NASA HANDBOOK

NASA-HDBK-8739.18

National Aeronautics and Space Administration
Washington, DC 20546

Approved: 2008-04-29
Expiration Date: 2013-04-29

Procedural Handbook for NASA Program and Project Management of Problems, Nonconformances, and Anomalies

Measurement System Identification: Metric

APPROVED FOR PUBLIC RELEASE – DISTRIBUTION IS UNLIMITED

This page intentionally left blank.

DOCUMENT HISTORY LOG

Status	Document Revision	Approval Date	Description
-	Baseline	2008-04-29	Initial Release

This page intentionally left blank.

FOREWORD

This handbook provides basic guidance for the management and resolution of problems, nonconformances, and anomalies for a NASA program or project. Technical data management as described herein includes elements of three of the seventeen process requirements outlined in NPR 7123.1, NASA Systems Engineering Processes and Requirements: the technical data management process, the technical assessment process, and the decision analysis process. For a list of the documents containing requirements that the instructions herein are based upon, see NPR 7123.1; NPD 8700.1, NASA Policy for Safety and Mission Success; NPR 8715.3, NASA General Safety Program Requirements; NPR 7120.5C, NASA Program and Project Management Processes and Requirements; NPD 8730.5, NASA Quality Assurance Policy Program; NPR 8735.2, Management of Government Quality Assurance Functions for NASA Contracts; and NPR 7120.6, Lessons Learned Process. Complete references can be found in Section 2.

This handbook applies to NASA Headquarters and NASA Centers, including Component Facilities, and the Jet Propulsion Laboratory and service providers to the extent specified in their contracts with NASA.

This handbook is intended for use specifically by programs and projects that provide aerospace products or capabilities, such as space and aeronautics systems, flight and ground systems, technology demonstration, validation, and operations as specified within the respective program or project planning and subsequent implementation requirements documentation.

Requests for information, corrections, or additions to this handbook shall be submitted via "Feedback" in the NASA Technical Standards System at **http://standards.nasa.gov**. This handbook was developed by the NASA Headquarters Office of Safety and Mission Assurance.

Bryan O'Connor
Chief, Safety and Mission Assurance

April 29, 2008
Approval Date

This page intentionally left blank.

TABLE OF CONTENTS

LIST OF FIGURES

Process for Management of Problems, Nonconformances, and Anomalies

1. SCOPE

1.1 Purpose

1.1.1 This document provides guidance for managing the complete life cycles of problems, nonconformances, and anomalies. Program and project managers are the primary audience.

1.1.2 This handbook is primarily designed to provide a description of technical data management and support processes. The document also provides an overview of the problem management process and a means to track problems within a program or project. Finally, this handbook provides guidance on methods and techniques for assessing and analyzing problems and provides suggestions for reporting and documenting problems for recurrence control.

1.1.3 This document should be used by the program or project manager as the primary reference for establishing and implementing systems or processes for the management of problems, nonconformances, and anomalies.

1.2 Applicability

a. This handbook may be used by all NASA Centers, Programs, Projects, …, as a reference for establishing and implementing systems or processes for the management of any problems, nonconformances, and anomalies which NASA has lead involvement and control or has partial involvement with control over design or operations via U.S. internal or international partnership agreements. This document has no automatic exclusions for any program or project due to limited funding, responsibility, or involvement of NASA in the program or project. NASA involvement includes design, manufacture, or funding of instruments, systems, hardware, software, operations, and processing.

b. This handbook has been designed to be cited in contract, program, and other Agency documents as a reference for guidance.

c. Within this handbook, the word "shall" which normally indicates a mandatory requirement is NOT used within the text. However, in the Examples, the word "shall" indicates places where mandatory requirements should be used in contractual documents. The word "should" indicates a suggested implementation, the word "may" indicates an optional implementation.

2. APPLICABLE AND REFERENCE DOCUMENTS

2.1 Applicable Documents

2.1.1 General

The documents listed in this section contain requirements that the instructions herein are based upon as cited in the text of Section 4. The latest issuance of cited documents is to be used unless otherwise approved by the assigned Technical Authority. The applicable documents are accessible via the NASA Online Directives Information System at http://nodis3.gsfc.nasa.gov/, or directly from the Standards Developing Organizations or other document distributors.

2.1.2 Government Documents

NPR 7123.1	NASA Systems Engineering Processes and Requirements.
NPD 8700.1	NASA Policy for Safety and Mission Success.
NPR 8715.3	NASA General Safety Program Requirements.
NPR 7120.5C	NASA Program and Project Management Processes and Requirements.
NPD 8730.5	NASA Quality Assurance Program Policy.
NPR 8735.2	Management of Government Quality Assurance Functions for NASA Contracts.
NPR 7120.6	Lessons Learned Process.

2.1.3 Non-Government Documents

Fayyad, Usama, Gregory Piatetsky-Shapiro, and Padhraic Smyth. Knowledge Discovery and Data Mining: Towards a Unifying Framework: Proceedings of the Second International Conference on Knowledge, Discovery and Data Mining. KDD, 1996.

Piatetsky-Shapiro, Gregory, Knowledge Discovery in Real Databases: A Report on the IJCAI-89 Workshop, AI Magazine, Winter 1991, pp. 68-70.

Larose, Daniel T., Discovering Knowledge in Data. New Jersey: Wiley, 1995.

"Cross Industry Standard Process for Data Mining," August 2000 (accessed 29 March 2006). http://www.crisp-dm.org/Process/index.htm

Bao, Ho Tu. Knowledge Discovery and Data Mining Techniques and Practice (accessed 29 March 2006). http://www.netnam.vn/unescocourse/knowlegde/2-3.htm

"Data Mining Techniques." Electronic Textbook, StatSoft. StatSoft Inc. 1984-2003. http://statsoft.com/textbook/stdatmin.html#eda

Heckert, Alan and James J. Filliben. Engineering Statistics Handbook. NIST SEMATECH (accessed 29 March 2006). http://www.itl.nist.gov/div898/handbook/eda/section1/eda11.htm

XL Miner V.3. Cytel Statistical Software (accessed 29 March 2006). http://www.resample.com/xlminer/help/PCA/pca_intro.htm

"Self-Organizing Map." Wikipedia, The Free Encyclopedia. 2006 (accessed 29 March 2006). http://en.wikipedia.org/wiki/Self-organizing_map

"Coefficient of Determination." Wikipedia, The Free Encyclopedia. 2006 (accessed 29 March 2006). http://en.wikipedia.org/wiki/R-squared

"Curve Fitting Toolbox." The Math Works. 1994-2006 (accessed 29 March 2006). http://www.mathworks.com/access/helpdesk/help/toolbox/curvefit/ch_fitt9.html

3. ACRONYMS AND DEFINITIONS

3.1 Acronyms

ATP	Acceptance Test Procedure
CART	Classification And Regression Trees
CIL	Critical Items List
COTS	Commercial Off-The-Shelf
CRISP-DM	Cross-Industry Standard Process for Data Mining
DAR	Deviations Approval Request
DRD	Data Requirements Description
EDA	Exploratory Data Analysis
ECP	Engineering Change Proposal
ET	External Tank
FMEA	Failure Mode Effects Analysis
FQT	Format Quality Testing
FSE	Flight Support Equipment
GIGO	Garbage In Garbage Out
GSE	Ground Support Equipment
HA	Hazard Analysis
HDBK	Handbook
HPOTP	High Pressure Oxidizer Turbo Pump
LCC	Launch Commit Criteria
LLI	Limited Life Item
LLIS	Lessons Learned Information System
LRU	Line Replacement Unit
MIS	Management Information System
MPI	Management Performance Indicator
MTBF	Mean Time Between Failures
MTBR	Mean Time Between Repairs
MTTR	Mean Time To Repairs
NASA	National Aeronautics and Space Administration
NESC	NASA Engineering and Safety Center
NPD	NASA Policy Directive

NPR	NASA Procedural Requirements
NSRS	NASA Safety Reporting System
OERU	Off-Earth Replacement Units
OMRSD	Operations and Maintenance Requirements and Specifications Document
ORU	Orbital Replacement Unit
OSE	Orbital Support Equipment
PCA	Principal Components Analysis
POS	Probability of Sufficiency
PRACA	Problem Reporting and Corrective Action
RCAR	Recurrence Control Action Request
RSRM	Reusable Solid Rocket Motor
RTAT	Repair Turn Around Time
SDO	Standards Developing Organizations
SMA	Safety and Mission Assurance
SOM	Self-Organizing Maps
SPR	Significant Problem Report
SPRL	Significant Problem Report List
SRP	Standard Repair Procedure
SRU	Shop Replacement Unit
SSME	Space Shuttle Main Engine
TAT	Turn Around Time
TSE	Test Support Equipment
UA	Unexplained Anomaly
UCR	Unsatisfactory Condition Report

3.2 Definitions

Anomaly: An unexpected event, hardware or software damage, a departure from established procedures or performance, or a deviation of system, subsystem, or hardware or software performance outside certified or approved design and performance specification limits.

Assessment: Review or audit process, using predetermined methods to determine whether specified criteria are being met. Evaluates hardware, human action, software, procedures, technical and programmatic documents, and the adequacy of their implementation. Used interchangeably with "evaluation."

Confidence Interval: The interval computed around an estimated parameter, which expresses the probability of including the true population within its bounds.

Corrective Action: Action to bring nonconforming items into a state of conformance and to eliminate the cause of the detected nonconformity or changes to design processes, work instructions, workmanship practices, training, inspections, tests, procedures, specifications, drawings, tools, equipment, facilities, resources, or material that result in preventing, minimizing, or limiting the potential for recurrence of a mishap.

Corrective Action Plan: A document that addresses root causes for findings and the actions to correct specific individual problems as well as actions taken to correct any potential systemic or process problems in order to prevent recurrence. This plan includes designation of a schedule for completing the actions as well as designating the responsible party (or parties) assigned to perform the actions.

Correlation: A measure of the strength of the relationships among data.

Critical Item: A system or subsystem with a failure mode effects analysis (FMEA) criticality of 1, 1S, 2 (with a single point failure), or 1R (if it fails redundancy screens). (See definitions below of "criticality categories.")

Criticality (of a failure): A measure of the severity of a failure in relation to mission performance, hazards to material or personnel, and maintenance cost.

Criticality Categories: A criticality category classification is assigned to every identified failure mode for each item analyzed for all mission phases. Criticality categories are assigned to provide a qualitative measure of the worst-case potential consequences resulting from item failure. The criticality categories are defined below:

Category	Potential Effect
1	Single failure that results in loss of human life, serious injury to flight or ground personnel, or loss of a major space mission resource (for example, Space Shuttle, International Space Station, or Hubble Space Telescope).
1R	Failure modes of like or unlike redundant items, if all failed, could lead to criticality category 1 consequences.
1S	Single failure in a safety or hazard monitoring system that causes the system to fail to detect or operate when needed during a hazardous condition and leads to criticality category 1 consequences.
1SR	Failure modes of like or unlike redundant items in a safety or hazard monitoring system, if all failed, could lead to criticality category 1S consequences.
2	Single failure that results in loss of one or more essential mission objectives as defined by the program office without resulting in criticality category 1 consequences.

2R Failure modes of like or unlike redundant items, if all failed, could lead to criticality category 2 consequences.

3 All other failure modes.

Data Mining: Data mining is the nontrivial extraction of implicit, previously unknown, and potentially useful information from data.

Failure: The inability of a system, component, process, or crew to perform its required functions within specified performance requirements.

Nonconformance: A condition of any article, material, process, or service in which one or more characteristics do not conform to requirements specified in the contract, drawings, specifications, or other approved documents. Includes failures, defects, anomalies, and malfunctions. Used interchangeably with "discrepancy" and "noncompliance."

- (Critical Nonconformance) - A nonconformance that is likely to result in hazardous or unsafe conditions for individuals using, maintaining, or depending upon the supplies or services, or is likely to prevent performance of a vital Agency mission
- (Major Nonconformance) - A nonconformance, other than critical, that is likely to result in failure or to materially reduce the usability of the supplies or services for their intended purpose
- (Minor Nonconformance) - A nonconformance that is not likely to materially reduce the usability of the supplies or services for their intended purpose, or is a departure from established standards having little bearing on the effective use or operation of the supplies or services

Problem: Any circumstance that fits or is suspected of fitting one of the following categories:

- Failure, including conditions that would result in waivers
- Unsatisfactory condition
- Unexplained anomaly (hardware or software)
- Overstress or potential overstress of hardware
- In-flight anomaly
- Any nonconformance (hardware, software, or process) that has been shown by a trend analysis to need recurrence control

Process Control: Adhering to established processes and procedures without deviation:

- (Process Control, In-Control) - A process in which statistical measures indicate that only common causes of variation are present. Typically, process performance is consistently within 3 standard deviations of the process mean
- (Process Control, Out-of-Control) - A process in which statistical measures indicate that special causes (non-random) of variation are present. Typically, process performance varies beyond 3 standard deviations of the process mean

Repair Turnaround Time (RTAT): The period between the time an item is removed from the system for off-line repair and the time that it is returned in ready-for-installation condition. RTAT includes the time an item is waiting for available shop time, diagnosis, parts, hands-on work, test, and final inspection.

Risk: The combination of 1) the probability (qualitative or quantitative) and associated uncertainty that a program or project will experience an undesired event such as cost overrun, schedule slippage, safety mishap, or failure to achieve a needed technological breakthrough, 2) the consequences, impact, and severity, and 3) uncertainties in both the probability and consequence.

R-squared: A quantitative measure of the correlation or goodness-of-fit of a trend model to actual data.

Significance Level: The maximum probability of accidentally rejecting a *true* null hypothesis (a decision known as a Type I Error). Sometimes called alpha error. Computed as 1-(confidence interval).

Significant Problem: Any problem that is considered to pose a serious risk to safety or mission success.

Trend Analysis: The analysis of trends in data. A trend can be any identified direction, tendency, or pattern. Trend analysis generally utilizes statistical analysis to extract the trends and patterns from recorded data. Clustering approaches, discrimination approaches, and pattern-matching approaches can also be used. Examples of trend analysis are time trend analysis, correlation analysis, categorization analysis, regression analysis, and kernel analysis.

4. THE MANAGEMENT OF PROBLEMS, NONCONFORMANCES, AND ANOMALIES

The intent of this document is to outline a process by which all problems, nonconformances, and anomalies (as defined in Section 3.2, hereafter referred to as "problems") affecting safety and mission success are identified, fully understood, and appropriately resolved. The problem management process includes the elements of technical data management for problems (hereafter referred to as "technical data management") as shown in Figure 4.-1, along with the support processes data mining and trend analysis that focus primarily on technical performance while addressing program or project management characteristics.

Problem management begins early in program or project formulation and should continue in a disciplined manner throughout all program and project life cycle phases. A long-range view of the program or project's mission success criteria and open communication among all members of the team, including stakeholders, are essential to successful problem management.

Program or project problem reporting and corrective action (including risk management processing) are governed and documented with the specific responsibilities through NASA and Center Safety and Mission Assurance (SMA) organizations, as specified in NPD 8730.5, and subsequently through contracts by NPR 8735.2. Coordination of the processes with this document and the quality assurance functional requirements will be a necessary part of problem reporting, nonconformance, and anomaly processing.

Problem management can be broken into two components: technical data management and support processes. Problem management is illustrated in Figure 4.-1. Technical data management (Section 4.1) consists of procedures and actions to be followed when a problem arises. Those actions include: problem identification, problem assessment and investigation analysis, root cause determination, problem report, corrective action and problem resolution, problem recurrence control and lessons learned, and problem disposition and closure. The supporting processes (Section 4.2), including data mining and trend analysis, are discussed in greater detail in Appendices D and E.

4.1 Technical Data Management

The technical data management process focuses on managing technical data that is generated and used in technical applications. The technical assessment process acts as a check on progress and satisfaction requirements. The decision analysis process incorporates data collection to aid in decision making. Each of these processes plays an important role in problem management.

There are seven basic steps in technical data management that support the technical assessment and decision analysis process. They are:

- Problem identification (4.1.1)
- Problem assessment and investigation analysis (4.1.2)
- Root cause determination (4.1.3)
- Problem report (4.1.4)
- Corrective action and problem resolution (4.1.5)

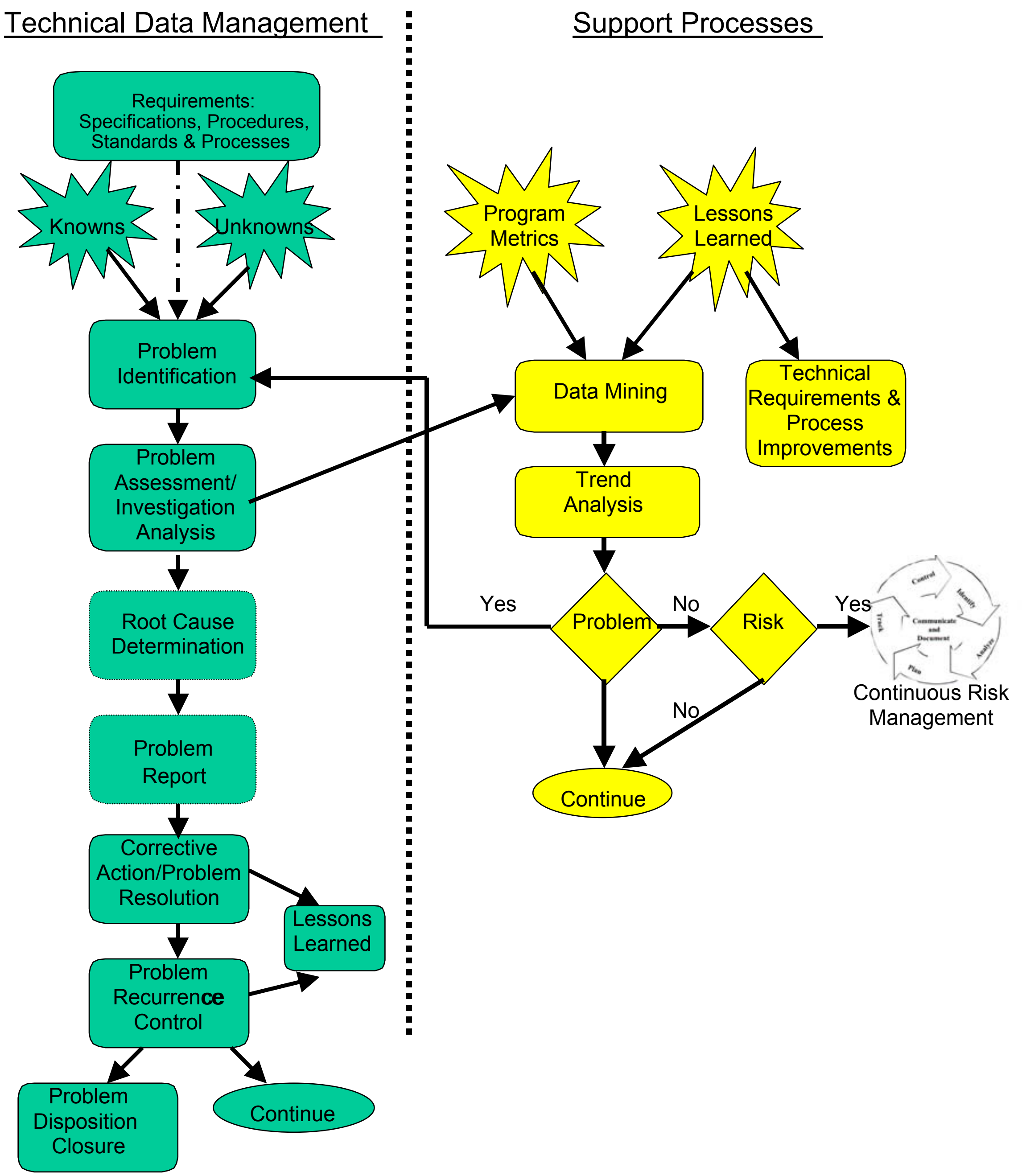


Figure 4-1 Problem Management Overview

- Problem recurrence control and lessons learned (4.1.6)
- Problem disposition and closure (4.1.7)

A sample technical data management plan can be found in Appendix A.

4.1.1 Problem Identification

Problem identification involves gathering information from the project so that problems can be described and measured. Identifying problems involves two activities:

- Capturing a statement of the problem
- Capturing the context of the problem

Finding and identifying problems within a program or project is a critical element of a program or project's mission success. Early identification of problems can lead to improved safety and quality and substantial savings of resources such as cost, schedule, and performance.

Once problems are identified and collected, they should be categorized. In the initial categorization, problems may be classified as “significant problems” or “less significant problems.” This aids in determining where to allocate resources most effectively. The term, significant problem, is defined in Section 3.2.

