shown in Figure 4.3.

In addition, intermediate measures of safety were used in common with other performance indicators studies; these broadly correspond to the overall indicators of safety discussed in Section 2.4.1. The intermediate measures of safety that were used are:

- No. of scrams
- No. of scrams from high power
- No. of scrams from low power
- No. of safety system actuations
- No. of significant events
- No. of safety system failures
- Forced outage rate
- Equipment forced outage rate
- No. of hours while critical
- Most recent SALP maintenance rating
- Equipment failure rate (NPRDS)
- Equivalent availability

Each of these is reported on a quarterly basis under current NRC programs. Most of these intermediate indicators represent one perspective or another of safety. For example, the frequencies of initiating events are represented by the numbers of scrams (at various power levels), the number of safety system actuations, and by the forced outage rate and equipment forced outage rate. Availability of safety equipment is represented by the numbers of safety system failures and the equipment failure rate. More complex situations are represented by the numbers of significant events; and "global" evaluations of safety are presented by the SALP rating.

These measures have been used by other performance indicator projects as their reference measures of safety.

4.4.3 Evaluation Process

The evaluation of the indicators against the measures of safety identified two principal values; these were (1) the estimated lead-lag time between changes in the indicator and changes in performance, and (2) the degree of association between the time-adjusted indicator and the measure of performance. The degree of association was obtained as a non-parametric (i.e., not based on assumptions of a normal distribution) measure that measures the probability that any associated changes in values are not the result of random variations. Kendall's Tau test is used for these estimates.

The evaluation of all the investigated lead/lag relationships was performed using a cross correlation method. This method was developed as a systems analysis method to evaluate time lags in data signals. Mathematically a cross correlation function X_k is calculated for two time series (U, V).

$$X_k = \text{COV}(U_t, V_{t+k})$$

$$[\text{Var}(U) * \text{Var}(V)]^{1/2}$$

Where:

K is a lead (or lag) time

COV (U, V) is the covariance of U, and

Var (U) is the variance of U.

X_k is calculated for a range of values of K. When X is a maximum, the two time series are "best matched."

The term "degree of association" indicates a statistic that describes the likelihood that two series of numbers are associated in terms of their trends. For example, if in two series, the entries are:

Series A	1	2	3	4	5...
Series B	2	3	4	5	6...

Then, in this case the two series are completely associated; (i.e., both have identical trends). In another case:

Series A	1	2	3	4	5...
Series B	2	3	5	4	6...

The series are associated to a significant degree. Lastly,

Series A	1	2	3	4	5...
Series B	5	4	3	2	1...

In this case, the series are completely negatively associated.

This degree of association is estimated by means of Kendall's Tau Statistic. This statistic estimates the likelihood that two series of data are associated in the manner just described. It is a well-known statistic and is in many tests on statistics. Kendall's Tau statistic is a non-parametric statistic; that is, its results are not based on underlying assumptions of normal distributions. Such a method was selected because there was no reason beforehand to assume that normal distributions would be relevant to performance indicators, especially with the sparseness of data.

A degree of association of 80% was selected in deciding which indicators were significantly associated with safety. In other words, the likelihood of the associated being random combinations was less than 20%. This criterion was selected because it was judged to be a reasonable basis for screening the indicators for their effectiveness for assessing safety. A significance level more stringent than this was judged to be likely to reject possibly valid indicators as a result of data deficiencies. A significance level less than 80% was judged to be likely to provide too many spurious relationships.

The evaluations of the cross correlation function and Kendall's Tau statistic were performed using two computer codes: PC-MATLAB (cross correlation) and Number Cruncher Statistical System (NCSS) (Kendall Tau).

5. Results of Preliminary Evaluations

5.1 Detailed Results

The detailed results of the analyses for each plant are shown in Tables B.1, B.2, B.3, and B.4; these tables are contained in Appendix B. Each plant has a similar set of four tables associated with it. Corresponding to these tables included in Appendix B are associated graphs for each plant. The graphs presented will be discussed in detail along with Table B.2.

Table B.2 identifies the maintenance-related analyses performed for the particular plant; these indicators are designated as:

5REWORK - The number of repeat maintenance actions identified from maintenance records;

9MTSERV - Mean time to return to service;

CSCRAM - Inadvertent scrams due to test and maintenance errors;

GHR - Gross heat rate, the ratio of the gross thermal output to the gross electrical output;

WPE - Wrong part events, as discussed in Section 3.

REALE - Realignment errors, where components in a train are incorrectly coupled to components in another train or another part of the system (such as where a test loop is left aligned during operation);

BACKECN - Backlog of engineering change notices (notices of intended plant modifications) required for continued design performance (not design improvements);

WRONGT/U - Wrong train/unit events as discussed in Section 3;

ESFACT - Inadvertent ESF actuations due to test and maintenance as discussed in Section 3.

The measures of safety (MOS) were the safety indicators presently used by NRC as discussed above. These are designated:

1SCRTL - Number of scrams.
2SCRHI - Number of scrams from high power (>15%)
1SCRLO - Number of scrams from low power (<15%)
2SSA - Number of safety system actuations
3SA - Number of significant events
4SSF - Number of safety system failures
5FOR - Forced outage rate
7EFO - Equipment forced outages
CRITHOUR - Numbers of hours critical
SALP-M - SALP maintenance rating
FRATE - Composition safety system failure rate
AVAIL - Availability (average) (NPRDS)

In addition, a rank value derived from the ASP program that reflects the safety significance of reportable events was used as an additional MOS. This is identified as RANK in the tables.

In Table B.1, those entries designated "N/A" are combinations of indicators and MOS that were not analyzed because of insufficient or no data. "NS" designated entries are those combinations where the analyses showed no significant degree of association. These entries marked "S" are combinations where significant degrees of association were found. For some indicators no events were reported in the data for the full period; with zero events, no relationship could be tested.

Table B.2 summarizes the relationships identified as significant. These summaries identify the time relationships given in the cross correlation analysis and the degree of association calculated using Kendall Tau tests. These results are also represented in graphical form as discussed below.

For each plant and each significant relationship identified, two set of graphs are presented. The first set of these graphs consists of a series of three graphs presented on a single page. The first two of the graphs on the page are plots of the raw data for the measure of safety values and maintenance indicator values. The third graph is a plot of the cross correlation function showing the value of the function at various time shifts. Note that the cross correlation function reaches a peak for the time shift the two time series best match. The second set of graphs for each significant relationship consists of a scatter plot of the particular measure of safety values versus the maintenance indicator where the maintenance indicator values have been shifted the number of quarters suggested by the cross correlation function maximum. Each of these scatter plots is overlaid with a regression line.

5.2 Summary of Results

Four of the nine selected maintenance indicators showed potential relationships to the range of measures of safety. These were:

Gross Heat Rate (GHR)
 ESF Actuations from Test and Maintenance
 Scrams from Test and Maintenance
 Mean Time to Return to Service (MTRS).

Of these, the scrams and the MTRS indicators only appeared in a few cases and were only considered as secondary indicators. The GHR was determined to be associated with the individual MOS in the following number of plant instances:

Availability	-	3 plants
EFO	-	2
FRATE	-	2
SSF	-	2
CRITHOUR	-	1
RANK	-	1
SALP	-	1
SE	-	1

The corresponding association between inadvertent the "ESF actuation from test and maintenance" indicator and the MOS is summarized below:

EFO	-	4 plants *
Availability	-	3 *
CRITHOUR	-	3
SCRHI	-	2
SSF	-	2 *
SCRTL	-	1
FOR	-	1

(* indicates leading relationship).

For three plants, there were insufficient ESF actuations to evaluate relationships.