Significant problems should receive the most imminent attention; however, less significant problems should also be captured and resolved in a timely fashion. Both levels of problems may be further categorized to allow for the prioritization as well as for detection of clustering of problems. Patterns among problems can indicate areas that are not designed or managed well (should problems require local SMA or NASA Engineering and Safety Center (NESC) support, they should be ranked and sorted to provide a Significant Problem Report List (SPRL) so that the NESC or SMA can provide timely assistance). Significant problems should be reported to the appropriate management level not only as information, but also for the purpose of requesting allocation of resources for significant problem resolution. All problems should be documented, managed, and resolved as discussed in the following sections.

Many identified problems are reportable; the minimum criteria for defining what is reportable can be found in Appendix B, Section 1. Programs and projects should use Appendix B, Section 1 to determine whether a problem is reportable.

4.1.2 Problem Assessment and Investigation Analysis

Problem assessment and analysis is a process by which problems are examined in further detail to determine their extent. This process is a subset of the overall project systems engineering technical assessment process (NPR 7123.1). Examples of investigation analysis and failure analysis can be found in Appendix B, Section 2. Problem assessment is as described below.

Assessment of problems incorporates three basic activities:

- Evaluating the attributes of problems
- Classifying problems
- Prioritizing problems

Additionally, problem assessment may reveal how problems relate to each other and how to assess which ones are significant. A significant problem is any problem that is of the highest category by virtue of the problem’s impact on personnel safety or mission success. In analyzing problems, program and project managers should at least ask the following questions:

- How soon do we need to act on this?

- What will happen if we do not act on this?
- How does this problem compare with other problems?

Early discussion of potential problems and their resolutions facilitates the optimal resolution identification later.

Before prioritizing, the problems need to be classified or grouped with similar problems. There are several purposes for classification, which include the following:

- To understand the nature of the problems and to group related problems in order to build more cost-effective mitigation plans
- To identify problems that are equivalent and combine them as appropriate
- To produce evidence of the "clustering" of problems around a specific area. This in turn may be evidence of systemic problems within that area
- To track and monitor problems by various elements of the program
    - For example, a functional area such as financial or safety may want to concentrate on the subset of problems within their functional area to assure that all these problems are adequately resolved

Prioritization provides a mechanism for sorting through a large number of items and determining which are the most important and, therefore, should be dealt with first. Several methods of prioritization exist and are either quantitative or qualitative.

4.1.3 Root Cause Determination

Root causes of problems need to be defined completely so that adequate corrective action can be implemented or an approved standard repair put in place. An example of guidance for root cause determination can be found in Appendix B, Section 3.

4.1.4 Problem Report

Problem reporting can take any of several forms including, but not limited to, formal reports, presentations, or memos. Selection criteria may vary across NASA programs and within programs by management level.

The NASA Safety Reporting System (NSRS) is one example of a problem reporting system used within the Agency. More information on how to report safety problems can be found in NPR 8715.3. An example of guidance for a problem report Data Requirements Description (DRD) can be found in Appendix B, Section 4.

4.1.4.1. <u>Problems Reported by Program or Technical Discipline</u>

Program managers, project managers, and engineers have an additional communication channel: the Technical Fellow. Problems that are considered to be significant enough to demand immediate attention by higher levels of management authority may be reported to any Technical Fellow. These problems will be evaluated for significance and, if found significant, the report forwarded will be identified as a Significant Problem Report (SPR).

4.1.4.2. Hazardous Problems, Anomalies, and Nonconformance Conditions

SPRs will focus on hazardous problems, anomalous conditions and nonconformances relating to flight hardware, flight software, launch system, facility systems, ground support equipment (GSE), and test operations, and focus on indicators reflecting institutional factors that may affect the program performance (safety and quality), schedules, or milestones.

4.1.4.3. Outputs - SPR

For each significant problem identified, the originating office will propose actions to eliminate the problem or mitigate a recurrence of the problem. Completion effective dates for the corrective action will be included in the SPR Summary.

4.1.4.4. SPR Format

An example of a SPR Template can be found in Appendix C.

4.1.4.5. SPR Summary

The SPR summary will track the progress of actions for addressing the significant problem and will serve as the official record for reducing the significance level and to track resolution. The originating office will close items from the SPR when deemed appropriate, provided no objection is voiced from an appropriate Technical Fellow, SMA, or the NESC.

4.1.5 Corrective Action and Problem Resolution

Corrective action and problem resolution requires the following steps:

- Develop a plan of action
- Assign the problem to a person capable of resolving the problem
- Determine the approach to resolve the problem and implement the plan

Problem resolution efforts will involve investments or tradeoffs and, as a result, require careful consideration. Once the plan of action is developed, and each problem is identified and assigned to a person capable of resolving the problem, it is that person's responsibility to determine the approach for problem resolution. Problem resolution efforts will involve an investment or tradeoff and will require careful consideration.

- If the problem was originally an accepted, identified risk, the problem resolution action will activate the contingency plan originally described in the risk planning element of the program or project's risk management plan
- If the problem did not arise from an accepted or identified risk, a corrective action plan will need to be developed

In either case, it is important to record the action plan within the problem management documentation. A description of all identified problems should be included in the program documentation that feeds into the data mining support process. Problem resolution is a subset of the overall project systems engineering decision analysis process (NPR 7123.1).

An example of corrective action within the example problem report can be found in Appendix B, Section 4. Problem Resolution is as described above (no example provided in Appendix B).

4.1.6 Problem Recurrence Control and Lessons Learned

Problem recurrence control involves capturing lessons learned during the process of problem resolution, ensuring the communication of these lessons across the entire Agency, and implementing corrective actions to ensure recurrence control. Lessons learned are captured in the Agency's Lessons Learned Information System (LLIS), located at http://www.nasa.gov/offices/oce/llis/home/. It is vital to enable the documentation of problem information so that there is a historical record of problems that have occurred. This serves several purposes:

- Allows for easier reporting of problems, significant and other
- Supplies lessons learned
- Documents progress in problem resolution
- Enhances the visibility of problems

An example of guidance for problem recurrence control can be found in Appendix B, Section 5. Lessons learned examples can be found on the LLIS website. NPR 8715.3 contains requirements for developing safety lessons learned. NPR 7120.6 contains requirements for capturing lessons learned.

4.1.6.1. Recurrence Control Concept

Recurrence control is preventative action, beyond remedial action, taken to preclude or minimize the recurrence of a problem in existing and future hardware or software. Examples of problem recurrence control may include:

- Changing designs, procedures, or processes in hardware and software
- Revising procedures and standards
- Processing engineering change orders of drawings
- Updating policy
- Training
- Tooling
- Scheduling maintenance

4.1.6.2. Recurrence Control Inputs

There are many useful sources of information that provide input for problem recurrence control including:

- Trend analysis results and output
- Data mining results and output
- Significant problem reporting output
- Risk mitigation trends (successes and failures)
- Problem resolution process inputs

- Problem resolution outputs

4.1.6.3. Recurrence Control Outputs

The output of a recurrence control process is a Recurrence Control Action Request (RCAR). This is a request initiated by SMA to responsible organizations to investigate a problem and identify the root cause and actions necessary to prevent recurrence. A RCAR is used to record the results of the investigation, to justify the rationale for not implementing a corrective action if appropriate, or to document actions taken to implement the corrective action.

4.1.7 Problem Disposition and Closure

After the appropriate steps have been taken to resolve the problem, it is important to properly disposition the problem. This step aids in tracking the problem through its life cycle to completion. An example of guidance for the disposition and closure of problems can be found in Appendix D, Section 6.

4.2 Support Processes

Implementing analysis tools and techniques for problem management provides a structured methodology to determine the best way to attain a desired result. Such analyses for problem resolution refer to the broad quantitative field arising from operations research and statistics that deals with modeling, optimizing, and analyzing decisions made by individuals, groups, and organizations. These analyses assist decision makers in making better and more informed decisions in complex situations under a high degree of uncertainty. The quality of the decisions is measured by their expected consequences, the uncertainty of the consequences, and the stated preferences of the decision makers. The two tools presented herein are data mining and trend analysis.

4.2.1 Data Mining

NASA has tasked all programs and projects to perform trending as one method to uncover adverse data patterns. The NESC has been tasked with performing independent trending across NASA programs and projects. The NASA culture provides a large degree of autonomy and independence for each individual program or project. As a result, a common database of pertinent information that should be reviewed in order to identify trends does not exist. NASA is not alone in this predicament; it has been estimated that 80% of all corporate data is unstructured. Therefore, some electronic mechanism to extract information from diverse data sources is required. Data mining fulfills this requirement.

Data mining can be described as the nontrivial extraction of implicit, previously unknown, and potentially useful information from data. This differs significantly from querying data where a querying language is used to ask multiple questions of the data. The distinction is that queries are useful when specific information is sought, whereas data mining is necessary when discovery of unknown, pertinent information is the goal.

Data mining is not the only component of an optimal solution to identify precursors to future problems. Data mining is merely the first step. Once the data mining effort discovers something, the subject matter experts are required to determine if the “something” actually constitutes a potential problem. The discovery of similar events in multiple sets of data may not be an indicator of a future problem. In fact, the "blind application of data mining methods (rightly criticized as 'data dredging' in the statistical literature) can be a dangerous activity easily leading to discovery of meaningless patterns" (Fayyad, et al.). Only the subject matter experts can determine which discoveries require further attention. However, the use of domain knowledge experts initially can severely limit discovery (Piatetsk-Shapiro, et al.); therefore, data mining should be the first step in the overall trending process.

Data mining can be automated with available commercial off-the-shelf (COTS) software as well as tools developed within NASA. Data mining includes various algorithms and methodologies for detecting "data behavior" patterns useful in formulating strategies. These techniques can include, but are not limited to, cluster detection, memory-based reasoning, statistical analysis, market basket analysis, genetic algorithms, link analysis, decision trees, and neural nets. Details on data mining theory can be found in Appendix D. Requirements that discuss capturing data can be found in NPR 7120.5C.

4.2.2 Trend Analysis

Trend analysis is an element of engineering investigation that provides continuous review of program factors. Trend analysis has two prime characteristics: investigation of actual events and comparative assessment of multiple events. Trend analysis is applied to program characteristics that vary in relation to time, sequence, or element performance. Trend analysis results are used to evaluate the operations of a program and its component systems by assessing past performance to establish baselines for current and future performance. When a valid trend exists, the accuracy of the analysis will increase as more time or event data are collected.

Trend analysis is also used to discover and confirm correlations between diverse factors. The degree to which trend analysis approaches can be applied to the elements of NASA's programs will vary and be based on many factors. However, it is important to avoid the situation where the data was available but not analyzed and, therefore, not presented for decision in a comprehensive manner.

Trend analysis is defined as the search for patterns and subtle relationships in data in order to infer relative data behavior. In a comparative study of the parts of a product or system and the tendency of a product or system to develop in a particular direction over time, it is a generic reference to any longitudinal (time series) analysis of data and can be applied to any industry. It widens the temporal scope of knowledge-based systems. NASA trend analysis comprises four interrelated elements:

- Performance
- Problem
- Supportability
- Programmatic

Analyses of these types can be found throughout the engineering community; however, organizing trend analysis into these specific groupings is a NASA-unique approach. Examples of trend analysis can be found in Appendix E. Requirements for performing trend analysis can be found in NPR 7120.5C.

APPENDIX A. Technical Data Management Plan

A.1 Overview

A technical data management plan is needed for documenting the approach, methodology, and responsibilities to implement a problem management program.

A.2 Technical Data Management Plan Template

The following outlines the key elements of a Technical Data Management Plan. These elements are suggested and include elements often found in a program or project plan (specifically section 2). Those items may be referenced to the program or project plan.

1 Introduction
- 1.1 Objective
- 1.2 Purpose
- 1.3 Document Organization
- 1.4 Summary of Technical Data Management Plan Information

2 Program or Mission Summary
- 2.1 Mission Description
  - 2.1.1 Mission Statement
  - 2.1.2 Mission Need
- 2.2 System Description
- 2.3 Requirements
  - 2.3.1 Summary Requirements
- 2.4 Integrated Schedule
- 2.5 Management Organization
- 2.6 Roles and Responsibilities

3 Technical Data Management Approach
- 3.1 Introduction
- 3.2 Process Overview
- 3.3 Processes and Tasks
  - 3.3.1. Problem Identification
  - 3.3.2 Problem Assessment and Investigation Analysis
  - 3.3.3 Root Cause Determination
  - 3.3.4 Problem Report
  - 3.3.5 Corrective Action and Problem Resolution
  - 3.3.6 Problem Recurrence Control and Lessons Learned
  - 3.3.7 Problem Disposition and Closure

4. Detailed Technical Data Management Procedures
- 4.1 Technical Data Management Documentation
  - 4.1.1 Forms
  - 4.1.2 Documentation Process
- 4.2 Technical Data Management Implementation
  - 4.2.1 Roles and Responsibilities
  - 4.2.2 Reporting

APPENDIX B. Technical Data Management Examples, Guidance and Criteria

B.1 Minimum Criteria for Reportable Identified Problems

For a NASA nonconformance (hardware, software, or process) on the ground to be considered Problem Reporting and Corrective Action (PRACA) reportable, one of the following applicability criteria should be satisfied by the hardware, software, or process exhibiting the nonconformance:

- Flight (including qualification and proto-flight hardware, software, or processes)
- Flight spare
- Flight Support Equipment
- Orbital Support Equipment
- Acceptance or Qualification Test Article
- Ground Support Equipment
- Fleet Leader
- Flight Software
- Any process related to the above

Once the applicable item (hardware or software) has reached one of the following manufacturing or development levels, its design is considered sufficiently mature for the item to be considered for PRACA reportability:

- First and subsequent instance of acceptance or qualification testing has started for the Orbital Replacement Unit, Line Replacement Unit, Shop Replacement Unit, or designated configuration item
- First and subsequent instance of primary or secondary structure qualification testing has started
- For GSE, after it has been accepted by the program
- For software, after it has been accepted by the program or project
- Any process accepted by the program or project is reportable at any time

B.2 Example of Investigation and Failure Analysis Requirements

- Perform and provide technical direction for an investigation to research all applicable data, records, and telemetry to fully characterize the problem
- Investigate the history of the hardware, software, or process to determine the following:
  - Previous occurrences of the anomaly
  - Similar or related problems on other hardware or software

    - Other problems that may have contributed to the specific nonconformance being investigated
- Assess the impacts of the problem upon other planned applications of the hardware or software or for similar hardware, software, or for the process
- Assess the adequacy of any interim corrective actions implemented to ensure that no other actions are required until permanent corrective action and recurrence control can be implemented or determined unnecessary
- Assess the problem history and any previous corrective actions implemented to ensure that generic design, fabrication, or other issues do not exist across the rest of the hardware fleet, with the software, or with the process
- Develop appropriate failure analysis, fault trees, and troubleshooting plans as required and document them in the problem report
- Determine if investigation activities may be terminated and if a failure analysis, fault tree, or troubleshooting plan are not necessary based on satisfying one of the following criteria:
    - The failure trend has been established, documented, and is understood without the need for further analysis
    - Performing additional investigation is not considered cost effective when weighed against the failure insight or risk reduction gained by further investigation
    - The governing program board has directed cessation of failure analysis
- Ensure that any system-level, vehicle-level, or other integration impacts have been adequately identified, analyzed, and resolved

B.3 Example of Guidance for Root Cause Determination

- Root cause should be investigated such that adequate corrective action can be implemented or an approved standard repair is in place
- The root cause investigation may be terminated once the problem has been isolated to the lowest appropriate level to support identification and implementation of effective corrective actions and recurrence controls based on analysis or engineering judgment. If it is determined that identifying the root cause is not necessary or feasible, rationale for terminating the root cause analysis should be documented in the problem report and should be consistent with one of the following criteria:
    - A limited life item designated as a wear-out item became degraded or failed after it has surpassed its assigned limited life parameter
    - Performing additional root cause analysis is not considered cost-effective when weighed against the failure insight or risk reduction gained by knowing the root cause
    - The governing program board has directed cessation of the root cause determination
- If the root cause cannot be identified with high confidence, the problem report should clearly document that the problem is an "Unexplained Anomaly" (UA), with the problem

report also documenting which of the possible causes, if any, are considered the most likely cause of the problem

- Where there is high confidence that the root cause of the problem has been identified but that its determination cannot be proved conclusively, the problem report should document why NASA is confident in the root cause identification and that there is no need to consider the problem a UA
- The problem report should include adequate documentation (for example, fault trees, troubleshooting procedures, and analysis results) for each possible root cause investigation.

B.4 Example of Guidance for Problem Reporting and Corrective Action

DATA REQUIREMENTS DESCRIPTION (DRD)

1. **PROGRAM**: CEV
2. **DRD NO.:** CEV-S-013
3. **DATA TYPE**: 1/3*
4. **DATE REVISED**: 04/25/2008
5. **PAGE**: 1
6. **TITLE**: PRACA Data

7. DESCRIPTION/USE: The Contractor shall report, evaluate, track, trend, process, and disposition CEV-related problems (such as hardware, software, firmware, test, operational process, and other as defined in Section 13.0) using the centralized, closed-loop NASA Constellation PRACA data system. A problem is defined as any circumstance that fits, or is suspected of fitting one of the following categories:

- Failure, including conditions that would result in waivers
- Unsatisfactory condition
- Unexplained anomaly (hardware or software)
- Overstress or potential overstress of hardware
- In-flight anomaly
- Any nonconformance (hardware, software, or process) that has been shown by a trend analysis to need recurrence control

8. **DISTRIBUTION**: As determined by the Contracting Officer

* Except for required problem disposition approvals, all data entered into the NASA Constellation PRACA data system are Type 3. All problem dispositions, with approval authorities as defined in CxP 70068, "Constellation Program Problem Reporting, Analysis, and Corrective Action (PRACA) Implementation Methodology," are Type 1 and require concurrence within the Constellation PRACA data system before they are considered official.

9. **INITIAL SUBMISSION**: Per Data Requirements Matrix

10. **SUBMISSION FREQUENCY**: Initial entry of problems (as identified in Section 13.1 below) into the closed-loop NASA Constellation PRACA data system is required within one working day of discovery. The goal should be that all problems would be reported on the same day and preferably the same shift that the problems are discovered. Updates are required as additional data becomes available to document progress in working assigned actions associated with problem analysis, resolution, and disposition.

11. **REMARKS**: No remarks with this example.

12. **INTERRELATIONSHIP**: No interrelationship with this example.

13. **DATA PREPARATION INFORMATION**: No data preparation information with this example.

13.1 **SCOP**E

A. **APPLICABLE TIMEFRAME**:

The Contractor shall commence use of the centralized, closed-loop NASA Constellation PRACA data system to report, evaluate, track, trend, process, and disposition CEV-related problems once any one of the following criteria are satisifed for an "applicable item:"

a. The baseline of the item exhibiting the problem has been accepted by NASA

b. The Government has formally accepted (Form DD-250) the item exhibiting the problem

c. Acceptance or qualification testing of the item has started. This includes first and subsequent instance of primary and secondary structure qualification

d. For GSE or Test Support Equipment (TSE), the item has been accepted or approved for use by the program

   Note: Problems that occur during acceptance or qualification testing of GSE and TSE are exempt from the NASA Constellation PRACA process as long as the item has not been approved for use by the program

e. For software, formal quality testing or equivalent has commenced

f. Integrated system and element testing has commenced, even when subsystem items and integration simulation hardware or software may not have been formally accepted by the Government

B. **PROBLEM SCOPE:**

The types of problems that the Contractor shall report into and process using the centralized NASA Constellation PRACA data system include, but are not necessarily limited to, any nonconformance, noncompliance, and variation to any requirement and specification (such as design, safety, configuration, operation, or test) for any of the elements listed in Section 13.1C, "Applicable Items."

C. **APPLICABLE ITEMS**:

The types of hardware, software, firmware, and other items which require the use of the NASA Constellation PRACA Data System to report and process problems include, but are not necessarily limited to:

a. Flight hardware and spares. Includes, but is not limited to, Off-Earth Replacement Units, LRUs, and shop replaceable units

b. Flight-like hardware

c. Firmware

d. Flight software

e. Test and simulation hardware and software

Note: Limited to simulation hardware and software used to 1) accept or qualify flight hardware or software; 2) develop and verify test, acceptance, qualification, or flight operational procedures; and 3) train and certify flight controllers, astronaut crews, or mission support engineers

f. Flight Support Equipment

g. Operational Support Equipment

h. GSE

i. TSE

j. Shop Aids and Tools

Note: Limited to those problems that occur while the shop aid or tool is being used to perform activities involving flight or flight-like hardware, software, and firmware

k. Qualification hardware

l. Flight-leader hardware, software, and integrated systems

m. Protoflight hardware

n. Prototype hardware

Note: Limited to prototype hardware that has the same part number as flight or flight-like hardware

o. Facility and flight, flight-like, flight-leader, qualification hardware interfaces (such as platforms)

p. Extra-planetary equipment and facilities

q. Payloads

r. Operational and programmatic requirements

s. Ground software

13.2 **APPLICABLE DOCUMENTS**:

CxP 70068, Constellation Program PRACA Implementation Methodology

NESC Taxonomy Working Group Report, dated 1/20/2006

13.3 CONTENTS:

The NASA Constellation Program Office will develop, provide, and maintain a centralized Constellation PRACA Data System that will provide an electronic method for the Contractor to enter, report, track, trend, fully document, and manage workflow associated with the investigation and resolution of problems (refer to Section 13.1) associated with all Constellation-related programs, projects, operations, and other activities. The system will be available to all authorized users in real-time to support problem-related activities including, but not limited to, initial problem report generation, data updates, data query, report generation, trend analysis, action assignment, action follow-up, trending, and data mining.