6. Further Analysis of Promising Indicators

Following the statistical analyses summarized in Section 5, the two indicators showing potential for acting as indicators (ESF actuations from test & maintenance errors and thermal performance (gross heat rate)) were subjected to further scrutiny to test whether they did actually reflect measures of safety, and whether there appeared to be some logical connections between the indicators and plant safety. The analysis approach was adopted as a result of recommendations and comments received from NRC staff in various review meetings held during the course of the project; it is an approach that was similar to that used in other NRC indicator programs.

In summary, the results of this analysis appear to indicate that there is a secure foundation for the ESF actuations-related indicator. This indicator appears to represent the effectiveness of plants to identify and eliminate root causes of equipment and other performance problems. In turn, this lack of effectiveness to manage plant reliability is one symptom of a problem organization.

The analysis also provided evidence that the indicator associated with thermal performance required additional work. The indicator measure used in the preliminary analysis had several deficiencies including its insensitivity and its tendency to be influenced by many factors other than the quality of maintenance. Additional work has been performed to formulate and evaluate an indicator which reflects thermal performance but does not have these deficiencies. This work appears promising but is not yet complete.

6.1 Inadvertent ESF Actuations from Test & Maintenance Errors

The ESF actuations indicator is a measure that reflects how plant staff and maintenance supervisors perceive the importance of ESF challenges and how they act to remedy their causes. Inadvertent ESF actuations are considered important because of their direct association with errors, such as accidental selection of the wrong cabinet or relay for maintenance tasks; they are a direct and unambiguous output of maintenance (often causing a scram); they reflect strongly the quality of maintenance; and they are significantly influenced by poor communications and training. This analysis has concentrated on two comparisons. First, what kind of performance related to maintenance was discussed in the narrative sections of relevant Systematic Assessments

of Licensee Performance (SALP), and, second, what kinds of remedial or corrective actions dominate the plants' responses to these events.

The nature of this indicator relates to the portrayal of management's responses to moderately significant events. Inadvertent ESF actuations in themselves are not immediately significant in terms of risk, though some fraction are accident sequence-initiating events. However, these events are reportable to the NRC; this reporting entails an administrative burden, as well as an additional burden to the operating staff caused by the requirement to restore the affected systems to their normal standby mode. For those actuations causing a scram, there is the additional costs of lost generation. In other words, inadvertent ESF actuations are events that plants will wish to avoid for operational, economic and administrative reasons.

A continuing number of these events with no apparent decreasing trend suggests several possible alternative explanations:

- (1) an absence of any true root-cause analyses that identify contributing factors and eliminate them;
- (2) an excessive workload that prevents the elimination of causes of these events; or
- (3) an implicit or explicit acceptance of inadvertent ESF actuations as being of no significance to plant performance.

Any of these are a concern in assessing the programmatic aspects of maintenance.

Ten plants were analyzed in this study. These include the original eight plants analyzed in the statistical study (A-H).

In addition, two plants (I,J) were added to extend the range of analysis to include plant with previously identified problems.

6.1.1 *Method of Analysis*

This section describes: (1) the sources of data; (2) the evaluation process; and (3) the methods of comparison.

Data Sources

The primary sources of data for this analysis were the full text descriptions of events contained in Licensee Events Reports (LERs), and the full text of the SALP reports spanning the period of analysis, from 1/1/85 to 7/1/88.

ESF actuations are occurrences that must be reported as LERs under the requirements of 10 CFR 50.73 (a) (2) (iv), which states "The Licensee shall report...any event or condition that resulted in manual or automatic actuation of any Engineered Safety Feature (ESF), including the Reactor Protection System (RPS). However, actuation of an ESF, including the RPS, that resulted from, and was part of, the preplanned sequence during testing or reactor operation need not be reported". This is clarified in NUREG-1022, (NRC, 1983) specifically to point out that if ESF actuations during testing occur in a way that is not part of the planned procedure, then they must be reported. Based on these requirements, then, all inadvertent ESF actuations that occur should be reported as LERs, including those resulting from test and maintenance.

In the statistical analyses described earlier, summary LERs coded as being related to ESF actuations were reviewed. For this part of the analysis, full texts of LERs associated with apparent test and maintenance errors were obtained and reviewed for reevaluation of the detailed description of the event. In addition, LERs for other ESF actuation events were obtained on a sampled basis to ensure that the use of summaries as the basis for selection was not biasing the results. In general, the summaries identified the events correctly, with one or two additions found. However, the number of summaries in the initial analyses did not appear biased compared with the full test reports.

The second data source was the texts of the SALP reports that spanned the period of concern in this analysis. These reports were obtained from the docket files in the NRC's Public Document Room. The emphasis in reviewing these reports was the SALP Board's qualitative opinions of the strengths, weaknesses and trends of performance related to the maintenance and testing functions. In the earlier SALP reports, these were described

under maintenance, surveillance, and quality programs and administrative controls. In the more recent SALP reports, the area of engineering and technical support was added; these new areas have been reviewed also. The purpose of concentrating on the qualitative assessments was to emphasize the technical performance issues seen during the inspections, rather than the quantitative score, which does not define the technical issues.

Evaluation Process

The review of ESF actions concentrated on identifying events that had the following characteristics:

- (1) they occur while the reactor was at a power greater than 0%,
- (2) the involvement of maintenance or testing activities was a significant contribution to the event.

Satisfying the first requirement, that the reactor was at power, was unambiguous; the reactor power at the time of the event was explicitly stated as item 10 on the LER cover sheet. Satisfying the second requirement, however, required an opinion based on the description in the LER. In most cases, the role of maintenance or testing errors was clearly stated. Examples include incorrect valving of instrument lines during calibration, inadvertent connections between terminal blocks during maintenance, and so on. However, there were some events where the role of maintenance was ambiguous. For example, one event occurred during painting, when a cover cloth fell and tripped equipment causing an ESF actuation. This event was included because it was judged that the maintenance function could have arranged precautionary measures to protect equipment controls before painting started. However, only three such ambiguous events were identified and included; the inclusion of this small number is not considered significant in terms of the overall results.

In addition, some events appear to be maintenance problems but were not included in this study. Those excluded were a series of events which occurred at plants E & F and were associated with poor (unreliability) performance of ammonia and chlorine gas detectors leading to frequency actuations (up to 12 per year) of the standby heat and vent system (an ESF system at these plants). In this analysis, these events were excluded because the

"real" cause was judged to be a poorly designed system. In this case the maintenance was related to preserving the availability of these detectors by maintaining microscopic clearances. Several "less-than-perfect" maintenances acts were reported as causing ESF actuation in these systems. However, in this analysis, these specific events were considered not to be relevant because the demands on maintenance were considered to be excessive. Further it was determined that their inclusion would not significantly change the overall results.

One issue that could possibly be of significance in evaluating ESF actuations is that there is no common definition of what comprises the ESF equipment across all plants, or even across types of plants. The exact scope of the ESF system is defined separately for each plant. Despite this variation, most plants have more or less a common scope that includes the Emergency Core Cooling Systems (ECCS), the Emergency Power System (EPS) and the Standby Heat and Vent Systems (SHVs). It was therefore judged that in practice, the variability in definitions was not significant.

The intent of the review of the SALP reports was to identify basic problems or strengths. These included the commitment of management and the availability of adequate resources, an effective planning and prioritization process, and adequate interfaces with other functional areas. Activities such as these are important in a plant having an effective self-improvement process as was discussed earlier. The analytical process involved was simply to read comments associated with these activities.

Method of Comparison

In this study, the comparison was made between the pattern of inadvertent ESF actuations and the commentary from the SALP reports on a plant-by-plant basis.

6.1.2 Results

For four plants, no event meeting the classification of inadvertent ESF actuation from test and maintenance was identified. These plants are D, G, H & J.

The other six plants reported events that met the classification used in this study. These plants, and their total number of events (1985-1988) are:

A-18

B-9

E-8

F-6

D-3

C-2

Each site is reviewed in turn.