The NASA Constellation PRACA Data System will facilitate, where practical, entry of data and use of the system by personnel in the field, such as personnel discovering a problem at a point of manufacturing, test, inspection, maintenance, or operation. In addition, the NASA Constellation PRACA Data System will incorporate, where feasible, the ability to automatically input part or serial number data using radio frequency identification or bar-code markings.

The NASA Constellation PRACA Data System will collect and the Contractor shall input into the NASA Constellation PRACA Data System various data fields consistent with Appendix A of the NESC Taxonomy Working Group Report (1/20/2006) to support the following:

- Documentation of initial problem and contributing environment conditions
- Documentation of the problem investigation (such as review, analysis, testing)
- Documentation of the root cause(s) of the problem (or rationale for why the root cause(s) cannot or will not be identified)
- Documentation of the remedial action(s) taken in direct response to the problem to allow continued activities or to safe the problematic item or activity
- Documentation of the corrective action(s) taken to resolve and prevent recurrence of the problem, or rationale for why corrective action(s) is not or will not be taken
- Actions assigned and responses from actionees to accomplish any of the activities associated with the problem investigation and resolution
- Information related to the problematic hardware or activity to allow for trending and querying of the data, such as, but not limited to, part numbers, serial numbers, part criticality, associated system or subsystem, type of hardware or software (such as flight, simulation), type of problem (such as operational, design, test)
- Identification of individuals involved in the reporting, investigation, and disposition of the problem, including individuals approving implementation of remedial actions, corrective actions, and disposition and closure of the problem
- Other information required by the PRACA Data System to allow use of the data to support

future real-time problem investigation and resolution, as well as future trending and analysis

13.4 **FORMAT**: The Contractor shall submit problem and supporting data directly into the NASA Constellation PRACA Data System as required by that system.

13.5 **MAINTENANCE**: Contractor-proposed changes to this DRD shall be submitted to NASA for approval. Complete re-issue of the document is required.

B.5 Example of Guidance for Recurrence Control

- Implement and provide technical direction for associated recurrence controls for problems, with the recurrence controls considered, implemented, and documented within the problem report
- Ensure through trend analysis that the recurrence controls implemented were effective in preventing or minimizing the recurrence of the problem
- Coordinate problem resolution activities to ensure that appropriate flight products (flight rules, operating procedures, malfunction procedures) and process-related documents (test procedures, maintenance documents) are updated when necessary to provide recurrence control associated with a problem report
- For recurrence controls not fully developed or implemented (training or process issues), open work with respect to recurrence control implementation should be tracked in a closed-loop corrective action system and referenced by a unique identifier number in the problem report

B.6 Example of Guidance for the Disposition and Closure of Problems

Essentially, a problem is typically resolved in one of three ways (note that this description is focused on hardware; however, these procedures apply to software and process as well):

- Return-to-Print. This is a generic term meaning that the problem has been fully resolved, and the hardware has been restored to full compliance with established design, certification, safety, and operating requirements. Returning a hardware item "to print" can involve, but is not limited to, scrapping or downgrading the anomalous hardware and replacing it with fully conforming hardware, or performing maintenance on the hardware to restore compliance. Returning to print can also include a UA that resulted in a nonconforming condition for a temporary period of time but was self-corrected and cannot be repeated or explained. In most cases, especially when no fault has been identified with the design, this is usually the most desirable disposition category from a technical perspective because it restores the approved baseline. However, restoring nonconforming hardware to full compliance is not always feasible due to technical, schedule, or cost limitations
- Change-the-Requirement. Given that much hardware is unique and has not been operated over the long-term in environments specific to NASA programs and projects, it is often discovered that a nonconforming condition is not only acceptable but should also be rolled into a design and specification change. Once such a change is made, the

anomalous condition is no longer considered anomalous and becomes fully compliant with the new requirements

- Waive-the-Requirement. When the risk of accepting continued use of the hardware is acceptable and returning the nonconforming hardware to a fully compliant condition is not feasible, the hardware may continue to be used "as is" in its nonconforming condition by formally waiving the need to comply with the governing design, safety, or operating requirements. This is typically the least desirable solution not only because the underlying nonconforming condition is not resolved, but also because issuing waivers makes tracking the design and configuration baseline and understanding the overall risk posture of the hardware more difficult

Within the PRACA process, the three resolution options described above are further broken down into six distinct categories, which include:

- Rework. This is the implementation of corrective actions that completely eliminate the nonconformance; change to design, specification, or other baseline and configuration-controlled requirements such that the hardware is now in a fully conforming condition with the requirements
- Repair. This is the implementation of corrective actions that restore acceptable functionality but some aspect of the hardware remains noncompliant with requirements; repairs not implemented per an approved standard repair procedure (SRP) or repaired by engineering Material Review Board action require a documented waiver to accept the nonconforming condition
- Scrap. This authorizes the nonconforming hardware to be discarded; typically, scrapped hardware is replaced with fully conforming hardware. Scrapping nonconformance hardware most often requires approval via a change request
- Use-As-Is. This authorizes the continued use of nonconforming hardware in its current physical and functional condition. Use-as-is dispositions require a documented waiver to accept continued use of the hardware
- Downgrade. Downgrading nonconforming hardware is typically done when it is not feasible to restore the hardware to flight-worthy condition but the hardware remains suitable for other ground-related activities (such as training). Downgrading nonconforming flight hardware should be authorized via a change request
- Explained or Unexplained Condition. Disposition where no corrective action is required to restore the hardware to a conforming condition and implementation of recurrence controls is not required. This usually reflects the case where a nonconforming condition was temporary and did not result in overstress of the hardware. This includes UAs that cannot be repeated, did not result in overstress of the hardware, and for which no corrective action or recurrence control is required

APPENDIX C. Significant Problem Report (SPR)

C.1 Overview

An SPR should be as complete as possible so as to foster understanding of the problem, the potential solution, and the nature of the decision being requested from those in a position to allocate resources to resolve the problem. The following pages contain reporting templates.

C.2 SPR Template

Attached to this appendix is a template that outlines the key elements of a SPR.

C.3 SPR Summary List Template

Attached to this appendix is a template that outlines the key elements of a SPR Summary List.

Significant Problem Report

Title						SPID Number	
Description							
Source							
Affected Program		Status		Identification Date		Assignment Date	
Identified by		Organization		Phone		Email	
Assigned to		Organization		Phone		Email	

Sample Only

Problem Category	
Problem Level	
Problem Consequence	
Problem Analysis	
Resolution Plan	
Success Metric	
Management Request	
Management Approval	
Printed: 6/01/06	

	TIMEFRAME		
CONSEQUENCE CLASS	A	B	C
I	1	1	2
II	1	2	3
III	2	3	4
IV	3	4	5
V	3	4	5

High Level	(red)
Medium Level	(yellow)
Low Level	(green)

Approval Date	
Page m of n	

Significant Problem Report Summary List					
SPID Number	Title	Description	Status	Level	Comments
Printed:	6/01/06				Page m of n

Sample Only

APPENDIX D. Data Mining Overview

D.1 Introduction

Data mining will allow NASA to uncover patterns and trends from its existing databases in order to employ problem management. Numerous consultants provide data mining services. Equally numerous software vendors claim to provide tools that are "one-size-fits-all" and can be used on any kind of dataset. In order to truly be successful in performing data mining, or in selecting appropriate tools for a data mining study, the most critical requirement is that the analyst fully understands the business, the objective of the study, and the data available to perform the study. Daniel Larose, one of the leading experts in data mining, states in his book:

> "Data mining is easy to do badly. A little knowledge is especially dangerous when it comes to applying powerful models based on large data sets. For example, analyses carried out on un-preprocessed data can lead to erroneous conclusions, or inappropriate analysis may be applied to data sets that call for a completely different approach, or models may be derived that are built upon wholly specious assumptions. If deployed, these errors in analysis can lead to very expensive failures" (Larose).

Data mining is appropriate for knowledge discovery; however, technical experts are required to interpret whether the results are meaningful. Larose is quoted heavily in this appendix. All wording in quotations are directly from Larose's book, unless otherwise denoted with a citation referencing another source.

D.2 Data Mining Methods

There are a number of well-accepted methods for performing data mining in various industries. The following list highlights a small sample of those available:

- Process model (D.3)
- Classification methods (D.4)
- Clustering techniques (D.5)
- Neural networks (D.6)
- Kohonen maps (D.7)
- Affinity analysis (D.8)

Other areas of interest when mining data include what to do when missing values, misclassifications, or outliers occur. Exploratory Data Analysis (EDA) is an approach to data analysis that uses mainly graphical techniques to explore the data. It allows the data to uncover its underlying structure.

Larose's book is a text which may be helpful in applying data mining techniques. The following sections provide an overview of various topics, several of which can be found in Larose's text.

D.3 Process Model

Several process models for data mining exist. The following diagram shows one model based on the Cross-Industry Standard Process for Data Mining (CRISP-DM). The model "provides an overview of the life cycle of a data mining project. It contains the corresponding phases of a project, their respective tasks, and relationships between these tasks" (http://www.crisp-dm.org/Process/index.htm). The phases, which can be seen in Figure D.3-1, are described as follows:

a. Business Understanding. In this phase, the project requirements and objectives are defined with respect to the business. The data mining problem is then defined from these goals. A plan to achieve the goals should be created

b. Data Understanding. Data is collected and analyzed. In this phase it is necessary to detect quality issues within the data or subsets. It is also the time to form initial hypotheses about the data

c. Data Preparation. All actions to achieve the final dataset occur in this phase. The cases and variables should be selected for analysis. The data should be cleaned in preparation for the modeling tools

d. Modeling. Modeling techniques are selected and applied after having been calibrated to optimize results. Multiple techniques can exist and may be applied to one type of data mining problem. It is often times necessary to move back and forth between the data preparation phase and the modeling phase in order to assure that the data is in the correct form

e. Evaluation. The model should be evaluated with respect to quality. The model should also be evaluated to confirm that it meets the objectives originally set forth. It is also important to determine whether some factor of the business has not been adequately considered. Finally, a decision on the use of the data mining results should be made

f. Deployment. The steps to make use of the model should first be understood by the user and then use of the model may occur. The user may be the data analyst or the customer

A specific sequence from phase to phase does not exist; looping back to a previous phase may be necessary. "The arrows indicate the most important and frequent dependencies between phases. The outer circle in the figure symbolizes the cyclic nature of data mining itself. The lessons learned during the process can trigger new, often more focused business questions. Subsequent data mining processes will benefit from the experiences of previous ones" (http://www.crisp-dm.org/Process/index.htm).

D.3.1 Data Processing

Data processing or data preparation as it is described in Step c. (D.3) of the model occurs after the data has been collected. "To be useful for data mining purposes, databases may need to undergo preprocessing, in the form of *data cleaning* and *data transformation*. Data mining often deals with data that hasn't been looked at for years, so that much of the data contains field values that have expired, are no longer relevant, or are simply missing. The overriding objective is to *minimize* 'garbage in, garbage out' *(GIGO)*: to minimize the 'garbage' that gets into our model so that we can minimize the amount of garbage that our models give out."

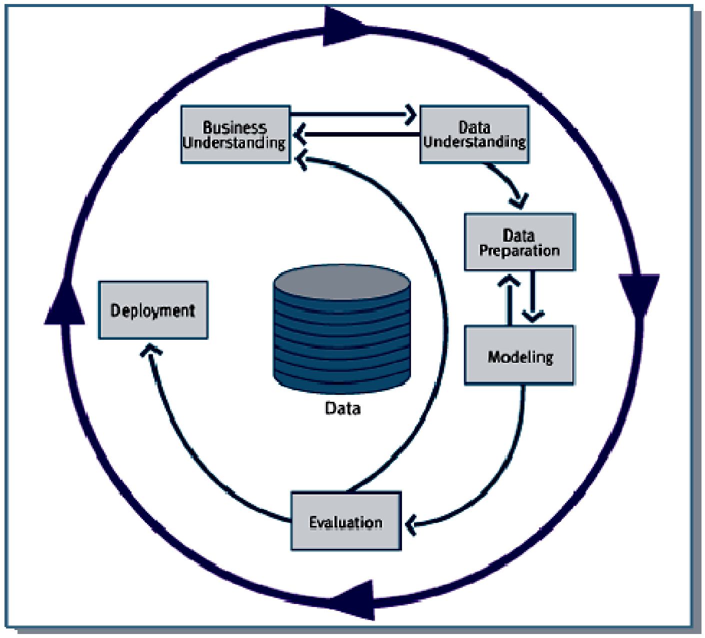


Figure D.3-1: Phases of the CRISP-DM Process Model

D.3.1.1. Data Cleaning

Data cleaning attempts to correct missing values, misclassifications, and outliers.

D.3.1.1.1. Missing Values

Data can be missing for a variety of reasons: not recorded, not measured, lost, or unknown. Most data mining methods do not work well when values are missing from a data set. It is common practice to simply omit a row or column that contains a missing value. However, ignoring the field completely can lead to biased data. A different approach focuses on replacing a missing value with some number, whether it is a constant, or the mean of the dataset. This solution also creates biased data. Their main flaw is that the substituted value is not the correct value. One option is to use logic methods, in which a substitute feature is found that approximately mimics the performance of the missing feature (ref: Bao). Another good approach is to find multiple solutions. If it is possible to isolate the missing values to a few fields, data mining software may be able to find multiple solutions. Various software packages are available for missing value analysis.

D.3.1.1.2. Misclassifications

It is necessary to confirm that classifications are valid and consistent. For example, "one of the functions of Insightful Miner's *missing values* node is to display a frequency distribution of the categorical variables available." Other software can factor in various weights for misclassifications in various categories. This function attempts to deal with situations in which one type of misclassification is more serious than another. Each approach can be a significant help in assuring proper classifications.

D.3.1.1.3. Outliers

"Outliers are extreme values that lie near the limits of the data range or go against the trend of the remaining data." Outliers are typically the result of faulty data collection, recording, or entry. It is important to identify outliers because they can drastically sway statistical analysis. Outliers usually inflate error rates, giving distorted results. "Problems in data mining can arise from outliers, especially with clustering based on distance calculations" (http://statsoft.com/textbook/stdadmin.html#edu).

"One graphical method for identifying outliers is a histogram of the variable. Sometimes two-dimensional scatter plots can help to reveal outliers in more than one variable."

D.3.1.2. Exploratory Data Analysis

EDA is an approach to data analysis that uses mainly graphical techniques to explore the data. It allows the data to uncover its underlying structure. EDA is useful when the analyst does not have *a priori* a hypothesis of the expected relationships among the variables. EDA takes many variables into account in order to examine any systematic patterns or interrelationships. The graphical techniques employed by EDA usually consist of "plotting the raw data, plotting simple statistics, and positioning such plots so as to maximize natural pattern-recognition abilities, such as using multiple plots per page" (Heckert, et al.).

D.3.1.2.1. Correlated Variables

Situations are often encountered in data mining in which there are a large number of variables. "In such situations, it is very likely that subsets of variables are highly correlated with each other. The accuracy and reliability of a classification or prediction model will suffer if highly correlated variables or variables that are unrelated to the outcome of interest are included" (http://www.resample/com/xlminer/help/PCA/pca_intro.htm). Principal components analysis (PCA) is used to decrease the dimensionality of a problem, by transforming interdependent, correlated variables into a smaller set of independent, uncorrelated variables. The dimensionality is reduced without a real loss of the data. It is important "to avoid feeding correlated variables into statistical models. At best, using correlated variables will overemphasize one data component; at worst, using correlated variables will cause the model to become unstable and deliver unreliable results."

Many statistical software programs make use of a matrix plot, which is a matrix of scatter plots for a set of numeric variables, to determine the relationship between variables. A regression tool provides an estimated regression equation, as well as a correlation value. R-squared is the quantitative measure of the correlation or goodness-of-fit of a trend model to actual data. When the R-squared statistic is precisely 1, which shows a perfect linear relationship, one of the two variables correlated may be eliminated. The selection of which variable to eliminate is arbitrary. Following this or a similar procedure will eliminate redundant variables. "A further benefit of doing so is that the dimensionality of the solution space is reduced."

D.3.1.2.2. Categorical Variables

"One of the primary reasons for performing exploratory data analysis is to investigate the variables, look at histograms of the numeric variables, examine the distributions of the categorical variables, and explore the relationships among sets of variables." Generally, categorical data can be noisy and very problematic. Using "one-of-N" or "thermometer" encoding presents additional bias. Because the encodings increase the total number of variables, the effectiveness of many clustering techniques is decreased. Various software packages are available to explore categorical variables.

D.3.1.2.3. Anomalous Fields

"Exploratory data analysis will sometimes uncover strange or anomalous records or fields which the earlier data cleaning phase may have missed. It is possible that the fields just contain bad data. Further communication with someone familiar with the data history should occur before inclusion of these variables in models or further data mining steps."

D.3.1.2.4. Numerical Variables

"Numerical summary measures include minimum and maximum; measures of center, such as mean, median, and mode; and measures of variability, such as standard deviation." Using software packages will aid in exploring the relationships and correlations between numerical predictive variables.

D.3.1.3. Data Transformation

Data transformation attempts to normalize the numerical dataset, so that each variable might be compared on a standardized scale. Various forms of normalization exist. Two examples are as follows:

- Min-max normalization. This approach "works by seeing how much greater the field value is than the minimum value and scaling this difference by the range"
- Z-Score standardization. This approach "works by taking the difference between the field value and the field mean value and scaling this difference by the standard deviation of the field value"

D.4 Classification Methods

The following sections provide examples of methods for classification. These examples are only a small selection of the methods available.

D.4.1 k-Nearest Neighbor Algorithm

"The *k*-nearest neighbor algorithm is most often used for classification, although it can be used for estimation and prediction. The *k*-nearest neighbor algorithm is an example of *instance-based learning*, in which the training data set is stored, so that a classification for a new unclassified record may be found simply by comparing it to the most similar records in the training set."

The *k*-nearest neighbor method offers advantages that other methods do not. It is robust to noisy data and effective even if the dataset is large because the higher values of *k* help to smooth out the data. However, the *k*-nearest neighbor method also has several drawbacks. It is not always easy to determine the number of nearest neighbors, *k*. The computation cost is also very high.

D.4.1.1. Distance Function

The distance function is an important metric when applying the *k*-nearest neighbor method. "Data analysts define distance metrics to measure similarity. A *distance metric* or *distance function* is a real-valued function *d*, such that for any coordinates *x*, *y*, and *z*:

- $d\,(x,y) \geq 0$, and $d\,(x,y) = 0$ if and only if $x = y$
- $d\,(x,y) = d\,(y,x)$
- $d\,(x,z) \leq d\,(y,x) + d\,(y,z)$

The most common distance function is *Euclidean distance*:

$$d_{\text{Euclidean}}\,(x,y) = \sqrt{\textstyle\sum}(x_i - y_i)^2$$

where x = $x_1, x_2, \ldots, x_m$, and y = $y_1, y_2, \ldots, y_m$ represent the *m* attribute values of two records."

"When measuring distance, however, certain attributes that have large values, such as" program cost, "can overwhelm the influence of other attributes which are measured on a smaller scale, such as" years of Space Shuttle service. "To avoid this, the data analyst should make sure to *normalize* the attribute values."

"For continuous variables, the *min-max normalization* or *Z-score standardization* may be used. For categorical variables, the Euclidean distance metric is not appropriate. Instead we may define a function, 'different from,' used to compare the *i*th attribute values of a pair of records, as follows:

$$\text{different}\,(x_i,y_i) = \begin{cases} 0 & \text{if } x_i = y_i \\ 1 & \text{otherwise} \end{cases}$$

where x_i and y_i are categorical values. We may then substitute different (x_i, y_i) for the *i*th term in the Euclidean distance metric above."