Plants A,B

As shown above, these sister units had the highest overall frequencies of Inadvertent ESF actuations from test and maintenance. In total, they are 50% higher than the next site (E&F) with plant A more than twice as high as the frequency for the next highest unit. The dates of occurrence for these events are shown graphically in Figures 6.1 (A) and Figure 6.2 (B) respectively. These figures also show the periods when the plant was shut down, and therefore the periods when no opportunity for events existed. by inspection of Figure 6.1, there is not an obvious trend towards a reduced frequency of events with time for A. The data associated with B are somewhat sparse so as to obscure any overall trend. The actual events (by LER number) and dates are summarized in Tables 6.1 and 6.2. In addition, these tables report (where identifiable) the type of corrective actions taken in response to the event.

The essence of the SALP assessments are as follows:

SALP (5/84-10/85): Maintenance is considered average overall, with an improving trend, and no major issues of concern were identified. For surveillance, the overall performance was average and improving, with no major issues. For Quality Programs, performance was average, with only document control being an issue of concern.

SALP (11/85-6/87): Minor weaknesses in work control were noted in maintenance, together with minor concerns associated with a lack of attention to

Figure 6.1, ESF Actuations, Plant A

* Occurred during plant startup

Figure 6.2, ESF Actuations, Plant B

Table 6.1, ESF Events, Plant A

<u>DATE</u>	<u>LER NUMBER</u>	<u>CORRECTIVE ACTION</u>
2/7/85	85-003	?
2/14/85	85-007	?
4/4/85	85-004	DISCIPLINE
11/2/85	85-059	DISCIPLINE
11/12/85	85-062	?
12/2/85	85-066	REVIEW/COUNSELING
12/27/85	85-067	PROCEDURE MOD
3/26/86	86-009	?
9/13/86	86-024	?
12/31/86	86-031	ADMIN PROCEDURE
1/27/87	87-002	?
2/2/87	87-003	?
6/12/87	87-017	PROCEDURE
8/21/87	87-021	DISCIPLINE/COUNSEL
12/26/87	87-022	COUNSEL/PROC REV
1/04/88	88-001	ADMIN PROCEDURE
3/09/88	88-009	COUNSELLING

Table 6.2, ESF Events, Plant B

<u>DATE</u>	<u>LER NUMBER</u>	<u>CORRECTIVE ACTION</u>
9/4/85	85-005	DISCIPLINE
9/6/85	85-006	TRAINING
12/5/85	85-013	?
7/23/86	86-018	DESIGN
8/23/86	86-020	COUNSELLING
9/19/86	86-021	DISCIPLINE
10/9/86	86-023	EQUIPMENT CHANGE
1/5/87	87-001	DISCIPLINE
12/7/87	87-012	TRAINING

detail on some poor work practices--no major concerns. Surveillance was noted as having some instances of late testing and missed fire watches--again, no major concerns. The Quality Programs need more management attention, due to occurrences of a lack of attention to detail, personnel errors and the occurrence of some "complicated" events.

SALP (7/87-8/88): Maintenance and surveillance (combined as one SALP area) reported safety systems as having failure rates far exceeding the national average. There is a complete failure to identify root causes, and no effective tracking process for repeat failures. The priority process is poor. The Engineering/Technical support is reported as having significant staffing and morale problems, with no management direction. "Keys issues" were not being addressed, and faulty information was provided to NRC. In the area of Safety Assessment/Quality Verification, many examples of management failure were cited, such as: failing to focus on corrective actions for equipment and design failures, ineffective involvement in safety reviews, poor communications, and a lack of attention to root causes of failures.

In summary, the first two SALP assessments were relatively benign. The last, covering July 1987 to August 1988, was very critical of all areas relating to the maintenance function. Of interest is the reported failures to address root causes, and the failure to establish effective priorities. Tables 6.1 and 6.2 show a tendency for this site to respond to events with "discipline" or "counseling". It is considered that simplistic responses like this shows a lack of real root cause study; it reflects the mind set that "all accidents are pilot error"--that is, broader underlying causes of error are ignored, with the result that they are not eliminated. The relatively high frequency of events, with no significant downward trend, is therefore not surprising.

Plants E & F

These sister plants had the second highest total frequency of events; the individual events are listed in Table 6.3 and 6.4. The patterns of the events as given in Figure 6.3 and 6.4 do not show any obvious trend because of the comparative sparseness of data; when combined, the two units show a relatively flat trend.

The associated SALP reports for this site are:

Figure 6.3, ESF Actuations, Plant E

Figure 6.4, ESF Actuations, Plant F

Table 6.3, ESF Events, Plant E

<u>DATE</u>	<u>LER NUMBER</u>	<u>CORRECTIVE ACTION</u>
2/8/85	85-017	DRAWING CONTROL
2/22/85	85-022	?
3/21/85	85-029	ADMIN, TRG*, COUNSEL
11/26/86	86-041	MECH FIX
1/20/87	87-007	TRG, PHYSICAL BARRIER
4/1/87	87-016	TRAINING
4/16/87	87-020	COUNSELLING
12/22/87	87-044	ADMIN CONTROL

*TRG - TRAINING

Table 6.4, ESF Events, Plant F

<u>DATE</u>	<u>LER NUMBER</u>	<u>CORRECTIVE ACTION</u>
2/13/85	85-008	PROCEDURE CHANGE
7/23/85	85-035	PROCEDURE CHANGE
7/30/85	85-036	INSTALLATION MOD
5/9/86	86-008	COUNSELLING
8/23/86	86-015	PROCEDURE CHANGE
10/17/87	87-018	COUNSELLING

SALP (5/84-9/85): In maintenance, there is reported a series of problems with procedures, both in their use and in their adequacy. Deficiencies in planning and follow-through of priorities; communications with other groups is poor. In summary, the assessment was concerned with an apparent decline in management controls. Surveillance was seen as having the same problems as maintenance, including continuing problems in personnel errors and weaknesses in planning and priorities. Regarding Quality Programs, overall management effectiveness was weak, with weaknesses in corrective actions, root-cause analyses, and policies and responsibilities.

SALP (10/85-11/86): Significant improvements were seen in management involvement for all three areas, though some deficiencies were seen in implementing corrective actions associated with surveillance.

SALP (11/86-3/88): Problems from lack of management attention are again reported for maintenance and engineering support. Management involvement in Surveillance and Quality Programs is considered effective, however. Particular problems include weaknesses in preventive maintenance programs, surveillance personnel errors and resolution of some technical issues. Strengths included levels of staffing, knowledge of personnel, and communications.

In summary, the three SALP assessments present a picture of a site partially improving, but with mixed degrees of management involvement through time. Corrective actions seem to show a tendency to correct or modify procedures or install barriers rather than rely on discipline or similar measures. It is noted that in these SALP assessments, the number of inadvertent ESF actuations were explicitly considered by the SALP panel, the only panel explicitly stating so.

Plant I

Plant I was out of operation for a large fraction of time during the period of analysis, and no SALP assessments were performed from 8/84 to 11/86 because of the prolonged outage. In the relatively short period of power operations, three events occurred in a three-month period as shown in Table 6.5 and Figure 6.5. The plant then entered a refueling outage in late 1987. Hence the time window for the occurrences of ESF actuations at power is limited.

Figure 6.5, ESF Actuations, Plant D

Table 6.5, ESF Events, Plant I

<u>DATE</u>	<u>LER NUMBER</u>	<u>CORRECTIVE ACTION</u>
12/22/86	86-043	ADMIN PROCEDURE
1/1/87	87-001	ADMIN PROCEDURE
3/13/87	87-006	TRAINING

The relevant SALP evaluation for Plant I was:

SALP (11/86-12/87): No major deficiencies were observed in the areas of management involvement for all relevant functional areas. Many pre-outage problems had been resolved, including work control, planning, staffing and problem solving. The one area identified as a concern was spare parts availability.