D.4.1.2. Combination Function

Once it has been determined "which records are most similar to the new, unclassified record, we need to establish how these similar records will combine to provide a classification decision for the new record. That is, we need a *combination function*. The most basic combination function is simple unweighted voting." One problem is that sometimes simple voting may result in a tie. Weighted voting is one alternative that makes a tie less likely to occur. In weighted voting, the "closer neighbors have a larger voice in the classification decision than do more distant neighbors."

D.4.1.3. k-Nearest Neighbor Algorithm for Estimation and Prediction

"Locally weighted averaging is one way to use the *k*-nearest neighbor algorithm for estimation and prediction. First, the value of *k* would be arbitrarily chosen. Then the estimated target value $\hat{y}$ is calculated as

$$\hat{y}_{new} = \textstyle\sum w_i y_i \,/\, \sum w_i$$

where $w_i = 1 \,/\, d\,(\text{new}, x_i)^2$ for existing records, $x_1, x_2, \ldots, x_k$."

D.4.1.4. Choosing a Value of *k*

"In choosing a small value of k, it is possible that the classification or estimation may be unduly affected by outliers or unusual observations ('noise'). With small k (such as $k = 1$), the algorithm will simply return the target value of the closest nearest observation, a process that may lead the algorithm toward overfitting, tending to memorize the training data set at the expense of the ability to generalize."

"On the other hand, choosing a value of k that is not too small will tend to smooth out any idiosyncratic behavior learned from the training set. However, if we take this too far and choose a value of k that is too large, locally interesting behavior will be overlooked. The data analyst needs to balance these considerations when choosing the value of k."

D.4.2 Decision Trees

A decision tree is an arrangement of tests that prescribes an appropriate test at every step in an analysis. Decision tree learning typically suits problems with the following characteristics:

- The data set provides a fixed set of attributes and their values
- The target function has discrete output values

In a decision tree, squares represent the decision to be made, circles represent chance events, and the branches extending from the squares represent the different choices available to the decision maker. The consequence of each decision is annotated at the end of the branches.

"One of the most attractive aspects of decision trees lies in their interpretability, especially with respect to the construction of *decision rules*. Decision rules can be constructed from a decision tree simply by traversing any given path from the root node to any leaf. The complete set of decision rules generated by a decision tree is equivalent (for classification purposes) to the decision tree itself."

D.4.2.1. Classification and Regression Trees

Regression trees are useful when the target variable is continuous. If the target variable is categorical, then a classification tree is more useful. "Classification and regression trees (CART) recursively partition the records in the training data set into subsets of records with similar values for the target attribute. The CART algorithm grows the tree by conducting for each decision node, an exhaustive search of all available variables and all possible splitting values, selecting the optimal split."

D.4.3 "C4.5" Algorithm

"The C4.5 algorithm is Quinlan's extension of his own ID3 (Iterative Dichotomiser 3) algorithm for generating decision trees. Just as with CART, the C4.5 algorithm recursively visits each decision node, selecting the optimal split, until no further splits are possible. However, there are interesting differences between CART and C4.5:

Unlike CART, the C4.5 algorithm is not restricted to binary splits. Whereas CART always produces a binary tree, C4.5 produces a tree of more variable shape

For categorical attributes, C4.5 by default produces a separate branch for each value of the categorical attribute. This may result in more "bushiness" than desired, since some values may have low frequency or may naturally be associated with other values

The C4.5 method for measuring node homogeneity is quite different from the CART method; the C4.5 method uses the concept of *information gain* or *entropy reduction* to select the optimal split."

D.5 Clustering Techniques

Clustering is the classification of similar objects into different groups, or clusters, so that the data in each set share some common trait. "Clustering differs from classification in that there is no target variable for clustering. The clustering task does not try to classify, estimate, or predict the value of a target variable. Clustering is often performed as a preliminary step in a data mining process, with the resulting clusters being used as further inputs into a different technique downstream, such as neural networks." Just a few of the available clustering techniques are described in the following sections.

D.5.1 Hierarchical Clustering Methods

Hierarchical clustering builds the hierarchy from the individual elements by progressively merging clusters. The result is a treelike structure of clusters. Agglomerative hierarchical clustering builds, beginning at the top of the tree. Divisive hierarchical algorithms break up by beginning at the bottom.

D.5.1.1. *k*-Means Clustering

The *k*-means algorithm is a straightforward approach that assigns each point to the cluster whose center is nearest. To use the algorithm, it is necessary to first determine *k* clusters and then determine the cluster centers. Then assign each point to the nearest cluster and recompute the center cluster. This process is repeated until some convergence criterion is met or there is no change in centers. This algorithm is good for use on large datasets due to its simplicity and speed.

D.6 Neural Networks

Artificial neural networks are non-linear, statistical math models that describe a function. They are modeled after the biological process of learning in the cognitive system and the neurological functions of the brain. Neural networks can be used to model complex relationships between inputs and outputs or to find patterns in data. These networks are capable of predicting new observations from existing observations of data through "learning."

The networks contain nodes, which are connected using assigned weights. As one node sends a signal to the next node, the output value is weighted according to the outgoing connection. The receiving node calculates the weighted sum of the values coming in from each connecting node. The networks learn by varying the weights of the connections.

"One of the advantages of using neural networks is that they are quite robust with respect to noisy data. Because the network contains many nodes (artificial neurons), with weights assigned to each connection, the network can 'learn' to work around these uninformative (or even erroneous) examples in the data set. However, unlike decision trees, which produce intuitive rules that are understandable to non specialists, neural networks are relatively opaque to human interpretation. Also, neural networks usually require longer 'training times' than decision trees, often extending into several hours."

D.6.1 Input and Output Encoding

When using neural networks, attribute values, including categorical variables, should take values between 0 and 1. To standardize the attribute values, the min-max normalization may be applied. "This works well as long as the minimum and maximum values are known and all potential new data are bounded between them. Neural networks are somewhat robust to minor violations of these boundaries. If more serious violations are expected, certain ad hoc solutions may be adopted, such as rejecting values that are outside the boundaries, or assigning such values to either the minimum or the maximum value."

"Categorical variables are more problematical, as might be expected. If the number of possible categories is not too large, one may use *indicator (flag) variables*. In general, categorical variables with k classes may be translated into k - 1 indicator variables, as long as the definition of the indicators is clearly defined. The neural network output nodes always return a continuous value between 0 and 1 as output."

D.6.2 Sigmoid Activation Function

Activation functions are used to introduce nonlinearity into the network. Nonlinearity makes multilayer networks very powerful. For back propagation learning, the activation function should be differentiable. "The sigmoid function is useful because it combines nearly linear behavior, curvilinear behavior, and nearly constant behavior, depending on the value of the input. The sigmoid function is as follows:

$$y = f(x) = 1 / (1 + e^{-x})$$

Moderate increments in the value of x produce varying increments in the value of $f(x)$, depending on the location of x. Near the center, moderate increments in the value of x produce moderate increments in the value of $f(x)$; however, near the extremes, moderate increments in the value of x produce tiny increments in the value of $f(x)$. The sigmoid function is sometimes called a *squashing function*, since it takes any real-valued input and returns an output bounded between zero and 1."

D.7 Kohonen Maps

D.7.1 Self-Organizing Maps (SOM)

SOM, sometimes referred to as Kohonen maps, were first described by Finnish researcher Tuevo Kohonen in 1982. SOM is a widely used method of clustering analysis, which is especially good at converting high-dimensional data into a low-dimensional, discrete map.

SOMs are based on a competitive network, in which each node with its weight vector closest to the vector of inputs is assigned a new weight, making it even closer to the input vector. The weights for the competing nodes, or neighbors, are also changed. The neighborhood size is called *R*. The further a node is from the winning node, the smaller the weight change. "The process is repeated for each input vector. The network winds up associating output nodes with groups or patterns in the input data set. If these patterns can be named, the names can be attached to the associated nodes in the trained net" (http://en.wikipedia.org/wiki/self-organizing _map).

The SOM has two modes of operation:

a. Training process. A map is built, and the neural network organizes itself using a competitive process. The network is given a large number of input vectors, otherwise, all input vectors should be administered several times

b. Mapping process. A new input vector given a location on the map is automatically classified or categorized. The neuron whose weight vector lies closest to the input vector is the winning neuron. (This can be simply determined by calculating the Euclidean distance between input vector and weight vector (http://en.wikipedia.org/wiki/self-organizing _map))

KOHONEN NETWORKS ALGORITHM

For each input vector x, do:

- *Competition.* For each output node j, calculate the value $D(w_j,x_n)$ of the scoring function. For example, for Euclidean distance, $D(w_j,x_n) = \sqrt{\sum (w_{ij} - x_{ni})^2}$. Find the winning node J that minimizes $D(w_j,x_n)$ over all output nodes.
- *Cooperation.* Identify all output nodes *j* within the neighborhood of *J* defined by the neighborhood size *R*. For these nodes, do the following for all input record fields:
- *Adaptation.* Adjust the weights:

$$w_{ij,new} = w_{ij,current} + \ (x_{ni} - w_{ij,current})$$

- Adjust the learning rate and neighborhood size, as needed.
- Stop when the termination criteria are met.

Figure D.7-1: Kohonen Networks Algorithm

D.7.2 Kohonen Networks Algorithm

"Kohonen networks are self-organizing maps that exhibit *Kohonen learning*. Suppose that we consider the set of *m* field values for the *n*th record to be an input vector $x_n = x_{n1}, x_{n2}, \ldots, x_{nm}$, and the current set of *m* weights for a particular output node *j* to be a weight vector $w_j = w_{1j}, w_{2j}, \ldots, w_{mj}$. In Kohonen learning, the nodes in the neighborhood of the winning node adjust their weights using a linear combination of the input vector and the current weight vector:

$$w_{ij,new} = w_{ij,current} + \ (x_{ni} - w_{ij,current})$$

where η , $0 < \eta < 1$, represents the *learning rate*, analogous to the neural networks case. Kohonen indicates the learning rate should be a decreasing function of training epochs (runs through the data set) and that linearly or geometrically decreasing η is satisfactory for most purposes."

"The algorithm for Kohonen networks is shown in the accompanying box. At initialization, the weights are randomly assigned, unless firm a priori knowledge exists regarding the proper value for the weight vectors. Also at initialization, the learning rate η and neighborhood size, *R,* are assigned. The value of *R* may start out moderately large but should decrease as the algorithm progresses. Note that nodes that do not attract a sufficient number of hits may be pruned, thereby improving algorithm efficiency."

D.8 Affinity Analysis

"Affinity analysis is the study of attributes or characteristics that 'go together.' Methods for affinity analysis, also known as market basket analysis, seek to uncover associations among these attributes; that is, they seek to uncover rules quantifying the relationship between two or more attributes. Association rules take the form 'If antecedent, then consequent,' along with a measure of the support and confidence associated with the rule."

"The number of possible association rules grows exponentially in the number of attributes. For example, if there are *k* number of attributes, we limit ourselves just to binary attributes, and we account only for the positive cases, there are on the order of $k \cdot 2^{k-1}$ possible association rules."

"The *a priori algorithm* for mining association rules, however, takes advantage of structure within the rules themselves to reduce the search problem to a more manageable size." There are two principal methods of representing this type of market basket data: using either the transactional data format or the tabular data format. The *transactional data format* requires only two fields, an *ID* field and a *content* field, with each record representing a single item only. In the *tabular data format*, each record represents a separate transaction, with as many 0 / 1 flag fields as there are items."

APPENDIX E. Trending Analysis Overview

E.1 Trend Analysis Procedures

This section, "Trend Analysis Procedures" is a companion to Section 4.2.2, "Trend Analysis." This section provides uniform guidance for conducting trend analyses for NASA aeronautics and space programs. It is for the use of NASA Headquarters and NASA field installations involved in the development and operation of these programs. Development of essential information so that NASA management can base critical risk-informed decisions affecting safety and mission success is necessary for the continued credibility and success of this Nation's aeronautics and space programs. It is the responsibility of the program or project manager to ensure those tasked with trend analysis activities have the appropriate knowledge, skills, and abilities to create valuable trend analysis output.

E.2 Purpose

The purpose of trend analysis is to analyze past performance to provide information that can be used to assess the current status and predict future performance. The purpose of this section is to establish a uniform, Agency-wide mechanism for providing NASA management with trend analysis data on which to base top-level decisions affecting the safety and success of developmental or operational space and aeronautical programs and projects and related payloads, and institutional support facilities.

This section provides guidance to help assure the proper use of trend analysis to support Agency operational functions. This document is not intended to restrict innovation or creativity as it applies to program and project activities; however, it is intended to capture and standardize application of trend analysis and consistency in data capture. For consistency across the Agency, the standardized process, procedures, and output formats should be followed.

This document provides guidance to help standardize trending across NASA. There are many trend analysis activities on-going at the local level. For meaningful trend analysis to be valuable, sharing of the information should exist to allow for the development of correlations.

E.3 Guidelines

E.3.1 Overview

The major goal of each NASA program management level is to achieve operational and research objectives while ensuring that all NASA and NASA-sponsored flight, orbital, and ground operations are conducted safely and with a full understanding of mission risks. Achievement of this goal is supported through rigorous engineering analyses and assessments. The NASA system of trend analysis addresses the institutional characteristics and performance of each program as well as progress toward improving the program and eliminating problems.

Trend analysis is an element of engineering investigation that provides continuing review of program factors. Trend analysis has two prime characteristics: investigation of actual events and comparative assessment of multiple events. Trend analysis is applied to program characteristics that vary in relation to time, sequence, or element performance. Trend analysis results are used to evaluate the operations of a program and its component systems by assessing past performance to establish baselines for current and future performance. When a valid trend exists, the accuracy of the analysis will increase as more time or event data are collected.

Trend analysis is a method for unearthing correlations between unique elements, parts, and procedures that are not obvious or may even be counterintuitive on the surface. Trend analysis is also used to discover and confirm correlations between diverse factors. The degree to which trend analysis approaches can be applied to the elements of NASA's programs and projects will vary and be based on many factors. However, it is important to avoid the situation where the data was available but not analyzed and, therefore, not presented for decision in a comprehensive manner.

E.3.2 Basic Guidelines for Trend Analysis

Trend analysis is a formal data analysis approach. It is not sufficient to simply plot quantitative data and superimpose a trend line. Trend analyses should measure correlation and goodness-of-fit, use normalization techniques, and qualitatively analyze results (that is, present the results, management, and technical reasons for the trends). Significant trend analyses should include an assessment from the cognizant engineer, technician, or analyst. When appropriate, trend predictions should be included.

Trend analysis requirements should be included in all program-planning phases to ensure the capability to provide timely analyses of testing or operational events. Planning for trend analysis should include selective data collection, development of data analysis systems, and the means for disseminating results. In addition, trend analysis should incorporate data cleansing techniques. Trend analysis using 'dirty data' or improper statistical techniques will only provide misleading information.

E.3.3 Elements of Trend Analyses

A trend analysis program consists of four discrete but interrelated elements. These four elements are designed to accomplish three objectives. These three objectives are as follows: to monitor the performance of hardware and software designs (to forecast anomalies or potential problems), to monitor identified problems in order to assess progress in problem resolution, and to monitor information related to problems related to the institutional factors connected with major NASA programs (when those institutional factors could impact safety or mission success).

The four discrete elements are referred to, respectively, as performance, problem, programmatic, and supportability trend analysis. Analyses of these elements can be found throughout the engineering community; however, organizing trend analysis into these specific groupings is a NASA-unique approach. A decision by a program or project manager, after consulting with the supporting staff, to employ any of the various trend analysis elements and corresponding techniques for trend analysis, should be made after a careful and deliberate consideration of the benefit and expected efficacy of the result compared with the effort made to achieve the result. The four elements of trend analyses that will be discussed in the following sections are as follows:

a. Performance Trend Analysis. This is analysis of data based upon the measurement of specific key performance parameters, which indicate the safe and effective operation of a critical process or item of hardware or software (Section E.4)

b. Problem Trend Analysis. This is analysis of the data based upon the number of problems occurring in the area under study. Its purpose is to identify the source of key problems and to track whether action taken to resolve the problems is effective (Section E.5)

c. Supportability Trend Analysis. This is analysis of the effectiveness of logistics elements in supporting NASA programs and projects. Supportability trend analysis is concerned with the recurrence of logistics problems and the effective control of these problems (Section E.6)

d. Programmatic Trend Analysis. This is analysis of institutional information relating to program schedule and to supporting personnel activities. These analyses are used to assess the impact of schedule pressures and major disruption in resource capability or the ability of the work force to respond in a predictably safe and reliable manner (Section E.7)

E.4 Performance Trend Analysis

E.4.1 Overview

The primary objective of performance trend analysis is to monitor hardware and software operations to forecast anomalies or potential problems of a specific system, subsystem, or component. This section describes performance trend analysis and reporting. A consistent approach is established for conducting performance trend analysis and reporting the results to NASA management.

Performance trend analyses provide a parametric assessment of hardware and software operations to forecast anomalies or potential problems. Trends are used to identify impending failure or performance degradation in hardware and software, particularly those that affect safety or mission success. Key characteristics or performance parameters (such as temperature, pressure, or erosion) are identified and evaluated to determine if they are good predictors of failure. In some cases, the characteristics are so critical to safety or mission success that real-time performance trend analyses should be conducted.

Performance trend analysis identifies measurable parameters that can indicate component or system degradation before failure. Sampling a parameter's values over time (either historical parameter values for the same hardware component or values recorded at discrete time intervals during a mission) can reveal significant trends in performance degradation before exceeding a redline limit or experiencing a failure.

Performance trend analysis can be used to detect certain types of progressive failure mechanisms before final failure in a system, subsystem, or component. These failure mechanisms include (but are not limited to): wear, erosion, under or over-temperature, under or overpressure, vibration, friction, leakage, material property change, calibration drive, contamination, or electrical resistance change.

E.4.2 Applications of Performance Trend Analysis

Applications of performance trend analysis include:

a. Performing prelaunch maintenance on systems, subsystems, and components based on early detection of degrading parameters to prevent either mission failure or launch commit criteria violation during launch countdown, resulting in launch delay or scrub

b. Maintaining a unit in service based on trend analysis surveillance of the degradation trend line, degradation characteristics, and redundancy. (Note that this application can be used even if a measurable unit parameter exceeds the turnaround functional test limit or normal removal time limits)

c. Providing data to support an objective mathematical risk analysis to yield a probability estimate for predicating remaining life, failure, and exceeding limits

E.4.3 Candidates for Performance Trend Analysis

Candidates for performance trend analysis should be based on the following primary selection criteria:

a. Criticality (requirement based on FMEA and critical items list (CIL) data) (E.4.3.1)

b. Availability and trendability of data (E.4.3.2)

c. Problem history and engineering judgment (E.4.3.4)

E.4.3.1. Criticality (Based on FMEA and CIL Data)

Priorities for performance trend analysis should be established based on concern (risk, safety, and mission success with cost, availability, or schedule as trade parameters only) and expected benefits. Where risk is a primary concern, criticality 1 items should be given highest priority followed by criticality 1R and 1S items.

E.4.3.2. Availability and Trendability of Sensor Data

A determination should be made on whether sensors are available from which to obtain performance data (that is instruments in place to sense measurable performance changes). When no sensors exist, the cost and benefits of developing and installing sensors should be considered. Common performance parameters that are well suited to performance trend analysis include: pressure, temperature, voltage, current, flow rate, torque and motion, given input and required output, and operating elapsed time and cycle (including on or off or open or closed cycle).

Sensor data should be analyzed to determine: a) the relationship to the condition being monitored, and b) whether these data show a trend in performance. Selected parameters should be capable of showing performance degradation (with a definable upper and lower limit) to allow scheduled corrective action before failure. Data sampling rates, transmission rates, and system, subsystem, or component degradation characteristics should be analyzed and compared to determine if the data can be trended to effectively show performance degradation.

E.4.3.3. Problem History and Engineering Judgment

Selection of candidates for trend analysis includes a search of problem reporting databases to identify systems, subsystems, or components with a high frequency of reported problems. Problem reporting records are to be reviewed for history of maintenance problems and component problems and anomalies. This review should focus on, but not be limited to, the following areas: launch delays, component removals, in-flight or on-orbit anomalies or failures, and ground checkout anomalies or failures.

E.4.4 Data Sources of Performance Trend Analysis

The data sources for performance trend analysis include, but are not limited to, the following: flight or orbital data, prelaunch countdown data, ground test, checkout and turnaround data, teardown inspection and analysis reports, acceptance test procedure, failure analyses, and problem reports (requirement including nonconformance, in-flight anomaly, and unsatisfactory condition reports (UCRs)).