The responses to these ESF actuations from test and maintenance errors were administrative or technical changes. No corrective actions relied on discipline or counselling.

Plant C

Only two events, on adjacent days in 1985, were identified. These two events were not apparently connected. These are listed in Table 6.6 and shown in Figure 6.6.

The associated SALP reports for Plant C:

SALP (11/84-3/86): The assessment identified strong management involvement in maintenance and surveillance programs. However, Quality Programs did not portray a strong management involvement, with problems reported in interfaces between the site and the corporate staffs, and in procedures and their control. Strengths in maintenance include the increased use of predictive maintenance programs. In surveillance, strengths were the use of a procedures adherence program and the improved understanding and compliance of personnel.

SALP (4/86-8/87): Management involvement in maintenance was reported strong, with evident use of predictive maintenance, excellent control of maintenance activities and increased effectiveness of planning. However, management's involvement and support for surveillance were considered poor, with some inadequate or poorly written procedures allowed to impede work. Limited experience of the engineering staff, with no effective feedback from the field, was a major concern. Increased management attention for Quality Programs was reported, with improved definitions in individual responsibilities.

Figure 6.6, ESF Actuations, Plant C

Table 6.6, ESF Events, Plant C

<u>DATE</u>	<u>LER NUMBER</u>	<u>CORRECTIVE ACTION</u>
10/8/85	85-019	COUNSELLING
10/9/85	85-020	EQUIPMENT CHANGE

Plants D, G, H

The associated SALP reports for these sister plants were:

No events were identified for any of the three units over the period of analysis.

SALP (9/84-2/86): Management involvement in all relevant areas was considered strong; no major issues were identified in any of these areas. Minor concerns noted include a considerable backlog of station modifications and some maintenance-induced scrams. In the Quality Programs areas, some loopholes were seen in the "Q" list, indicating a possible lack of attention to detail, and some lateness in submitting NRC reports.

SALP (3/86-7/87): Management involvement was again reported to be high for all relevant areas. The maintenance area had implemented several aggressive improvement programs associated with preventive maintenance, procedures, and failure analysis. Surveillance procedures were subject to thorough reviews, causing them to be well-written and effective. Senior management was involved in any surveillance problems. The Quality Programs area demonstrated many initiatives, with a consistent presence of station management in planning, and a consistent allocations of resources. The station had adopted a program of Self Initiated Technical Audits. There was a reluctance to resolve some technical issues, however.

In summary, the SALP assessments present a "glowing" picture of excellence. With no events, no pattern of corrective actions can be seen.

Plant J

No events were identified in the period of analysis. It is noted that this plant was shut down from May 1986 to April 1987 at NRC's direction following the multiple-failures event in May 1986. This came close on the heels of a 5-month refueling outage starting in December 1985.

The associated SALP reports for Plant J were:

SALP (11/84-10/85): Generally, problems were identified with all relevant areas, with a systematic lack of attention by management being reported. In maintenance, no trending of performance was performed and no effective preventive maintenance was in place. Modifications were not reviewed for safety implications; limited corrective actions were performed. In surveillance, a lack of attention to detail led to some incorrect procedures, compounded by a lack of training. In Quality Programs, there was a failure to prevent recurring problems, a failure to implement corrective actions associated with known weaknesses, and an unwillingness to assign a high visibility to action items.

SALP (11/85-4/87): As noted above, for most of this period the plant was shut down. The SALP assessment concentrated on programmatic changes during the outages. In particular, a major increase in the level of management involvement was reported. This level of involvement communicated new expectations of performance. In maintenance, some problems remained in staffing levels, and there was a tendency to "analyze problems away" rather than fix them. In Quality Programs, problems remained, including a lack of attention to personnel errors.

SALP (5/87-5/88): Management involvement was reported to be continued extensive. New initiatives in maintenance and engineering analyses were effectiveness. The effectiveness of planning and use of preventive maintenance programs were significant. Organizational changes had increased the functional effectiveness of maintenance. In surveillance, procedural weaknesses had been identified and addressed; staffing was more than adequate. The engineering support staff were solution-oriented and open-minded. Effective redesign or upgrade programs were in place. Overall coordination was good. The Quality Programs area was seen as continuing to improve, with increased corporate management support leading to reduced turnover and improved morale. Deficiencies in quality activities were largely self-corrected and minor in nature.

In summary, Plant J appears to be a major "turn-around" plant in terms of the performance of its functional areas. By the end of the last SALP assessment cycle, all areas were considered strong.

6.1.3 Discussion

In most cases, there is a strong relationship between the occurrence of inadvertent ESF actuations from test and maintenance errors, and the SALP evaluations for areas associated with maintenance functions. For Plants E & F, the periods having higher frequencies of inadvertent ESF actuations from test and maintenance errors correspond with the periods of programmatic performance deficiencies--specifically the first third and last third of the analytical period are the periods having the higher frequencies of actuations. These are the times corresponding to marked deficiencies in the maintenance programs identified by the SALP analyses, and particularly, weaknesses in management effectiveness.

At the opposite end of the spectrum, the three units, D, G, & H had no relevant ESF actuations over the period of analysis. In turn, the SALP analyses present a picture of a highly effective maintenance programs with considerable management participation.

In between, Plant C had only two events in the period of analysis. This plant was assessed as generally having a strong involvement by management, with one or two isolated areas of weakness.

In the case of Plant I, it is difficult to draw inferences because of the prolonged outage periods. However, in a relatively short period of operation, there were three events reported. The SALP assessment for the operational period does not show either a very weak or a very strong maintenance involvement.

There appear to be two possible anomalies. The first is the first two SALP assessments at Plant A & B, the second is the "absence" of events at Plant J in early 1985. The SALP assessments of Plant A & B (5/84-10/85) and (4/85-6/87) are considered optimistic compared with personal reports from ex-maintenance staff of those plants for that period. The final SALP report (7/87-8/88) is considered more representative of the whole period of analysis. The reason for optimism in the earlier reports is not known.

Concerning Plant J, the absence of relevant ESF events was not expected for the initial period given the SALP assessment. It is recognized that this period of 11 months was when the plant's emphasis was strongly on continued power generation. This emphasis may have had a strong influence on the avoidance of error types which could have the

potential for causing a shutdown. However, because the plant was off-line for much of the period of analysis, the plant performance over an 11-month period is not considered as much of an indicator.

Of particular interest was the tendency of Plants A & B to report corrective actions in terms of "discipline" or "counselling". This can be considered a failure to investigate root causes in any detailed or systematic way--dismissing an event simply as "personnel error" has increasingly been considered an inadequate explanation for man-machine failures. The failure on the part of the plant to take any apparently effective steps to reduce the frequency of these events via changes in procedure or design indicates no effort for real root-cause removal is in place.

The use of discipline and its less punitive equivalent, counselling, are small examples of what has been termed "denial actions" (Westrum, 1988). Denial actions are characteristic of an organization that avoids recognition that a potential hazard exists and punishes anyone who may indicate otherwise. In Westrum's taxonomy, the alternatives to denial behavior are "repair actions" and "reform actions". Repair actions are where hazards are recognized and their immediate causes are fixed, and reform actions are where problems and hazards are recognized as symptoms of larger problems, which are then sought out and remedied. Westrum cites three types of actions as characteristics of three types of organizations. These are: pathological, calculative and generative. Pathological organizations exhibit suppression of "bad news", with isolation of the messengers and the message. Calculative organizations apply "by the book" solutions to problems and do not suppress the bad news; however, they do not take the opportunity to search out broader issues and learn larger lessons, as do generative organizations.

Of course, virtually all plants will use counselling or discipline sometimes in response to events involving personnel errors. A recent study by INEL (1989) of cause/correction codes taken from approximately 1,230 LERs in 1988, indicated that, on average, approximately 11% of LER events involved counselling or discipline. For LERs involving maintenance, between 11 to 15% of events were reported to have discipline/counselling as corrective actions, and for LERs involving personnel errors in maintenance, between 18 to 30% of events involved discipline/counseling. In summary, plants reported the use of discipline counselling as corrective actions in the range of 11 to 30% LERs. In this study, discipline was reported by Plants A & B in 56% of the events for which corrective actions can be identified from the report. This apparently excessive

reliance on discipline/counselling methods in itself acts as an indication of denial behavior occurring in at least part of the organization in this site.

The report cited above describes a parallel study to identify performance indicators from the causes and corrective actions reported in all LERs from whatever cause. This work involved coding and analysis of approximately 1,230 LERs reported by 109 units in the first six months of 1988. This analysis has focused on a specific combination of event types (inadvertent ESF actuations) involving specific causes (test and maintenance errors) with consideration of one specific corrective action (discipline/counselling). This unique combination is not explicitly considered in the cause/correction code study; rather, the elements are analyzed in aggregation (e.g., all maintenance errors).

This current study, by focusing on one class of LER events, provides a benchmark comparison for events of (more or less) equal significance between plants since not all LER events are of equal importance to safety. By blending all LER events, the results do not separate causes and corrections of less important and more important events. In addition, no reasons are given as to why the programmatic aspects of behavior may be important to safety. This conceptual relationship is important for NRC to understand what messages are being conveyed by the indicators.

In current NRC studies, the total number of ESF actuations is monitored as a performance indicator. The indicator proposed here is significantly different from that indicator in two respects: (1) many ESF actuations occur while the plant is shut down during I&C testing that have no safety relevance, and (2) many inadvertent ESF Actuations occur that do not involve errors in test and maintenance. By including all ESF actuations, however caused and at whatever power, the link to maintenance programs effectiveness is lost, both in terms of the overt relationship and the statistical ties to safety. For these reasons, NRC is expected to drop the total ESF actuation frequency as an existing indicator.

6.1:4 Future Steps

NRC should consider undertaking a trial program of using the frequency of inadvertent ESF actuations due to test and maintenance errors as a programmatic performance indicator of maintenance. Such a trial program should monitor the performance of a significant number of plants (say 25 to 30) for a SALP cycle for each plant, and then

compare the SALP assessment against the frequency of inadvertent ESF actuations calculated on a 3-month trailing average for periods with the reactor at power. A three-month trailing average provides a continuously updated measure representing the average number of events for the three most recent months. It is defined by:

$$M_0 = (M_{-1} + M_{-2} + M_{-3})/3$$

where M_0 is the current value of the three-month trailing average

M_{-1} is the number of events in the previous month

M_{-2} is the number in the second prior month, and

M_{-3} is the number in the third prior month.

At the end of each month, M_0 is recalculated by deleting the oldest month's value and adding the most recent month's value. Such a calculational method is used in other current NRC indicators.

In addition, NRC should evaluate the corrective actions reported by the plants to these events as a supplementary indicator. Specifically when discipline or counselling is greater than about 20-25% of the predominant corrective actions, the plant should be monitored more closely.

6.2 Thermal Performance Indicator

The work described earlier in this report led to gross heat rate being proposed as a promising leading indicator of nuclear plant safety trends. Gross heat rate, the electrical generation industry's conventional way of expressing the overall thermal efficiency of a thermal power plant, is defined as the ratio of the total thermal energy input to the plant, to its total electrical energy output during some interval of time. For nuclear plants, the thermal input is the heat energy produced by the reactor core. (Since heat rate is the algebraic reciprocal of thermal efficiency if consistent units are used, the two quantities are equivalent, and the terms will be used interchangeably below.)

The rationale for thermal efficiency as a trending indicator of nuclear plant safety is that over the long term, both safety and thermal efficiency depend on high-quality plant maintenance (broadly defined to include maintenance, and all other plant functions which help to preserve operability), so trends in plant safety should follow trends in thermal efficiency. In fact, thermal efficiency should be more sensitive to maintenance quality than safety, for two reasons. First, the non-safety-related portion of the plant continuously generates information on its own status during operation, while safety depends on either passive features (e.g., containment) or standby stems whose status is only known intermittently. Therefore, declining efficiency will presumably be noticed and corrected first. Second, for business reasons the licensee presumably places at least as much priority on thermal efficiency as safety. In other words, long-term neglect of the maintenance needed for the plant to fulfill its basic mission--generating electrical energy efficiently--may indicate even greater neglect of the nuclear safety functions, which are secondary to the plant's reason for being (however critical they may be from the NRC perspective). It is also well-established that reactor scrams and safety system challenges are mainly attributable to balance-of-plant (BOP) system and component failures. Thus it was suggested that gross heat rate would be a comprehensive indicator of maintenance performance in its broadest sense.

However, gross heat rate or thermal efficiency *per se*, while suggestive, has three deficiencies as an overall indicator of maintenance quality. First, heat rate is essentially meaningless when the plant is shut down. Second, a power plant has its maximum thermal efficiency at its design load, so extended off-design loading for reasons external to the plant (e.g., transmission grid problems or low power demand) could produce a misleading decline in efficiency. Third, the NRC is not currently collecting the data required to calculate heat rate. The ideal indicator would reflect thermal efficiency, deal realistically with shutdowns and load fluctuations, and preferably be based on data the NRC is already receiving in licensees' monthly operating reports.

Therefore, a reevaluation of the definition of gross heat rate was made to identify how the concept of plant performance efficiency could be best represented. This led to a reinterpretation in terms of Daily Power Loss (DPL), which is a measure of the difference between the maximum dependable electrical power output and the actual electrical power sent out, calculated on a daily basis. Second, the DPL values were plotted against the dates of occurrence of outages and power reductions, to observe the effects of equipment failures on this parameter. Third, an analytical process was

developed for calculating parameters associated with DPL that could be performed on a continuing basis.

6.2.1 Definition of Indicator

One of the two candidate indicators selected in Section 5 for further analysis was gross heat rate, defined to be the ratio of the thermal power output to the electrical power output. This is the definition used by INPO, which tracks this indicator. The underlying thinking behind this indicator is that it reflects trends in inefficiencies in the power conversion system (turbine generator, condenser, feed train). However, this definition has significant drawbacks, including the lack of sensitivity of the ratio to the range of changes in thermal inefficiencies, and the fact that the ratio changes significantly with changes in power level without changes in actual system inefficiencies. Thermal losses in the system are of two kinds: fixed and variable. The fixed losses include the thermal insulation losses (the steam and feed temperatures are constant) and mechanical (friction) losses. Variable losses include the idealized efficiency losses inherent in a two-phase thermal (Rankine) cycle, and losses related to mass flow rates. For any power level, the ratio of thermal to electrical power output will be comprised of the same fixed losses and a unique variable loss. Hence the ratio for the same fixed losses and a constant inefficiency rate (as a percent of power) will vary.

Because of this lack of constancy in the ratio, alternative interpretations of gross heat rate were proposed. These interpretations centered on trying to represent power or efficiency losses in ways that were reflections of maintenance effectiveness. These interpretations were:

1. Losses in Excess of Ideal Losses: In any real system, losses can be calculated that would result from perfect operation; that is, losses purely associated with the two-phase cycle. Actual losses in excess of these idealized losses can be considered (in principle) to be controllable. These include losses from steam leaks, lost insulation, etc. In practice, not all losses are truly controllable; for example, frictional flow losses cannot be reduced though they are in excess of the "ideal" losses. However, the relative fraction of the total "excess" losses that are controllable by maintenance are much higher than the remainder of the total losses, which are dominated by the thermal cycle losses.