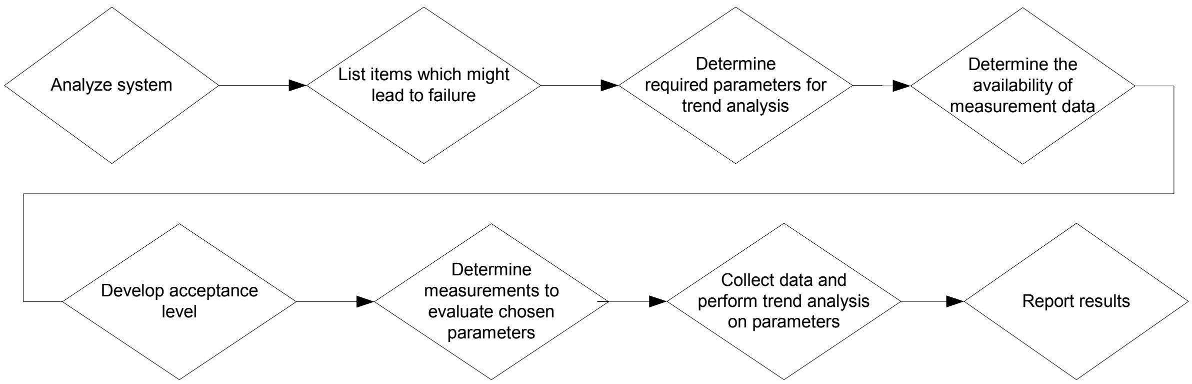


Figure E.4-1: Performance Trend Analysis Process Flow

E.4.5 Considerations for Performance Trend Analysis

The following subsections discuss factors that should be considered when conducting performance trend analysis.

E.4.5.1. Indirect Parameter Indicator

There may be cases where a direct indicator of component performance does not exist; however, performance can be tracked through indirect indicators (for instance, pressure may be an indirect indicator of temperature). In these cases, a mathematical relationship between the parameters, including advisory limits, should be developed for trend analysis.

E.4.5.2. Complementary Performance Data

Many systems contain complementary or interrelated parameters. As a system (or subsystem) changes state, two or more parameters may change in a proportional or inverse proportional relationship. These complementary parameters can be used to verify the trend of a tracked parameter, thus providing redundancy and increasing confidence in the trend data.

E.4.5.3. Trend Limits Adjustment (Based on Operating History)

Operating historical performance data gathered for performance trend analysis can be used to evaluate operating limits when it demonstrates that actual performance variability is less than was anticipated when limits were set originally.

E.4.5.4. Normalizing and Correction Factors

The operating state, output, or load (about or through which a system, subsystem, or component fluctuates) often cannot be controlled to achieve consistent trend data. Factors such as ambient or on-orbit conditions may affect data variability from one checkout or orbit to the next. For these cases, it may be possible to determine a normalized state, output, or load. If the relationship of the actual or normalized operating states is known, the performance trend parameter can be corrected upward or downward to reflect a normalized state. Using data from the normalized state will result in consistent trend data from checkout-to-checkout or orbit-to-orbit.

E.4.5.5. Performance Measurements

Whenever performance data are recorded, an attempt should be made to verify the stability and slope of data approaching or departing the recorded data point. Use of a data buffer is recommended to evaluate pre-event data in verifying the slope of data approaching or departing the recorded data point. Additionally, data filtering and persistence counters should be used to verify that the data point is not a noise spike. Whenever a performance advisory limit is exceeded, complementary data should be recorded to verify sensor condition. Other considerations include but are not limited to: data sampling rates, data sampling resolution, compression effect on resolution, data or system stability and calibration.

E.4.6 Procedures for Performance Trend Analysis

The basic steps (see flow process in Figure E.4-1) in performance trend analysis are as follows:

a. Analyze hardware or software systems to identify items that could lead to a critical or costly failure

b. Prepare a list of these items as candidates for performance trend analysis. Select the items to be analyzed from the list of possible candidates

c. Determine the parameters to be used in judging whether an item's performance is degrading at a rate sufficient to warrant management attention. When these parameters are critical to safety or mission success, strong consideration should be given to performance trend analysis

d. Determine if measurement data are available for the selected performance parameters. A performance parameter may be a directly measurable factor, or a relationship between two or more parameters (i.e., pressure versus time, temperature versus pressure, etc.) based on an algorithm. If measurement data are not available, determine the feasibility of establishing a system to measure the parameters. If feasible, then implement the measurement(s)

e. Establish the performance baseline (acceptance levels or bounds)

f. Determine the measurements necessary to evaluate the chosen parameters. The principal elements for performance trend analysis include: sensor data, time, age, and cycle data collected from design and project operating elements, together with problem reports in associated databases

g. Collect, measure, and record the data and conduct a performance trend analysis to predict an impending failure or ascertain the aging or degradation of an item. If the parameter being trended exceeds the historical limits or is below the performance baseline, the item could experience a failure. At this point, the decision should be made to either retain or replace the item

h. Report the results using charts, graphs, and recommendations

E.4.6.1. Reporting Performance Trend Analysis Results

E.4.6.1.1. Format

A trend analysis chart should display the parameters and health indicators, with appropriate analysis parameters plotted and annotated. When performance degradation of a system, subsystem, or component has been identified, the pertinent charts (or reports) should include, but not be limited to: item, part number, serial number, criticality, failure mode, failure effect, assessment, and action required.

E.4.6.1.2. Frequency

The data analyses, trend charts, and report should be made available to program or project management via regular and special reports. Program or project management should establish routine reporting requirements. Once established, the trend report should be updated at regular intervals. Performance trends should be reported periodically, normally by month or mission event. However, trend reports may be required more frequently, such as when trend data indicate rapid change. NASA management should be alerted in a timely manner of any performance trend analysis results that may affect safety.

E.5 Problem Trend Analysis

E.5.1 Introduction to Problem Trend Analysis

Problem Trend Analysis is intended to identify recurring problems and assess progress in problem resolution or recurrence control. This type of analysis will focus on where the key problems are occurring and their frequency. When the frequency of problems is increasing without an immediately obvious cause, the responsible party should consider initiating a SPR and commence an investigation to determine the cause of the adverse trend. Problem analyses (such as Pareto analyses) can be helpful as a starting point for focusing attention and determining where other analyses, such as performance trend analysis, can be of significant benefit.

E.5.2 Objective of Problem Trend Analysis

The objective of the problem analysis approach is to provide a historical overview of problems in an easy to understand graphical format. The overview should assist in decision-making relative to design effectiveness, process, or procedural changes over time (and the initiation of corrective action to improve trends).

E.5.3 Candidates for Problem Trend Analysis

Candidate items should be comprehensively screened for selection because it is not feasible, meaningful, or cost-effective to perform problem trend analysis on all NASA items and failure modes. Basic criteria for item selection include: problem frequency, criticality, engineering judgment, and unique program or project requirements. The candidate selection process is shown in Figure E.5-1. Descriptions of the process flow elements are as follows:

- Search the center problem report database and other databases to identify failures
- Determine whether to monitor the item for possible future trend or to delete or discard the item completely
- Monitor and observe the item until there is justification to repeat the screening process
- Delete or remove item from consideration for trend analysis
- Review failures obtained from problem report database search to determine whether launch delays were encountered regardless of criticality
- Review failures and decide whether to trend, discard, or monitor based on the technical aspects of failure history of an item
- Review failure frequency over time to determine whether trend analysis is feasible.
- Determine if sufficient failures are available to depict effects based on Engineering Change Proposals (ECPs)
- Consider date of last occurrence to decide whether to trend
- Process customer request for trend analysis without the restrictions applied to other trend analysis sources
- Implement actual trend analysis of selected item
- Decide whether to monitor or discontinue item from further consideration for trend analysis with customer input
- Apply the Five Step Trend Analysis Approach, which is described and illustrated in E.5.6.1

Review documentation for trending candidates – documentation examples include: indentured parts lists, FMEA/CIL, Launch Commit Criteria, Hazard Analysis, NASA Center, prime contractor or subcontractor problem reports, program or project meetings.

E.5.4 Data Sources for Problem Trend Analysis

The primary sources for problem trend data are the failure or problem reporting and corrective action systems, such as PRACA, supported by other databases as required. Unless the trend analysis is uniquely directed toward the contractor's internal operation, it is preferred to use the problem reports written during and after component level acceptance testing.

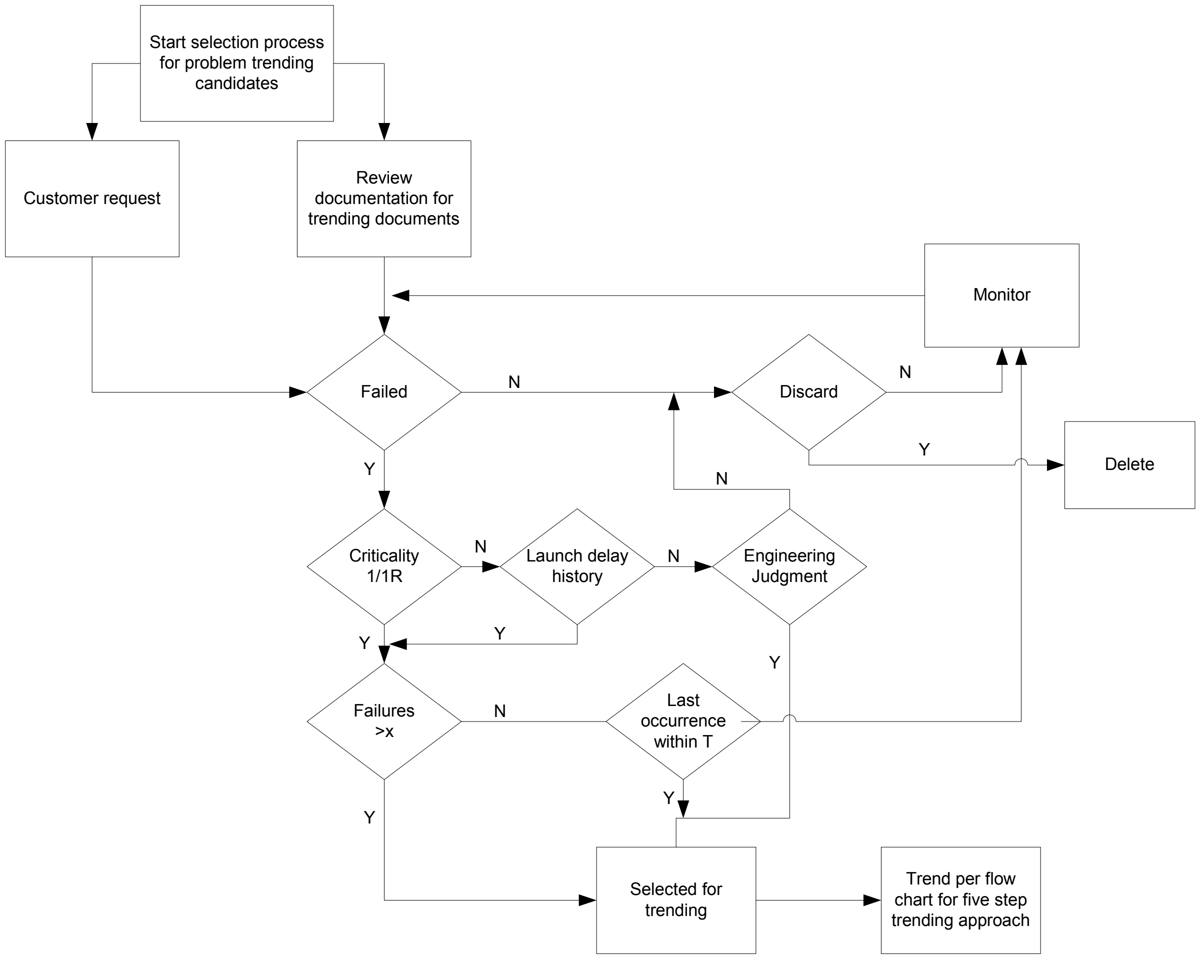


Figure E.5-1: Problem Trend Analysis Selection Process Flow

E.5.5 Considerations for Problem Trend Analysis

Fundamental areas of consideration that should be included in problem trend analysis are as follows: level of analysis (system, subsystem, or component), engineering judgment, statistical analysis, conflict between engineering judgment and statistical analysis, data normalization, goodness-of-fit, and trend direction. These fundamental areas are discussed in the following subsections.

E.5.5.1. Level of Problem Trend Analysis

Trend analysis should consider specific failure modes (with knowledge of the failure mechanism or causes) to effectively evaluate a trend and make specific recommendations for corrective action. To evaluate the effectiveness of corrective actions such as design or process changes, problem trend analysis should be performed at the lowest system, subsystem, or component level for which problem data are available for the failure mode involved. There are two methods for evaluating a trend: engineering judgment and statistical analysis.

E.5.5.2. Engineering Judgment

Engineering judgment is the basis for identifying a trend and classifying it as adverse or favorable. It applies when:

a. Sample size is not sufficient for statistical trend analysis

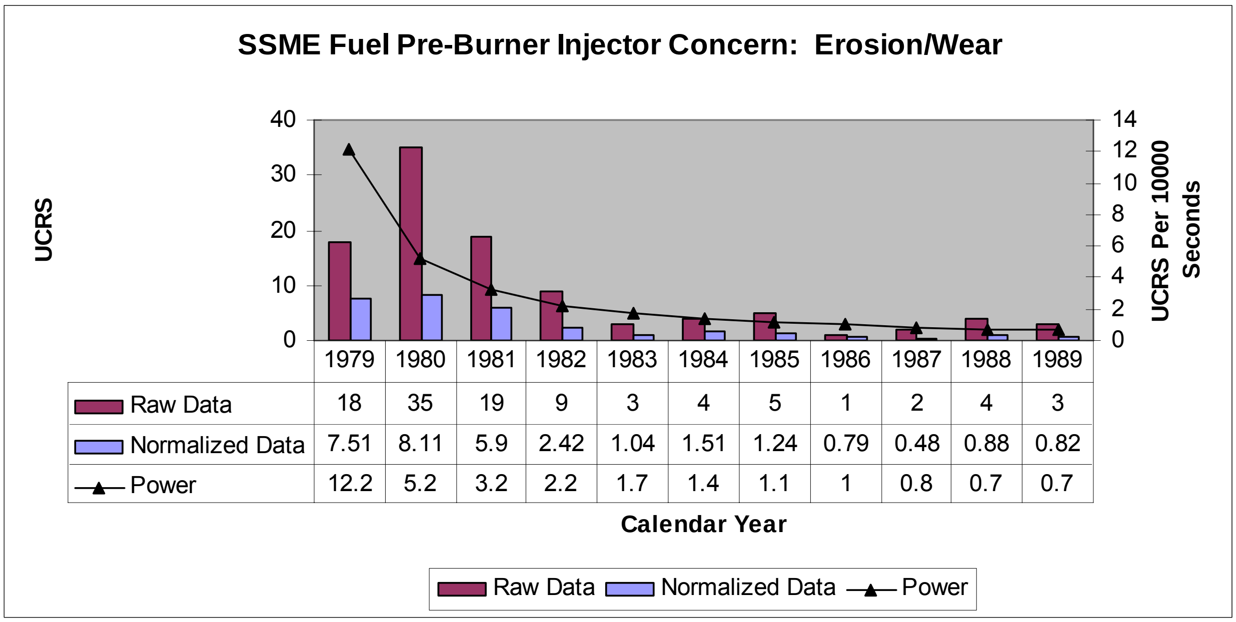
b. Failure mode and root cause are well understood

c. Corrective action is well understood

d. Statistically downward-trend levels are above zero, with one or more problem reports per year in most of the recent years trended (see example in Figure E.5-2)

e. Sufficient failure-free tests or inspections have been conducted to verify effectiveness of corrective action

Where practical, the results of engineering judgment should be verified by statistical analysis.



	1979	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989
Raw Data	18	35	19	9	3	4	5	1	2	4	3
Normalized Data	7.51	8.11	5.9	2.42	1.04	1.51	1.24	0.79	0.48	0.88	0.82
Power	12.2	5.2	3.2	2.2	1.7	1.4	1.1	1	0.8	0.7	0.7

Figure E.5-2: SSME Fuel Pre-Burner Injector Concern: Erosion and Wear

E.5.5.3. Statistical Analysis

Statistical analysis of a trend should be based on a sample of at least 20 problems; however, a minimum of 5 problems (with at least 5 years of data or 5 sets of mission hardware) could suffice. If corrective action is required based on a trend analysis, the failure mode(s) that constitute(s) the greatest area(s) of concern should be identified for trend analysis.

E.5.5.4. Conflict between Engineering Judgment and Statistical Analysis

Normally, engineering judgment and statistical analysis methods should yield the same trend conclusion (adverse or favorable). However, if there is a conflict in trend direction, engineering judgment usually is preferred for small sample sizes and statistical analysis for large sample sizes. There is no substitute for engineering judgment in assessing the importance of a trend. As an example, for extremely serious conditions, a favorable trend may only indicate that a situation is slowly improving where, in actuality, a more rapid trend of improvement is required.

E.5.5.5. Data Normalization

Before problem trend analysis, the quantity of problem reports per time interval (week, month, and year) or per set of opportunity should be normalized. Examples of normalized data are: problems per 10,000 seconds of run time, problems per 100 tests or inspections, problems per mission or flow, number of firings per year or number of end items delivered per month. Data should be normalized at the lowest possible assembly level.

E.5.5.6. Goodness of Fit

Goodness of fit of the trend to the data points is determined using the R-squared value. The R-squared statistic gives a measure of how well the trend model fits the data. With regard to quantification of fit, R-squared ranges from 0 to 1, with 1 indicating a perfect fit. It is imperative to consider more than linear trend modeling. The highest R-squared value should be selected from one of the following trend models: Linear, Exponential, Power (geometric), Logarithmic, or Positive Parabolic.

E.5.5.7. Trend Direction

Trend direction should be determined using the sign of the R-squared value. If R-squared is less than the value in Figure E.5-3, the trend may be declared level. If R-squared is more than the value, it would be declared upward or downward, depending on the R-squared value sign (positive or negative, respectively). "The R-squared statistic, R-squared = 1-SSE / SST, where SSE is the sum of square errors and SST is the total sum of squares. If the model fits the series badly, the model error sum of squares, SSE, may be larger than SST and the R-squared statistic will be negative" (http://en.wikipedia.org/wiki/R-squared).

> "Note that it is possible to get a negative R-squared for equations that do not contain a constant term. Because R-squared is defined as the proportion of variance explained by the fit, if the fit is actually worse than just fitting a horizontal line then R-squared is negative. In this case, R-squared cannot be interpreted as the square of a correlation. Such situations indicate that a constant term should be added to the model" (http://www.mathworks.com/access/helpdesk/help/toolbox/curvefit/ch_fitt9.html).

Figure E.5-3 is based on the number of sample points and confidence interval (100 *(1 - α) %, where α is the significance level). It is important to note that the table is valid only for two-parameter models, that is, models with two unknown variables a and b, such as in $y = a + bx$ or $y = ae^{bx}$. To use Figure E.5-3, locate the number of observational values (N) and the desired level of significance (almost always use α=.05). The corresponding table entry gives the minimum R-squared value to conclude that the fit is significant. For example, where N=8, then R-squared should be at least .5 (here α=.05) to conclude that the trend model is an adequate description of the data.

Generally, a line is good for fitting upward trends; however, downward trends often are better fitted using one of the nonlinear models. If the R-squared value is not statistically significant, it should be inferred that the trend is level or adverse. However, engineering judgment should still be applied.

Minimum R-squared Values for Significant Trending Fit			
NUMBER OF DATA POINTS (N)	α=.01	α=.025	α=.05
4	.98	.95	.90
5	.92	.85	.77
6	.84	.75	.66
7	.76	.67	.57
8	.70	.59	.50
9	.64	.54	.44
10	.59	.49	.40
11	.54	.44	.36
12	.50	.41	.33
13	.47	.38	.31
14	.44	.35	.28
15	.41	.33	.26
20	.31	.25	.20
25	.26	.20	.16
30	.21	.17	.13

Figure E.5-3: Minimum R-Squared Values for Significant Trending Fit

The determination of the adverse or favorable nature of a trend depends on the system that is being trended. A system that is expected to sustain a certain level of random failures would have an adverse trend if the failure rate increases or is predicted to exceed the design failure rate. A critical system that is maintained and operated to avoid all failures would have an adverse trend if a failure mode reoccurs subsequent to the institution of failure recurrence control after the first failure. Only a result of "no problems reported in the failure mode" would be favorable; any upward or level trend would be considered adverse.

E.5.6 Procedures for Problem Trend Analysis

E.5.6.1. Hierarchical Approach and the Five-Step Method

1. Figure E.5-4 shows an example of typical steps used to identify a component failure mode for trend analysis. Based on the highest frequency of problem reports at each hierarchical level, one might select the element (if applicable) followed by the system, subsystem, component, and finally, the failure mode.
2. There are many valid methods of performing problem analysis; the five-step method is the recommended approach for achieving consistency throughout NASA (Figure E.5-5). This should not preclude the use of other methods that may be more applicable in particular circumstances.

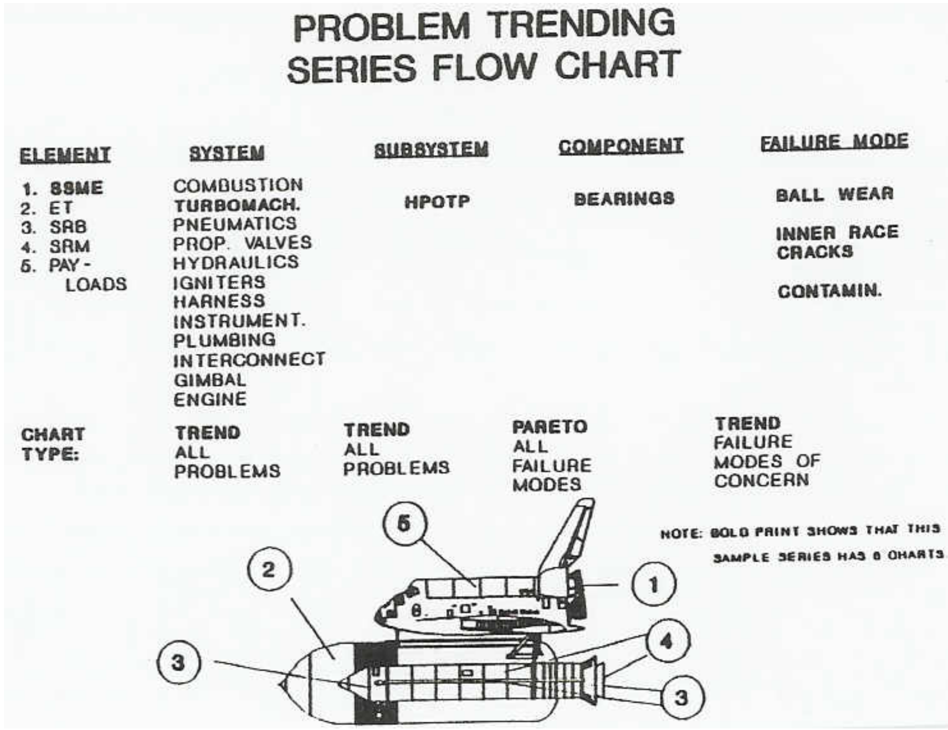


Figure E.5-4: Problem Trending Series Flow Chart

3. The five-step method of problem trend analysis comprises the following activities:
    a. Research appropriate database(s) and extract data (E.5.6.1.1)
    b. Construct a normalized subsystem-level or component-level trend chart (E.5.6.1.2)
    c. Construct a Pareto chart of failure modes or causes and identify area(s) of concern (E.5.6.1.3)
    d. Construct a normalized trend chart for each area of concern and failure mode (E.5.6.1.4)
    e. Prepare a summary assessment of the problem trend, including (E.5.6.1.6):
        - Suspected failure mode(s)
        - Root cause(s)
        - Recommended or actual corrective action(s)

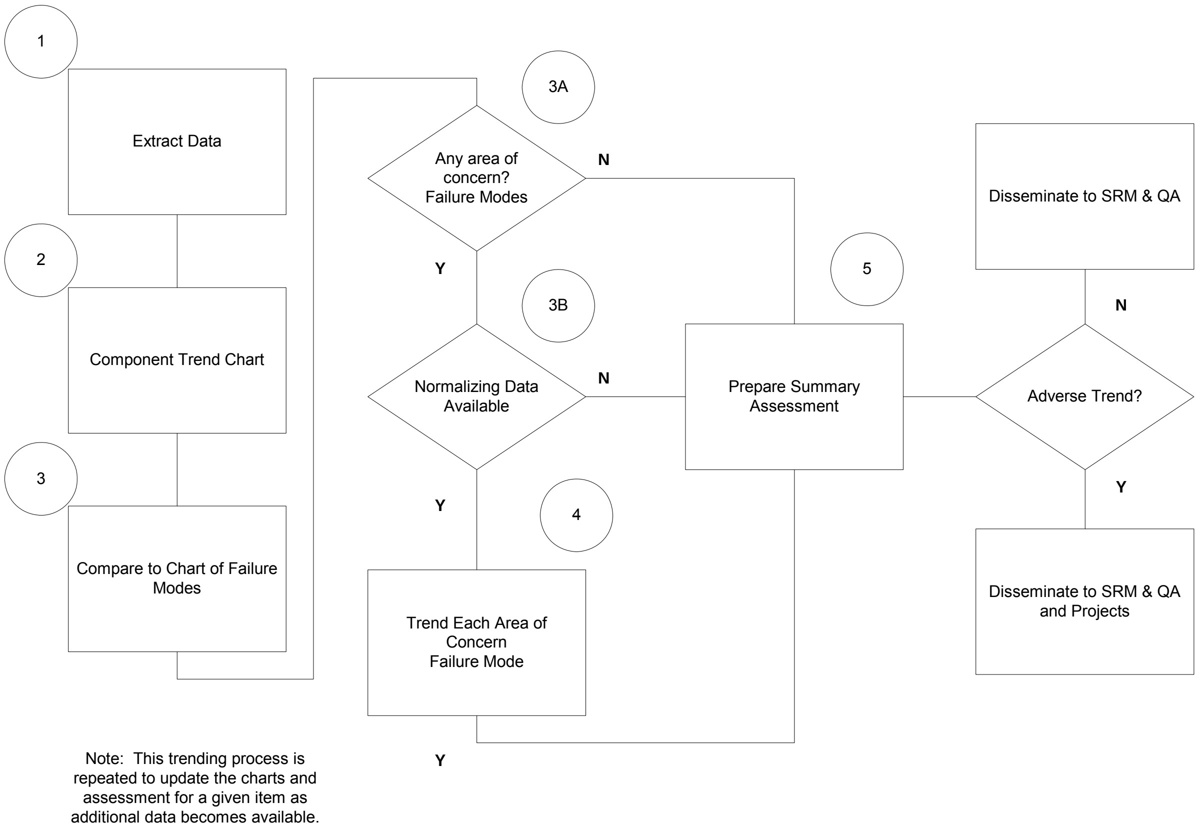


Figure E.5-5: 5-Step Trending Method Flow Chart

E.5.6.1.1. Step 1: Research Database and Extract Data

Automated data search and manual activities are necessary to obtain data for problem trend analysis. Primary considerations in Step 1 are as follows:

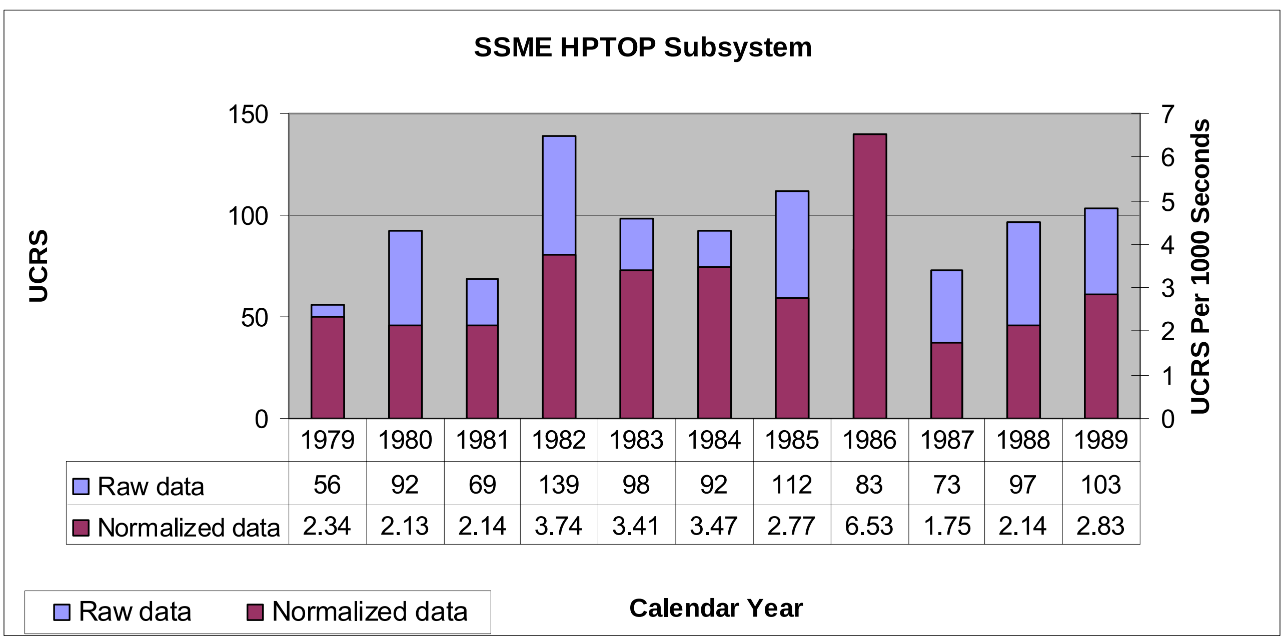
1. Ground Rules for Data Inclusion or Exclusion. In researching the data for trend analysis of a given component, the primary data source is usually the problem report database for the cognizant design center. A second source of data may be the launch center problem report database for flight component problem reports. In-flight or on-orbit anomalies are available from the cognizant design, launch, and operations centers. Ground rules used in excluding data should be noted, for example:
    - Pre-acceptance test problems
    - Facility or test equipment problems
    - Non-flight configuration problems
2. Data Search. The data search should begin with the problem report databases and include other applicable problem reports (such as NASA reports, contractor data). As a minimum, the data query should include:
    - Calendar period or mission number
    - FMEA code

- o LRU part number
- o Word search for failed component or failure mode

3. Manual Activities. Manual activities include, but are not limited to:
   - o Excluding nonapplicable problems
   - o Reading problem reports to verify correct failure modes
   - o Reviewing FMEA for assignment of new criticality categories
   - o Obtaining time or cycle data or number of units inspected or tested for normalization

E.5.6.1.2. Step 2: Construct A Subsystem or Component-Level Normalized Trend Analysis Chart

The chart includes all problems (except those excluded by ground rules) on a selected subsystem or component, without identification of failure modes. Prior to trend analysis, the problem frequency is normalized by run time, cycles, sets of mission flight or orbital hardware, inspections, or other parameters. Both the raw data (quantity of problems) and normalized data are displayed (example, Figure E.5-6). The trend direction (normalized data) may be determined by observation, or either a linear trend line or curve may be plotted. Trend direction is established by plotting all failure modes; a single corrective action is not applicable. The trend direction is observed only for information relative to overall condition of the subsystem and component.



	1979	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989
Raw data	56	92	69	139	98	92	112	83	73	97	103
Normalized data	2.34	2.13	2.14	3.74	3.41	3.47	2.77	6.53	1.75	2.14	2.83

Figure E.5-6: SSME HPOTP Subsystems

E.5.6.1.3. Step 3: Construct a Pareto chart of Failure Modes or Causes and Identify Area(s) of Concern

1. The Pareto chart (example, Figure E.5-7) shows frequency of all observed failure modes or causes and identifies each failure mode or cause that is (from an engineering viewpoint) an area of concern. If the database cannot sort data by failure mode or cause, it may be necessary to read each problem report on a failed component. Reviewing problem reports also may be necessary when cause codes are available because different engineers can assign different failure mode codes to identical failures.

4. As a minimum, the Pareto chart should indicate the following for each area of concern failure mode:

   f. Quantity of criticality 1 and 1R problem reports by failure mode

   g. Percent of all problem reports by failure mode

   h. Quantity of problems reported by year (or mission)

   i. Problem report closure status (quantity open and quantity closed)

   j. Date of last failure

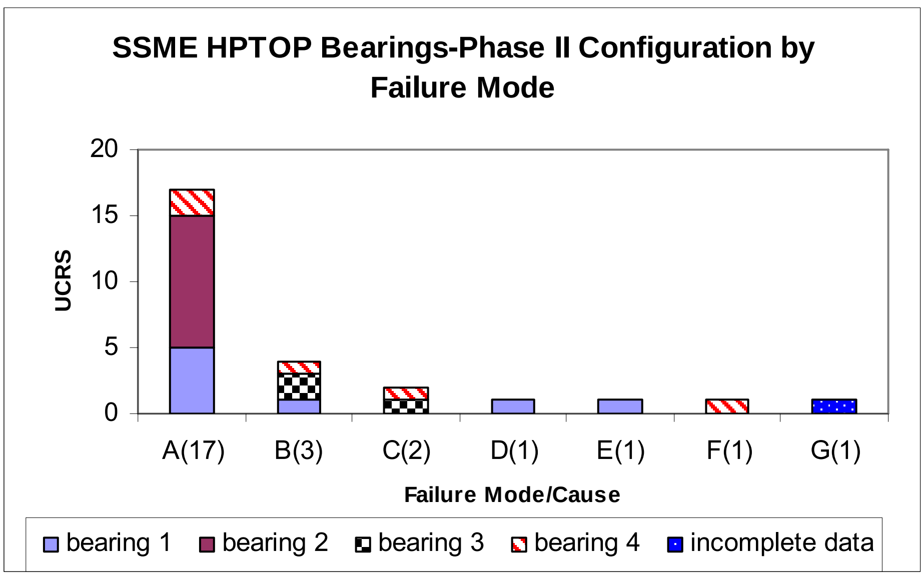
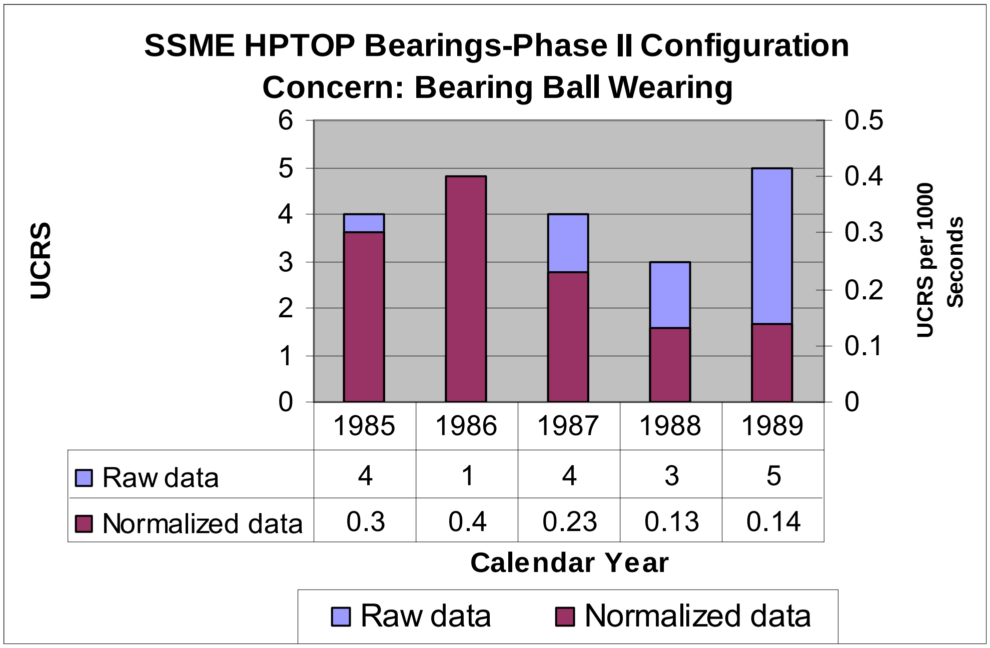


Figure E.5-7: SSME High HPOTP Bearings-Phase II Configuration by Failure Mode

E.5.6.1.4. Step 4: Construct a Normalized Problem Trend Chart for Area(s) of Concentration

A chart (example, Figure E.5-8) is prepared for each failure mode or cause identified as an area of concern. Chart preparation should consider data normalization, R-squared values, design, process, procedure changes, and engineering judgment.



	1985	1986	1987	1988	1989
Raw data	4	1	4	3	5
Normalized data	0.3	0.4	0.23	0.13	0.14

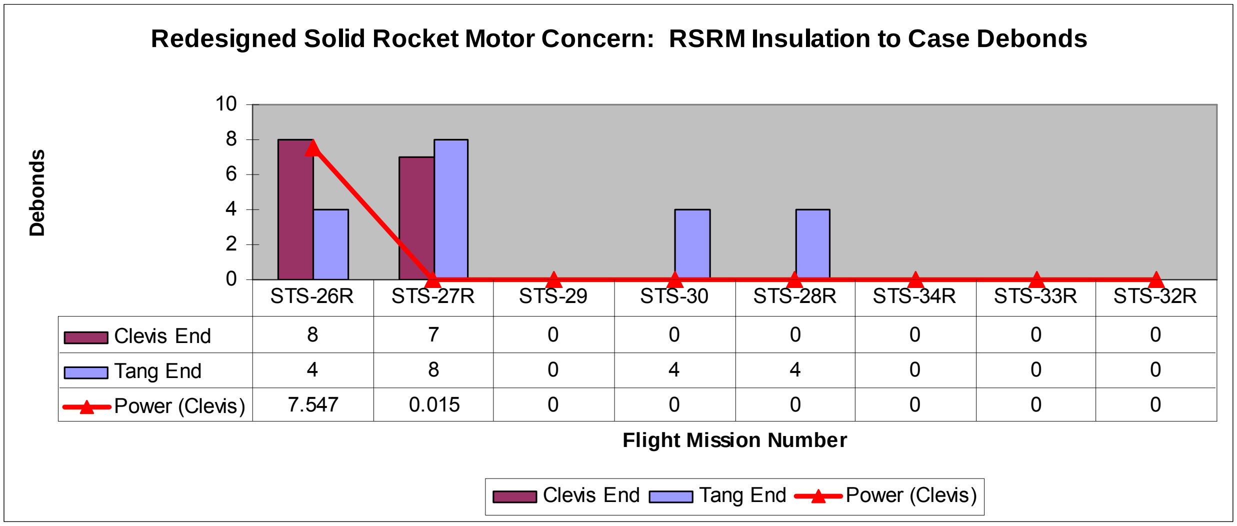
Figure E.5-8: SSME HPOTP Bearings-Phase II Configuration Concern: Bearing Ball Wear

1. Data Normalization. It is important to normalize trend data whenever possible to eliminate misleading trends. Usually, low-cycle fatigue problems are normalized by exposure cycles (quantity of tests), and high-cycle fatigue problems by operating time of exposure. In the event that problem reporting in a given area is reduced or discontinues, consideration should be given to normalizing for the reduced reporting.

   For example, if 20 percent of applicable problems during and after the acceptance test procedure were due to a process that is no longer reported, the subsequent trend data should be adjusted upward (multiplied) by 1.00 / 0.80 = 1.25.

2. R-squared Values. For each trend, only the models for which the fitted points have no negative values can be candidates for selection. When R-squared values for any of the five models (linear, exponential, power, logarithmic, or positive parabolic) are approximately the same (difference ≤ 0.020), the one that best fits the extreme right data point would be selected.

3. Design, Process, and Procedure Changes. Design, process, or procedure changes that could eliminate the failure mode should be shown at the appropriate point on the trend chart (example, Figure E.5-9). Usually, it is desirable to show raw data and normalized data both prior to and after the design change on a failure mode trend chart. Only the normalized data are trended. It is not recommended to show a trend line or curve on the trend chart unless the trend is declared statistically increasing or decreasing. It is important to determine trend direction after the last major change point.