For a system like power generation in a power plant, the best representation of thermal losses is a Rankine cycle (Holman, 1969). In a Rankine cycle, the ideal efficiency, η , is:

$$\eta = \frac{W_E}{h_1 - h_2}$$

where W_E is the external work, and h_1 , h_2 are the enthalpies of the steam entering the turbine, and the saturated water at the turbine exhaust pressure. If W_E represents the net electrical load (with appropriate units conversion), then a daily average value for η can be calculated. The actual daily thermal power output (W_T) is then required to calculate the excess losses ($Loss_{xs}$) by the following:

$$Loss_{xs} = (\eta W_T) - W_E$$

Hence knowing the daily average thermal and net electrical outputs, and the plant design parameters that define h_1 , h_2 (turbine entrance temperature and pressure, condenser inlet temperature), $Loss_{xs}$ can be calculated.

2. Monthly operating reports contain one piece of information related to thermal efficiency. This information is reported on a daily basis so it can be trended over the short term. The measure of interest is the average daily power level (ADP), defined as the quotient of total electrical energy output in MWh during the 24-hour period divided by 24 hours. Unfortunately, ADP is no more an ideal maintenance indicator than heat rate, because non-maintenance factors such as planned refueling outages and load changes can affect it. However, ADP has the advantage of tracking another maintenance-related parameter in addition to thermal efficiency, such as the incidence of forced outages of major equipment in the plant, because virtually any major equipment failure precipitates at least a power reduction if not a shutdown. Therefore, ADP--or some parameter closely related to it--should be an integrated indicator of both plant thermal efficiency and the availability of major equipment, and

hence a more sensitive indicator of trends in maintenance quality than thermal efficiency alone. The study analysts therefore concentrated on attempting to extract meaningful maintenance quality indications from patterns discernible in ADP data.

For several reasons which will be discussed below, it was found more convenient to analyze a parameter which is referred to as "daily power loss" (DPL) rather than working directly with ADP. Daily power loss is a normalized, inverted measure of average daily power level. It is defined by: $DPL = MDC - ADP$ where MDC is the plant's Maximum Dependable Capacity. DPL is equivalent to ADP in terms of information content, but has the following three advantages. First, higher DPL relates to poorer performance, which is consistent with the sense of the other indicators being considered. Second, the normalization by plant MDC eliminates two irrelevant plant-specific factors: plant capacity and design heat rate, so DPL is consistent from plant to plant. Third, the DPL waveform is symmetrical about the horizontal axis, which reduces the computation load by making the coefficients of all even-numbered terms in a Fourier series representation of the waveform equal to zero.

Given these alternatives, the Daily Power Loss was selected because: (1) it is simple, (2) it requires fewer assumptions, and (3) data are readily available.

6.2.2 Analytical Process

This section summarizes the analytical methods and data used in this analysis.

Analytical Methods

The approach taken in this study was to use techniques from system modeling, to explore the data as if the plant were a signal generator. In other words, the parameter DPL is treated as a signal representing the dynamic behavior of the power generation system. This approach was taken for the following reasons: (1) DPL data appear to display both periodic and stochastic elements; (2) DPL is the result of the reliability of many components, most of which are in series; and (3) it would provide a simple and

convenient analytical tool for future use by the NRC. The signal processing technique used for this study was Fourier analysis.

By way of review, Fourier analysis depends on the principle that any periodic mathematical function, no matter how complicated, can be represented as the sum of a possibly infinite series of sinusoidal terms. That is, if a complicated signal can be represented by a periodic function, it can be decomposed into a series of simple sine waves and cosine waves, each characterized by an amplitude, a frequency, and a phase angle. Thus Fourier analysis is a very useful tool for identifying and quantifying periodically occurring patterns in a noisy signal, because the frequency components which contribute significantly to the composite signal have relatively large amplitudes. The analysis involves determining the amplitudes and phase angles corresponding to each frequency.

The most convenient way to obtain the Fourier coefficients of a series of discrete equal-interval samples of a periodic function is the Fast Fourier Transform (FFT), a numerical method which lends itself to computer implementation. The FFT routine used to process the DPL data from the various plants was that available as part of MATLAB™, a commercial mathematical analysis software package running on an IBM-AT personal computer.

An example of the DPL value over the course of one year at one plant is shown in Figure 6.7. This example is taken from Plant S, Unit 1 for the year 1985.

Figure 6.8 shows the same chart, with summary descriptions added for the principal causes of changes in DPL.

The frequencies and their associated amplitudes for the data in Figure 6.8 are shown in Table 6.7. An additional output from an FFT analysis is a plot of the relative magnitudes for the various frequencies; this is called a Power Spectral Density (PSD) distribution. The PSD plot for the data in Figure 6.8 is shown in Figure 6.9. As can be seen, the dominant frequencies listed in Table 6.7 are correspondingly shown in Figure 6.9.

Figure 6.7, DPL Data, Plant S, 1985

Data from "Power lost data sheet"

Figure 6.8, DPL Data Plus Causes, Plant S, 1985

Figure 6.9, Power Density Spectrum, Plant S, 1985

Table 6.7, Plant S, 1985

1st Dominant Frequency	=	0.59E-2/day (169 day period)
Amplitude	=	2710
2nd Dominant Frequency	=	0.98E-2/day (102 day period)
Amplitude	=	1629
3rd Dominant Frequency	=	0.016/day (62.5 day period)
Amplitude	=	736

[Note: Amplitudes are in arbitrary units.]

Data Sources

Monthly operating reports were obtained for eight units for the years 1985-1987. These were selected to obtain a representative sample of plant designs, nuclear steam supply system (NSSS) suppliers, and "good" and "bad" plants. These reports identified the maximum dependable power level and the average daily power load. Major plant outages and forced power reductions are described in summary form. The plants analyzed are identified as plants S-Z.

6.2.3 Results

The results of the FFT analysis are summarized in Table 6.8, which shows the three highest frequencies for each plant analyzed in the study.

The evaluation of Daily Power Loss for the individual plants as well as a detailed description of the method is provided in Attachment 1. The results for each plant corresponding to Figures 6.7 and 6.9 and Table 6.7 are presented in that attachment. The plots corresponding to Figure 6.8 show that the principal causes of changes in DPL are (1) refueling outages, and (2) balance-of-plant equipment reliability. Other factors like fuel conservation and repairs to safety equipment also contribute but are less significant.

The FFT analyses of the power density spectrum are shown in Attachment A for each plant. In each case there appear to be frequencies identifiable for each period. However, since these were obtained from calendar-year data, they show phenomena associated with the window and the periodicity of the refueling outages. This can be seen when frequency plants for years involving refueling outages are synthesized, as in Figure 6.10. This kind of plot is the sum of the three dominant frequencies identified in the PSD. Figure 6.10 and Figure 6.11 are the plots for one plant in two adjacent years. As can be seen for the year involving the refueling outage (Figure 6.10), the frequency plot is characterized by a long period cycle, whereas in Figure 6.11, no such cycles are evident.

6.2.4 Discussion

At this point in the analysis, it is not obvious how to interpret the synthesized frequency plots. Many of the graphs are dominated by factors like the refueling cycle. However, for two plant periods, there are results not involving refueling outages. These are for

Figure 6.10, Combined Frequency Plot, Plant S, 1986

Figure 6.11, Combined Frequency Plot, Plant S, 1987

Table 6.8, Data for Cosine Plots of Yearly Frequencies

PLANT	YEAR	FIRST AMPLITUDE	SECOND AMPLITUDE	THIRD AMPLITUDE	FIRST FREQUENCY	SECOND FREQUENCY	THIRD FREQUENCY
S	1985	.2710 E+4	.1629 E+4	.7357 E+3	.5859 E-2	.9766 E-2	.1563 E-1
	1986	.7060 E+4	.8390 E+3	.5513 E+3	.3906 E-2	.7813 E-2	.1367 E-1
	1987	.1268 E+4	.1196 E+4	.9148 E+3	.1172 E-1	.1563 E-1	.7813 E-2
T	1985	.4124 E+4	.9345 E+3	.4503 E+3	.3906 E-2	.1172 E-1	.1758 E-1
	1986	.4780 E+3	.4128 E+3	.4691 E+3	.9766 E-2	.1367 E-1	.3711 E-1
	1987	.9530 E+4	.5378 E+3	.9543 E+2	.3906 E-2	.1367 E-1	.1758 E-1
U	1985	.5181 E+4	.2144 E+4	.2940 E+3	.3906 E-2	.1367 E-1	.2743 E-1
	1986 ⁽¹⁾	.1322 E+1	.1322 E+1	.1321 E+1	.1953 E-2	.3906 E-2	.5859 E-2
	1987	.1683 E+4	.7065 E+3	.4571 E+3	.9766 E-2	.3906 E-2	.1758 E-1
V	1985	.7081 E+4	.1221 E+4	.3832 E+3	.5859 E-2	.1172 E-1	.2539 E-2
	1986	.6969 E+4	.1616 E+4	.4817 E+3	.5859 E-2	.1367 E-1	.1758 E-1
	1987	.9686 E+4	.5013 E+3	.5102 E+3	.5859 E-2	.1367 E-1	.2148 E-1
W	1985	.2329 E+4	.2294 E+4	.1212 E+4	.3906 E-2	.9766 E-2	.2148 E-1
	1986	.7704 E+4	.2169 E+3	.1392 E+3	.3906 E-2	.1563 E-1	.2930 E-1
	1987	.4078 E+4	.5053 E+3	.1683 E+3	.5859 E-2	.2539 E-1	.1758 E-1

Table 6.8, Data for Cosine Plots of Yearly Frequencies

PLANT	YEAR	FIRST AMPLITUDE	SECOND AMPLITUDE	THIRD AMPLITUDE	FIRST FREQUENCY	SECOND FREQUENCY	THIRD FREQUENCY
X	1985	.1905 E+3	.1721 E+3	.1304 E+3	.7813 E-2	.1367 E-1	.2539 E-1
	1986	.5118 E+4	.3257 E+4	.1068 E+4	.7813 E-2	.3906 E-2	.1563 E-2
	1987 ^②	.2955 E+4	.1387 E+4	.4734 E+3	.7813 E-2	.1172 E-1	.1758 E-1
Y	1985	.1951 E+4	.1912 E+4	.1094 E+4	.3906 E-2	.9766 E-2	.1758 E-1
	1986	.1739 E+4	.2136 E+3	.1040 E+2	.5859 E-2	.3125 E-1	.4492 E-1
	1987	.3478 E+4	.1176 E+4	.8373 E+3	.5859 E-2	.9766 E-2	.1953 E-1
Z	1985	.4319 E+4	.1278 E+4	.3548 E+3	.5859 E-2	.9766 E-2	.2148 E-1
	1986	.4025 E+4	.3463 E+3	.3182 E+3	.5859 E-2	.2539 E-1	.4102 E-1
	1987	.8957 E+4	.1973 E+4	.4691 E+3	.5859 E-2	.9766 E-2	.1758 E-1
	1985						
	1986						
	1987						

NOTES: ① Operated for only 5 days in the calendar year

② Several days estimated due to quality of hardcopy data

Plant X, 1985 and for Plant T, 1986. The DPL plot for Plant X, 1985 is shown in Figure 6.12 and for Plant T, 1986, is shown in Figure 6.13; their equivalent PSD plots are shown in Figures 6.14 and 6.15, and their summed dominant frequencies are shown in Figures 6.16 and 6.17.

The essential difference between Figures 6.16 and 6.17 is that for Plant X there are small numbers of both large (>50%) and small (<50%) loss events, whereas for Plant T, there are a smaller number of large loss events but a larger number of small loss events. This is revealed in the plots of dominant frequencies, where Plant T shows both higher frequency content and slightly higher amplitudes.

Plant T has been found in other studies, including the inadvertent ESF actuations analysis described earlier, to have a weak maintenance program whereas Plant X has a strong maintenance program. It is tentatively considered that higher frequencies of small losses may indicate a tendency to fix problems by "band-aiding" (hence the small losses).

Further investigation is needed to confirm the nature of the DPL measure as an indicator, to explore its relationship to the maintenance function, and to develop a calculational process for DPL as an indicator for NRC use. This additional work is discussed more completely in Section 7.

6.2.5 Next Steps

Based on the results described above, the following suggestions are made for the next steps.

NRC may consider to undertake a trial program of using the frequency of inadvertent ESF actuations due to test and maintenance errors as a programmatic performance indicator of maintenance. Such a trial program should monitor the performance of a significant number of plants (say 25 to 30) for a SALP cycle for each plant, and then compare the SALP assessment against the frequency of inadvertent ESF actuations calculated on a 3-month trailing average for periods with the reactor at power.

In addition, NRC may consider the corrective actions reported by the plants to these events as a supplementary indicator. Specifically where discipline or counselling is

Figure 6.12, DPL Plot, Plant X, 1985

Figure 6.13, DPL Plot, Plant I, 1986

Figure 6.14, Power Density Spectrum, Plant X, 1985

Figure 6.15, Power Density Spectrum, Plant T, 1986

Figure 6.16, Combined Frequency Plot, Plant X, 1985

Figure 6.17, Combined Frequency Plot, Plant T, 1986

greater than about 20-25% of the predominant corrective actions, the plant should be monitored more closely for its effectiveness in performance root causes analyses.

For the DPL indicator additional development work is required. This additional work should focus on several issues:

(1) What alternative types of analytical processes could be used to capture the key characteristics of DPL; that is, its ability to trend and reflect the reliability of power generation in an objective and simple fashion. The use of Fourier transforms has merit because it is simple calculationally and it focuses on factors like the frequencies and durations of outages. However, other methods of analyzing DPL can be considered, such as moving-average integrals of power losses.

(2) What alternative "windows of analysis" are appropriate, and what "filtering" should be performed using the FFT method. For example, our experience to date suggests that the analysis should be performed between refueling or other extended outages. Separate analyses should be performed including and excluding reactor scrams from non-maintenance causes (e.g., operational errors).

(3) Hypotheses concerning the relationship between DPL and safety require testing. Possible hypotheses include: (1) frequent, small losses indicate the use of "band-aids" rather than substantive fixes; (2) the availability of BOP systems exhibit a periodicity as a result of a central tendency for component failures--lower frequencies represent aggregated lower failure rates as a result of more effective maintenance.

(4) How can the measures of DPL be expressed as an indicator that can be trended by NRC, and what would be an appropriate alert level.

7. Summary and Conclusions

Firstly, it is believed important to have a systematic basis for identifying and evaluating the potential usefulness and worth of indicators. By having such an approach, consideration can be given to "top issues" of interest, such as the frequency of maintenance, monitoring the process or the outputs of safety, and so on, and then indicators may be selected accordingly. Those indicators selected in this work for evaluation are unlikely to be uniquely "correct" in a general sense; however, the frameworks presented here would allow other analysts to make their own selections based on their own preference schemes. Use of the frameworks allows other analysts in the industry to express preferences with a common basis for discussion.

Most of the candidate indicators selected on the basis of this approach directly provide a connection to measures of safety, which provides a further logical basis for their development.

The experience in the data analysis indicates the qualitative sensitivity of indicators to interpretation of data items. In the course of evaluating the data, several cases were found where data were ambiguous and required interpretation by plant data experts. In some cases, data were already sparse and such interpretations may have a significant result in the evaluation of indicators. It is recognized that limitations in the data-collection process led to an inability to test some indicators because of an absence of data. Additional data-collection visits would have permitted additional tests but would have put further burden on plant personnel during a period of several NRC visits.