	STS-26R	STS-27R	STS-29	STS-30	STS-28R	STS-34R	STS-33R	STS-32R
Clevis End	8	7	0	0	0	0	0	0
Tang End	4	8	0	4	4	0	0	0
Power (Clevis)	7.547	0.015	0	0	0	0	0	0

Note: Clevis end statistically significantly downward trend R-squared (power) = 0.74, Tang end statistically non-significant trend

Figure E.5-9: Redesigned Solid Rocket Motor Concern: RSRM Insulation to Case Debonds

4. Engineering Judgment. If the failure mode, root cause, and corrective action are well understood and the number of subsequent tests (or seconds or inspections) without failure is considered sufficient, trends with few data points that have ended with zero failures may be declared as downward.

   a. The example illustrated in Figure E.5-10 involves quantities of case-to-insulation debonds on the Reuseable Solid Rocket Motor (RSRM) based on occurrences on successive sets of mission flight hardware. The plotted data indicate process change points on RSRM segments. Engineering knowledge of the changes plus six clevis end failure-free flights after the grit-blasting change indicates a statistically verified downward trend. Although initially considered downward, the tang end trend is not statistically significant and, therefore, is identified as an adverse trend

   b. Figure E.5-9 is a backup chart useful to show location of trended problems (in this case, by flight vehicle and RSRM segment)

E.5.6.1.5. Step 5: Prepare Summary Assessment of Problem Trend Analysis with Recommendations

The sample summary assessment is provided in Figure E.5-11. The following are proposed inputs for a summary assessment:

1. Data source if other than cognizant Center PRACA database. If applicable, provide ground rules for excluded problem reports
2. Component and failure mode(s) trended, including quantity of problem reports
3. CIL Code Number
4. Failure mode(s) criticality and date of last failure

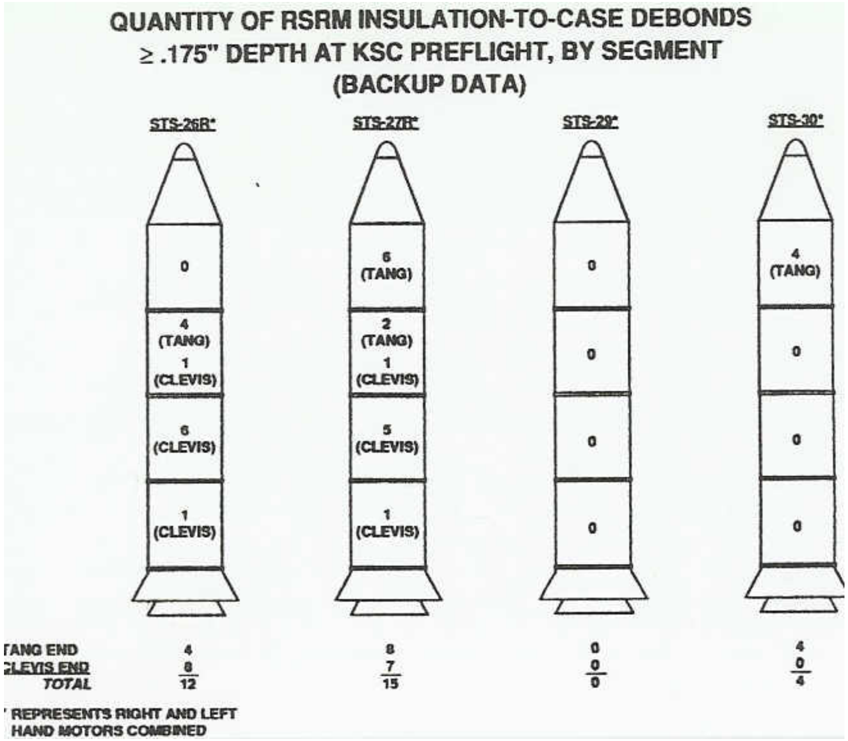


Figure E.5-10: Quantity of RSRM Insulation-to-Case Debonds ≥ 0.175" Depth at KSC Preflight, by Segment (Backup Data)

5. Primary failure cause or subcause
6. Design, process, and procedure changes, with effectivity. Indicate if any data prior to such changes are excluded
7. Trend direction (increasing, level, or descending)
8. Trend evaluation (adverse, acceptable, or favorable)
9. Recurrence control action

If applicable, include a statement regarding additional data (trend analysis update) needed to evaluate the trend direction.

As applicable, recommendations based on engineering analysis of the trend and a statement regarding additional resources required to correct an adverse trend should be included. When the failure mode for the area of concern can be characterized by a variable (dimension, load, voltage), recommend performance trend analysis of the variable versus run time, cycles, or inspections. An option is to correlate the variable with influence parameters (pressure, temperature, and critical dimension).

E.5.7 Reporting

The format described and illustrated in Step 5 in the process (E.5.6.1.5) should be used in reporting problem trend analysis.

Summary Assessment (Sample)

Failure Mode: HPOTP – Loss of support or position

This failure mode is FMEA Criticality 1.

CIL Item number: B400-13

Failure cause A: HPOTP phase II bearing anomalies

Failure subcause #1:

- Bearing ball wear:

  17 UCRS: Most recent failure occurred in September 1989. Excessive wear caused by low to negative coolant vapor margin. At least 10 of these 17 UCRS were written on pump – end bearing #2. The latest recurrence control is to limit bearing operating life to 2568 seconds by Deviations Approval Request with replacement of the 4 HPOTP bearings prior to each flight. Trend is adverse (Level).

- Recommendation:

  Rocketdyne, Pratt & Whitney, and MSFC direct bearing testing so as to identify design changes that would increase bearing life by decreasing ball wear. Performance trending of ball wear vs. run time and correlations of ball wear with influence parameters such as internal clearance and lox coolant flow should be updated.

Figure E.5-11: Summary Assessment (Sample)

E.5.7.1. Frequency

The frequency of problem trend analysis reporting is determined by program needs; as a minimum, an overall problem or project trend analysis should be reported monthly. Cyclic programs and projects such as Space Shuttle missions also should report problem trend analysis based on cycles. Where programs are comprised of major elements, the elements should be reported in addition to the overall project reporting requirements.

E.5.7.2. Reporting Results

Each trend analysis organization should establish a method of dissemination that meets their specific requirements. When reporting problem trend analysis results in support of management decisions, include the following activities: coordinate early trend analysis products; establish a routine periodic hard copy distribution of current trend charts; as applicable, maintain a display of selected current trend charts; provide trend charts for real-time support of mission reviews; and provide immediate distribution of charts identifying adverse trends.

E.5.7.3. Maintaining Problem Trend Analysis Status

When selection of items for trend analysis is complete, it is essential to maintain a status or accounting system. A suggested format for this effort is provided (see Figure E.5-12). Descriptions of column headings are as follows:

Element. Selection criteria for items trended

Planned. Number of deficient hardware items to be trended. Some planned items may not be trended because of insufficient data points, redesign, or other uses. The quantity in this column is equal to the sum of the next three columns

Currently Trended. Number of items for which at least one trend chart exists

In-Process. Number of items for which trend analysis is underway but no trend chart exists

Inactive. Number of items planned for trend analysis, but which are neither trended nor in-process. (This category may include items that were trended, but have been temporarily discontinued)

Remarks. Any pertinent explanatory notes

Problem Trending Program Status					
Element	Planned	Currently Trended	In Process	Inactive	Remarks
SSME					
F- Frequency	10	5	1	4	
C - Criticality	2	2	0	0	
E - Engineering	19	18	0	1	
M - MSFC	4	4	0	0	
ET					
F- Frequency	8	5	0	3	
C - Criticality	1	1	0	0	
E - Engineering	1	1	0	0	
M - MSFC	0	0	0	0	
SRB					
F- Frequency	13	2	0	11	Inactive items are low criticality and frequency
C – Criticality	1	1	0	0	
E – Engineering	1	0	0	1	
M – MSFC	0	0	0	0	
RSRM					Inactive items are low criticality, low frequency, and awaiting results of substantial redesign
F- Frequency	16	0	0	16	
C - Criticality	2	2	0	0	
E - Engineering	0	0	0	0	
M - MSFC	1	1	0	0	
Grand Total	**79**	**42**	**1**	**36**	

Note: Quantities on this table address the number of series of items. A series is identified on series flow chart.

Figure E.5-12: Problem Trending Program Status

E.6 Supportability Trend Analysis

E.6.1 Introduction to Supportability Trend Analysis

Supportability trend analysis assesses effectiveness of logistics elements in supporting NASA programs and projects. Supportability trend analysis focuses on the recurrence of logistics problems and the effective control of these problems. Common logistics elements include, but are not limited to: maintenance, supply support, support equipment, facilities management and maintenance, support personnel and training, packaging, handling, storage, and transportation, technical data support, automated data processing hardware and software support, and logistics engineering support.

E.6.2 Objectives of Supportability Trend Analysis

The primary objectives of supportability trending analysis are to:

- Monitor the current health of support systems
- Forecast support problems to enable resolution with minimum adverse effect
- Determine which support elements can be improved to optimize the system availability over its operating life
- Measure effects of system reliability and maintainability on supportability and identify areas for improvement, and analyze current support systems to estimate future requirements
- Identify the relationship between support and other program or project factors

E.6.3 Candidates for Supportability Trend Analysis

Because elements of supportability trend analysis are based on the common elements of logistics support and logistics engineering, the candidates for this analysis are generally well known. Candidates for trend analysis should be selected to provide an accurate measurement of the effectiveness of the support elements and the reliability and maintainability design factors. Examples of common candidates for supportability trend analysis include: repair turnaround time, scheduled maintenance activity, unscheduled maintenance activity, modifications, zero balance inventory items, cannibalization, technical document changes, fill rate, impending loss of spare and repair capability, personnel skill adequacy, and repetitive failures. Examples of supportability trend analysis candidates used to evaluate system reliability, maintainability, and availability support characteristics include: mean time between failures (MTBF), mean time to repair (MTTR), and mean time between repairs (MTBR). Priorities should be established based on the area of concern (risk, safety, cost, availability, and schedule) and the expected benefits of the trend analysis. Where risk criticality is a primary concern, criticality 1 items should be given highest priority followed by criticality 1R and 1S items.

A prime concern in supportability trend analysis is the determination of the extent of analysis and identification of the appropriate parameter variation that should be measured. Selected parameters should be measurable and capable of showing sufficient variation to be useful in monitoring the factor under analysis.

E.6.4 Data Sources for Supportability Trend Analysis

Data sources for supportability trend analysis may be found in: equipment problem reports, work authorization documents, contractual acceptance records, shipping and receiving reports, payment records for maintenance, transportation records, inventory and issue turn-in records, training course attendance records, technical documentation error reporting, and consumable replenishment records.

Each program or project should recognize the relationship between these data sources and the supportability factors. Recognizing the relationships should lead to an understanding that analysis of supportability data is often as important to a program or project as performance data.

E.6.5 Considerations of Supportability Trend Analysis

There are many factors to be considered for a supportability trend analysis, including: maintenance operations, selection criteria, line items and spare parts, indirect indicators, complementary data, trend limits, normalization factors, causes of delayed data, data accuracy, and corrective action.

E.6.6 Procedures for Supportability Trend Analysis

The basic steps in supportability trend analysis are:

a. Analyze the operations and support systems to identify items that could lead to a system failure, schedule delay, or cost increase if support degrades

b. List these items as candidates for supportability trend analysis

c. Select items from the list of possible candidates. Provide the list of items to the program or project office

d. Determine the parameters to be used in judging whether the item's supportability is fluctuating at a rate sufficient to warrant management attention. When these parameters are critical to safety or mission success, strong consideration should be given to the feasibility of performing trend analysis

e. Determine if measurement data are available for the selected supportability parameters. Supportability parameters may be directly measurable factors or the relationships between two or more parameters based on an algorithm. If measurement data are not available, determine the feasibility of establishing a system to measure the parameters

f. Establish the supportability baselines and limits. Original baselines and limits should be taken directly from program and project support requirements. The following documents are examples of the type of sources that should be reviewed to determine what values represent acceptable supportability for each indicator: logistics support plans, design criteria, program requirements documents, specifications, and intermediate and depot maintenance requirement documents

g. Determine the measurements necessary to evaluate the chosen parameters

h. Collect, measure, and record the data and perform a supportability trend analysis to determine if the parameter being trended exceeds the historical limits or falls below the supportability baseline. If so, immediate management attention may be needed to correct the situation. If the values are within limits but the trend indicates that they may exceed the limits in the future, this early warning would allow management to implement preventative measures before the situation deteriorates

i. Report the results using charts, graphs, and recommendations

E.6.7 Reporting for Supportability Trend Analysis

E.6.7.1. Format

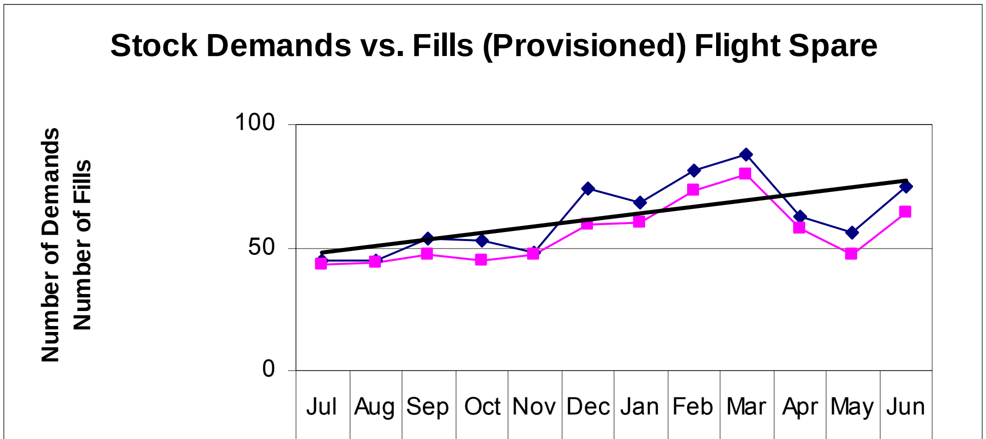
The supportability element chart should depict an historical trend of substantiated data on the characteristic being measured with realistic program or project control limits for that subsystem or repair location. When an adverse trend has been identified or exceeded, a detailed analysis should be provided, including a discussion of what corrective action, if any, is required.

E.6.7.2. Basic Supportability Analyses Examples

The following listed items provide examples of common supportability trend analysis reports that are used. These examples are not the only forms of supportability trend analysis that can be performed and reported. For simplicity, months are used to exemplify time periods and missions to exemplify events. Where reusable vehicles are involved, vehicle differences may require analyses by vehicle as well as overall analyses by vehicle type.

a. LRU, Spares, or Line Item Demands Filled per Month, Mission, or Vehicle. This sample report analyzes the number of demands that were filled for LRUs or spares or line items, generated by planned and unplanned work requirements. Analyses of line items should clarify whether or not the numbers reflect the quantities of each line item (refer to Figure E.6-1)

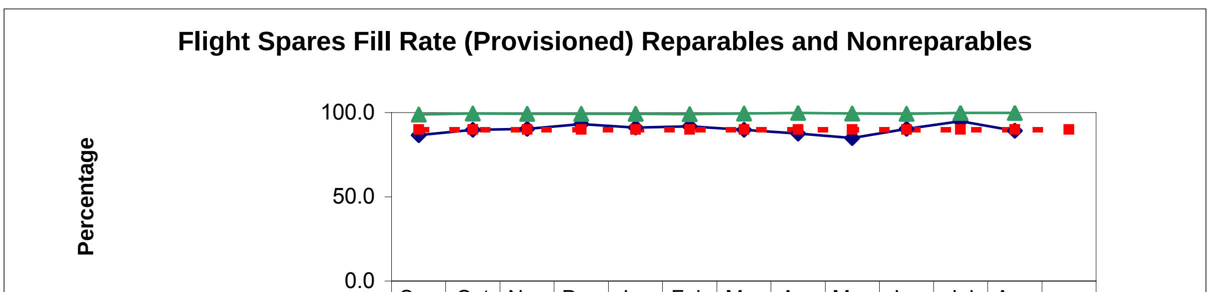
b. LRU, Spares, or Line Item Demand Fill Rate per Month or Mission. The previous report is useful for inventory management; this sample report (Figure E.6-2) is most useful as a measure of effectiveness for the supply support system. This report displays the data from the previous report on a percentage scale on the ordinate (y axis) and time or event or mission sequence on the abscissa (x axis). By measuring the percentage of the demands actually filled, this report shows the ability of the support system to meet the demand for replacement items. Normally, a supply support system cannot meet all demands; therefore, a program or project goal or limit is set, based on a trade-off of cost and availability. This analysis shows supportability of the supply system relative to the program or project goal. As a form of supportability trend analysis, this report can be used to anticipate when a supply support system may degrade below the acceptable probability of sufficiency factors specified in program or project documents



	Jul	Aug	Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May	Jun
Demands	45	45	54	53	48	74	68	81	88	63	56	75
Fills	43	44	47	45	47	59	60	73	80	58	47	64

Note: Data source XYZ Corps
Contact: Code WE X1905
Data as of: 01/91
Prepared 01/91
Analysis: Spares inventory continues to support on-site stock demands
Recommendation: Continue to monitor fill rate, no action required

Figure E.6-1: Stock Demands vs. Fills (Provisioned) Flight Spare

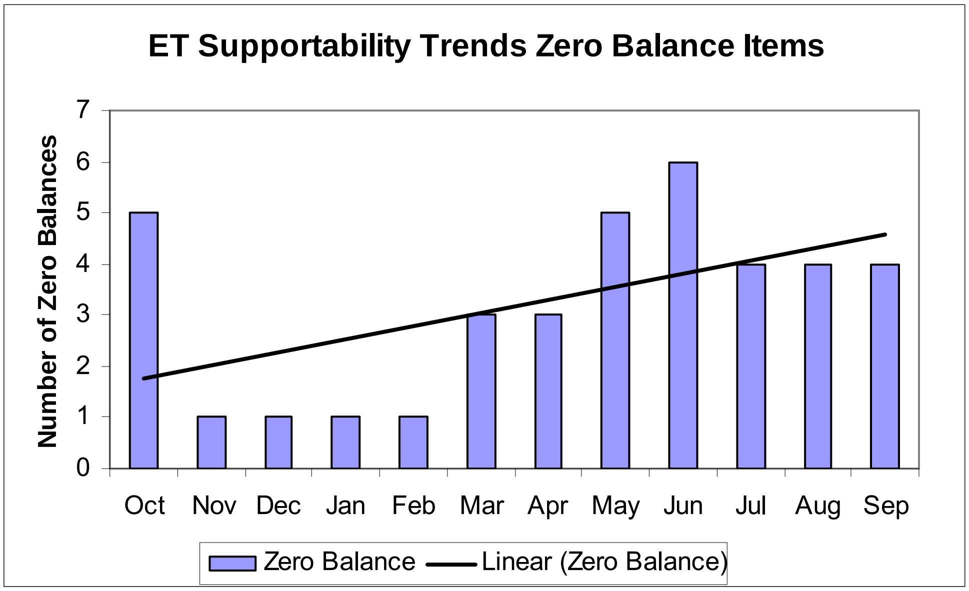


	Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	
Repairables	86.6	89.8	90.3	93.1	91.0	91.9	89.8	87.5	84.9	90.4	94.9	89.1	
Program Goal 90%	90	90	90	90	90	90	90	90	90	90	90	90	90
Nonrepairables	98.9	99.5	99.2	99.2	99.2	99.0	99.4	99.8	99.5	99.2	99.8	99.8	

Note: Data source XYZ Corps
Contact: Code WE X1905
Data as of: 01/91
Prepared 01/91
Recommendation: Continue to monitor fill rate and repair turnaround time (RTAT)

Figure E.6-2: Flight Spares Fill Rate (Provisioned) Reparables and Nonreparables

c. Zero Balance. This sample report (Figure E.6-3) provides the trend of out-of-stock line items in the spares and supply inventory of provisioned items. Historical and projected trends are included. The total number and individual part numbers may be detailed by criticality codes such as 1, 1R, or 1S

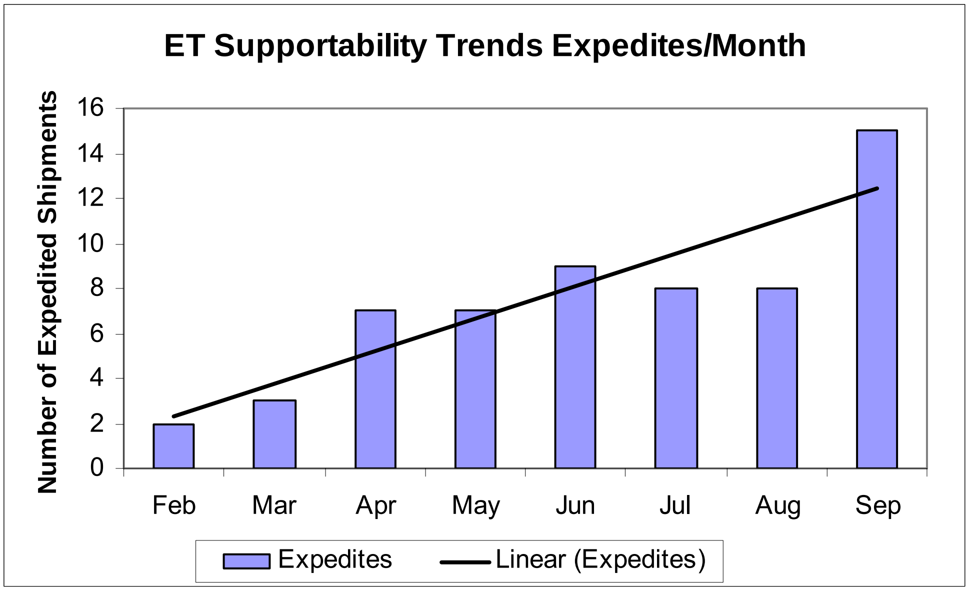


Note: Data source XYZ Corps
Contact: Code WE X1905
Data as of: 01/91
Prepared 01/91
Definition: Zero balance indicates no spares in stock: All spares have been consumed.
Analysis: Trend is positive indicating a possible impact to supportability
Recommendation: Continue to Monitor