The limited number of associations of indicators and measures of safety is in itself not surprising. Output indicators (in the context of the process model) are relatively rare events. In the case of inadvertent ESF actuations, for example, three plants had no events in the timescale of concern. In itself, this may imply that the plants have very effective self-correcting systems; in most cases, similar events had occurred earlier than the period of analysis. Therefore the lack of a statistical relationship should not negate future consideration of an indicator.

In the case of inadvertent ESF actuations from test and maintenance errors, the frequency of such events is strongly associated with the quality of the maintenance programs as assessed in the narrative descriptions of SALP reports. It is explicitly and directly

associated with maintenance activities, and focuses on events when the reactor is at power. Corrective actions based on discipline or counselling significantly in excess of industry norms corresponds to denial behavior that is a symptom of organizations that are unwilling to actively pursue safety.

The plant performance-related indicator of Daily Power Loss is intended to capture the availability and reliability of the balance-of-plant equipment in an integrated manner. It does this because the largest cause of forced power reductions is associated with problems of balance-of-plant reliability. In most cases, failures of individual components are sufficiently rare for their failure rates not to act as indicators; the DPL measure indicator captures almost all such failures in total. Use of the fast Fourier Transform technology should allow the discrimination of higher or lower frequencies of failures and of longer or shorter durations of outage. It is recognized that additional work is required for this indicator, to improve its calculational process, to confirm its relationship with other maintenance-related measures, and to develop an indicator form for NRC use.

In summary, the use of a systematic approach to indicator selection has been effective in identifying programmatic performance indicators. This systematic approach permits analysts to select their own preferences and factors important in their own study. In this case, nine candidate indicators were selected from an initial list of 78, with two demonstrating a consistency of indication in statistical analyses. These two indicators were then subjected to additional analyses to provide a basis for explaining the relationships between the programmatic indicators and safety. In the case of inadvertent ESF actuations from test and maintenance errors, this development work is considered complete and ready for trial use. For the plant performance indicator, Daily Power Loss, additional work is required to refine the analytical process and to further explore the relationship between the indicator, maintenance activities and safety.

8. Bibliography

NRC Documents

An Investigation of the Contributors to Wrong Unit or Wrong Train Events, NUREG-1192, US Nuclear Regulatory Commission, Washington, D.C., April 1986.

Boccio, J.L., and M.A. Azarm, *Performance Indicator Program--Data Collection Work Plant (Draft)*, Brookhaven National Laboratory, Upton, NY, May 1988.

Bogel, A.J., et al., Analysis of Japanese-US Nuclear Power Plant Maintenance, NUREG/CR-3883, Battelle Human Affairs Research Centers, Seattle, WA, June 1985.

Licensee Event Report System, NUREG-1022, US Nuclear Regulatory Commission, Washington, D.C., 1983.

Loss of Integrated Control System Power and Overcooling Transient at Rancho Seco on December 26, 1985, NUREG-1195, US Nuclear Regulatory Commission, Washington, D.C., February, 1986.

Loss of Power and Water Hammer Event at San Onofre, Unit 1, on November 21, 1985, NUREG-1190, US Nuclear Regulatory Commission, Washington, D.C., January 1986.

Murphy, G.A., and J.W. Cletcher, II, Operating Experience Review of Power Operated Relief Valves and Block Valves in Nuclear Power Plants, NUREG/CR-4692, Oak Ridge National Laboratory, Oak Ridge, TN, October 1987.

NRC Rulemaking on Maintenance of Nuclear Power Plants, Proceedings of Public Workshop, NUREG/CP-0099, US NRC, Washington D.C., July 1988.

NRC Augmented Inspection Team (AIT) Reports, December 1985-January 1988, US Nuclear Regulatory Commission, Washington, D.C. (21 AIT reports reviewed).

NRC Information Notices, January 1986 to June 1988, US Nuclear Regulatory Commission, Washington, D.C. (Approximately 150 notices reviewed).

Olson, J., et al., *Development of Programmatic Performance Indicators*, NUREG/CR-5241, Battelle Human Affairs Research Center, Seattle, WA, October 1988.

Performance Indicator Program Plan for Nuclear Power Plants, US NRC Interoffice Task Group on Performance Indicators, US Nuclear Regulatory Commission, Washington, D.C., September 1986.

Report to Congress on Abnormal Occurrences, January-March 1987, NUREG-0090, Vol. 10, No. 1, US Nuclear Regulatory Commission, Washington, D.C., October, 1987.

Safety Evaluation Report related to the Restart of Davis-Besse Nuclear Power Station, Unit 1, following the event of June 9, 1985, (Docket No. 50-346) US Nuclear Regulatory Commission, Washington, D.C., June 1986.

Seigel, A.I., et al., *Front-End Analysis for the Nuclear Power Plant Maintenance Personnel Reliability Model*, NUREG/CR-2669, Oak Ridge National Laboratory, Oak Ridge, TN, August 1983.

Status of Maintenance in the US Nuclear Power Industry, 1985, NUREG-1212, US Nuclear Regulatory Commission, Washington, D.C., June 1986.

Stello, V., *Guidance on the Use of Performance Indicators*, EDO Announcement #30, US Nuclear Regulatory Commission, Washington, D.C., February 5, 1988.

Stello, V., *Memorandum to Chairman Zech and Commissioners Roberts, Carr and Rogers: Proposed Rulemaking for the Maintenance of Nuclear Power Plants*, US Nuclear Regulatory Commission, Washington, D.C., June 27, 1988.

Systematic Assessment of Licensee Performance (SALP), Inspection Report No. 86001, Palisades Nuclear Generating Station, Consumers Power Company, (Docket No. 50-255), US Nuclear Regulatory Commission, Washington, D.C., February 1986.

Transcript of Proceedings, Public Workshop for NRC Rulemaking on Maintenance of Nuclear Power Plants, US Nuclear Regulatory Commission, Washington, D.C., July 1988.

Vesely, W.E., et al., *Measures of the Risk Impacts of Testing and Maintenance Activities*, NUREG/CR-3541, Battelle Columbus Laboratories, Columbus, OH, November 1983.

Other Sources

Bauman, M.B., et al., *Survey and Analysis of Work Structure in Nuclear Power Plants*, EPRI NP-3141, Electric Power Research Institute, Palo Alto, CA, June 1983.

Bignell, V., et al., *Catastrophic Failures*, The Open University, Milton Keynes (UK), 1977.

Engel, R.J., et al., *The Causes of Nuclear Power Plant Unavailability (WCAP-11716)*, Westinghouse Electric Corporation, Pittsburgh, PA, January 1988.

Geiwitz, J., et al., *Handbook for Evaluating the Proficiency of Maintenance Personnel*, EPRI NP-5710, Electric Power Research Institute, Palo Alto, CA, March 1988.

Human Factors Review of Power Plant Maintainability, EPRI NP-1567, Electric Power Research Institute, Palo Alto, CA, 1981.

Inaba, K., *Observations and Recommendations on the Proposed Rulemaking for the Maintenance of Nuclear Power Plants*, Xyzyz Information Corporation, Canoga Park, CA, July 1988.

Kletz, T.A., *What Went Wrong?* Gulf Publishing Company, Houston, TX, 1985. *Nuclear News*, Monthly "Power" and "International" Sections, American Nuclear Society, La Grange, IL, June 1984-May 1988 (Approximately 31 relevant articles reviewed).

Poister, T.H., *Performance Monitoring*, Lexington Books, Lexington, MA, 1982.

Reason, J., *Human Error*, Cambridge University Press, New York, 1989 (In Press).

Utility-specific indicator reports (as available).

Westrum, R., *Organizational and Inter-Organizational Thought*, World Bank Workshop on Risk Management & Safety Control, Washington, D.C., 1988.