Figure E.6-3: ET Supportability Trends Zero Balance Items

d. Expedite Actions Per Month. An expedite request should be filled within 24 hours. This sample report (Figure E.6-4) shows the expedite supply actions by month for the past year and highlights the top 10 expedite requests (whether filled or not), including those replaced by cannibalization action or withdrawn when they were not filled. Specific items that required two or more expedite actions during the past year often are reported

e. Number of Items Cannibalized per Month or Mission. This sample report (Figure E.6-5) provides a history of the number of cannibalized items by month and mission or event with projected trends. This information is presented in a line graph report with detailed part number listings as background data



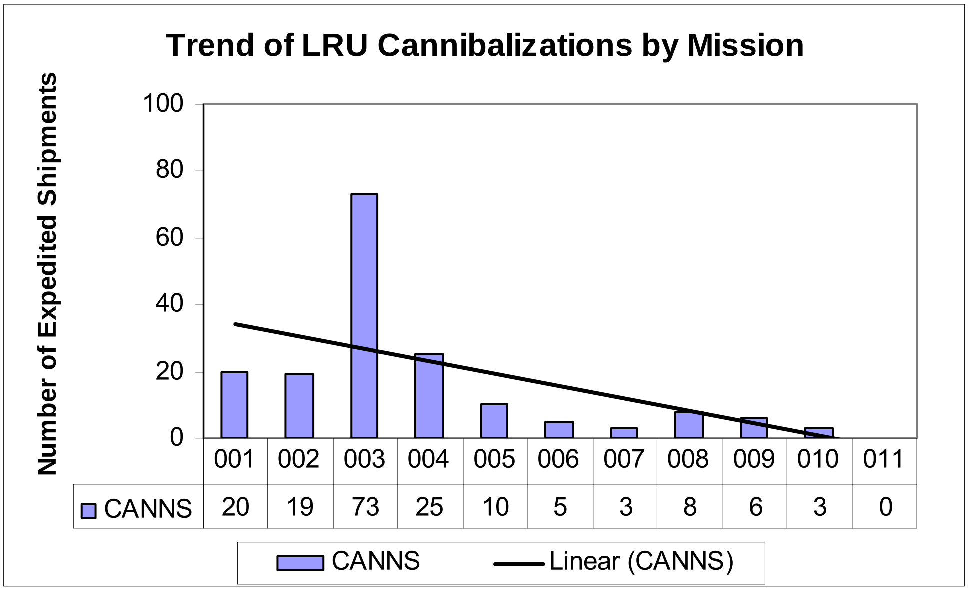
Note: Data source XYZ Corps
Contact: Code WE X1905
Data as of: 01/91
Prepared 01/91
Definition: Data not tracked prior to February 1990
Analysis: Upward trend due to launch site problems: Hydrogen leak
Recommendation: Continue to Monitor

Figure E.6-4: ET Supportability Trends Expedites per Month

f. Maintenance Tasks per Month or Mission. This sample report (Figure E.6-6) details the total number of scheduled or unscheduled maintenance tasks and modification tasks competed per month or mission

g. Maintenance Tasks Completed, Deferred, or Waived. This sample report (Figure E.6-7) supplements the previous one by comparing completed tasks with the deferred and waived tasks. The breakout of tasks shows capability of the support program to maintain a repetitive operation. As an example, if the overall number of completed tasks tends to remain level while the number of deferred tasks increases, program management has an indication that the support system does not have the required capacity. The shortfall is being accommodated by the increasing number of deferrals

h. Requirement Changes per Month or Mission. This sample report (Figure E.6-8) shows the number of Operations and Maintenance Requirements and Specifications Document (OMRSD) requirements changes per month or mission. It delineates the number of changes submitted versus approved for each major element (such as work package, major system, power system, Orbiter, External Tank (ET), Space Shuttle Main Engine (SSME)). This report also can show the number of waivers and exceptions by month or mission, and the number of new requests



Note: Data source XYZ Corps
Contact: Code WE X1905
Data as of: 01/91
Prepared 01/91
Analysis: Decrease in cannibalizations due to January 1987 provisioning policy change
Recommendation: Monitor the spares system to ensure proper corrective action

Figure E.6-5: Trend of LRU Cannibalizations by Mission

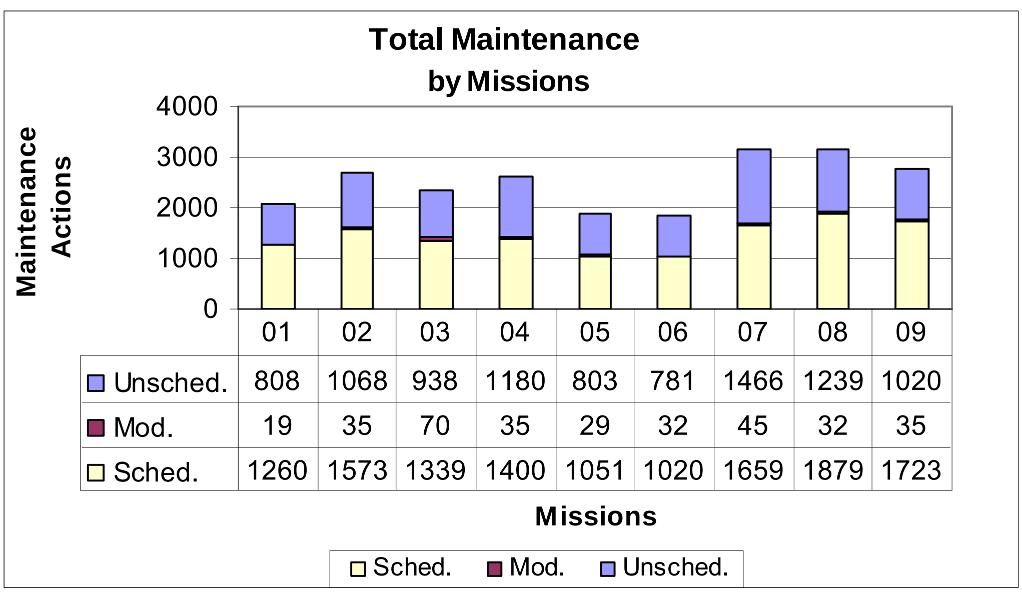


Figure E.6-6: Total Maintenance

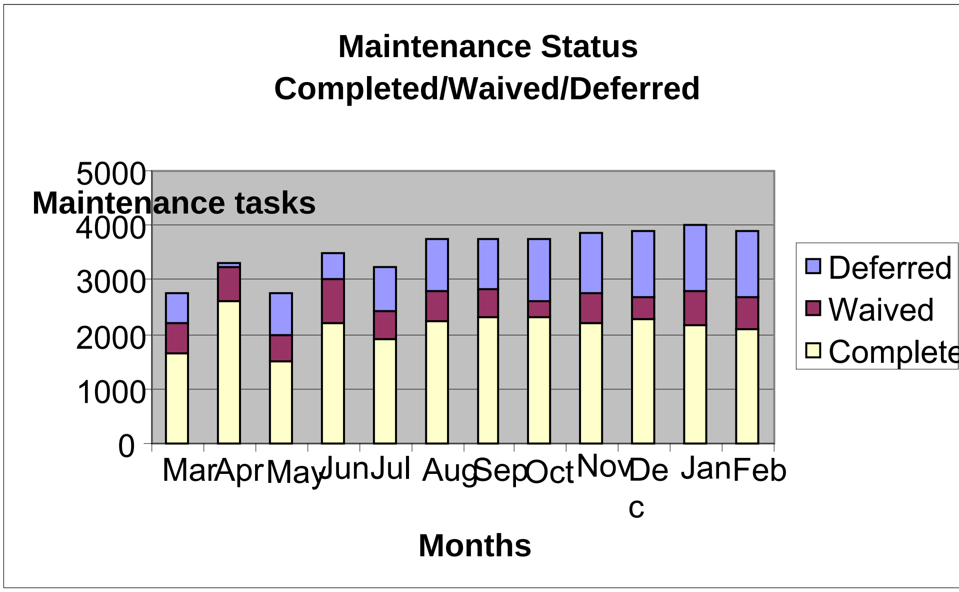
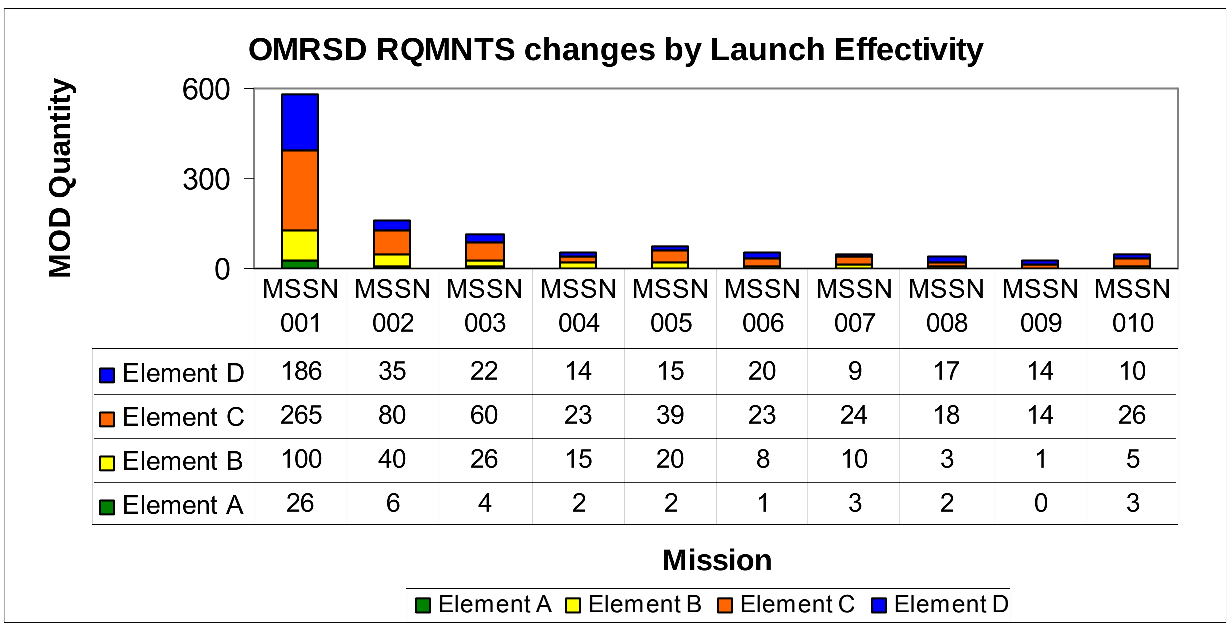


Figure E.6-7: Maintenance Status, Completed, Waived, or Deferred



	MSSN 001	MSSN 002	MSSN 003	MSSN 004	MSSN 005	MSSN 006	MSSN 007	MSSN 008	MSSN 009	MSSN 010
Element D	186	35	22	14	15	20	9	17	14	10
Element C	265	80	60	23	39	23	24	18	14	26
Element B	100	40	26	15	20	8	10	3	1	5
Element A	26	6	4	2	2	1	3	2	0	3

Note: MSSN 001 was purposely omitted from the flow average calculations.

Figure E.6-8: Operations and Maintenance Requirements and Specifications Document Changes by Launch Effectivity

i. Crew Maintenance Time per Month or Mission. This sample report (Figure E.6-9) shows the total number of workforce hours expended per month for on-orbit maintenance by the crew and the average number of hours per individual actually performing maintenance tasks. Control limits on crew time for space flight system maintenance are specified in the program or project function and resource allocation requirements. For launch-and-return missions, the maintenance should be normalized as maintenance time per flight hour.

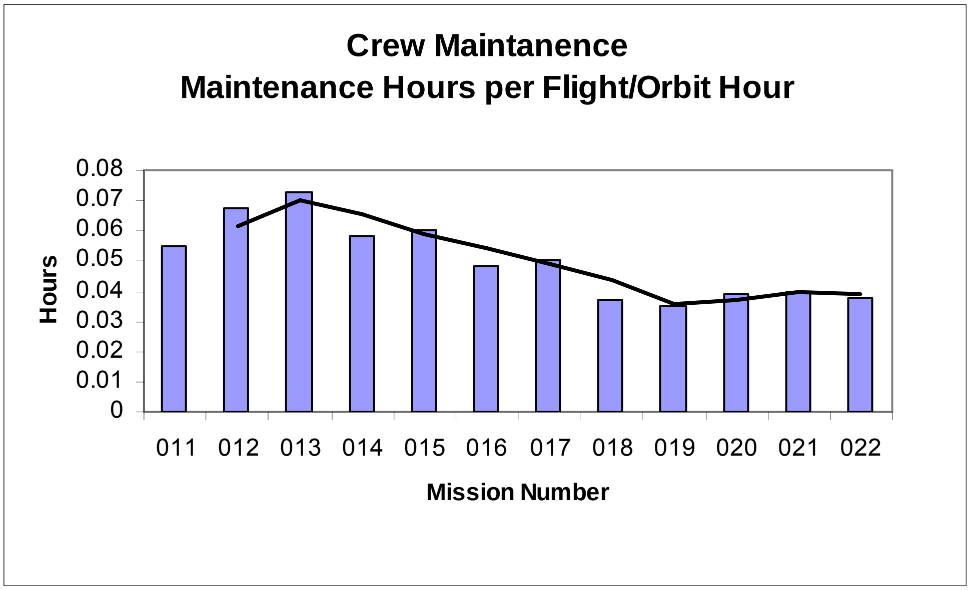
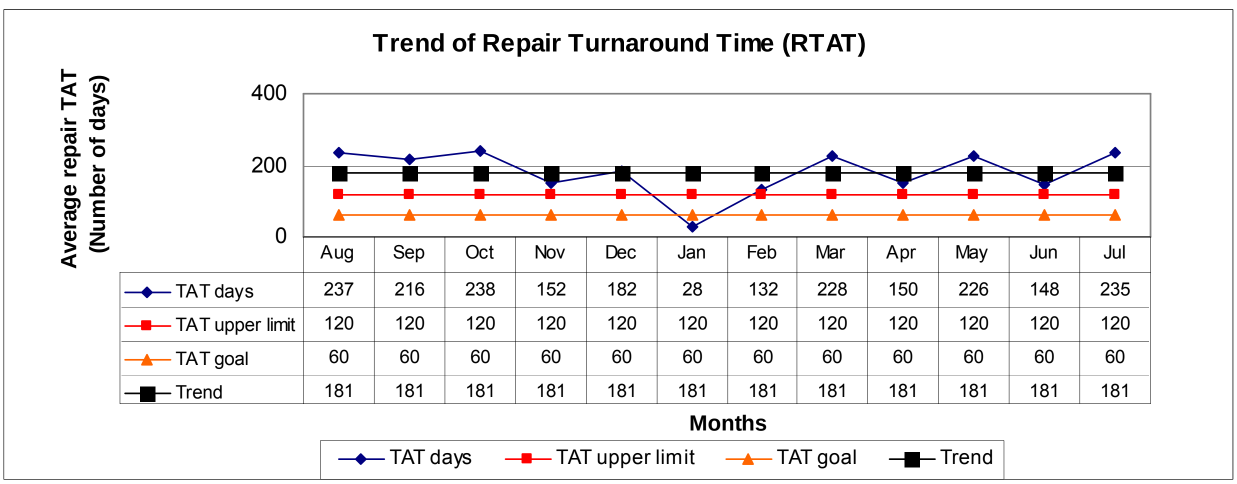


Figure E.6-9: Crew Maintenance

j. Turn Around Time (TAT) per Repair per Month. This sample report (Figure E.6-10) shows the status and trends for the repair TAT per month



	Aug	Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May	Jun	Jul
TAT days	237	216	238	152	182	28	132	228	150	226	148	235
TAT upper limit	120	120	120	120	120	120	120	120	120	120	120	120
TAT goal	60	60	60	60	60	60	60	60	60	60	60	60
Trend	181	181	181	181	181	181	181	181	181	181	181	181

Analysis: The sharp decrease in January 1989 RTAT reflects the lack of reported Original Equipment Manufacturer (OEM) activity and not reduced repair time. Since OEM RTAT is much higher than XYZ RTAT, decreased OEM activity will cause a decrease in the program RTAT trend. Recommendations: develop new RTAT trend charts. Base trend on open and closed Preproduction Readiness Reviews (PRR's).

Figure E.6-10: Trend of Repair TAT

k. Maintenance Action by Causes per Month or Mission. This sample report (Figure E.6-11) illustrates the breakout of support problem causes. It shows if any cause has an unfavorable trend in comparison to other causes

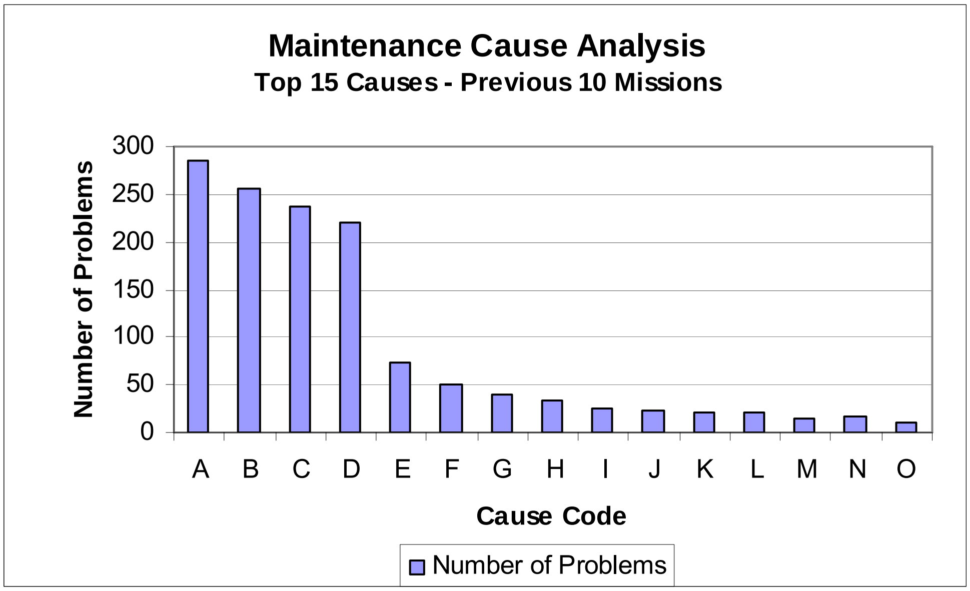


Figure E.6-11: Maintenance Cause Analysis

E.6.7.2.1. Frequency

The data analyses, trend charts, and the above sample reports should be made available to the program or project via regular and special reports as required. The program or project managers should establish routine reporting requirements. Once established, the trend reports should be updated at regular intervals, usually monthly and by mission or event. When trend data indicated rapid change or that timely availability of trend analysis is required, the trend reports may be prepared on a more frequent basis. Copies of the trend reports should be made available to the appropriate management individuals. NASA management should be alerted in a timely manner of any supportability trend analysis results that may affect safety.

E.7 Programmatic Trend Analysis

E.7.1 Introduction to Programmatic Trend Analysis

Programmatic trend analysis is a tool to assess program information such as schedule elements, employee utilization and attrition rates, overtime, noncompliance with operating procedures, equipment damage, mishaps or injuries, past program performance, and any similar data to identify problems in applying resources to comply with procedural requirements and management program schedules.

E.7.2 Objectives of Programmatic Trend Analysis

The principal objective of this analysis is to provide a medium that accurately and quantitatively monitors the programmatic posture and provides management visibility to determine the current and projected health of the human support element. Other important objectives include:

a. Increase management awareness of inappropriate demands on human resources (workload or schedules) required to support the program or project and associated hardware and software

b. Prevent possible compromises or delays in mission schedules caused by dysfunctional responses by the human element to stress

c. Support management in identifying schedule, human resource allocation, or experience or qualification mismatches that could have potential adverse effect on the program schedule or performance. This may require procedural, assignment, or schedule modifications to maintain or enhance performance

d. Support management in identifying areas requiring attention such as damage, mishaps, or injury rates. Determine the correlation over time or other potential program-related indicators

e. Support proposed program or project improvement changes

f. Support management in identifying and monitoring program or project management performance indicators over time to assure process controls. These indicators directly point out the ability of an end product to perform safely and reliably

E.7.3 Candidates for Programmatic Trend Analysis

Programmatic data should be used to monitor and report on (but is not limited to) the following areas:

- Workforce strength by specialty, experience, qualification, certification, and grade
- Personnel attrition and turnover rates by discipline
- Schedule changes, slippage, or overages
- Overtime usage versus approved policy
- Incidents such as damage, fire, mishap, or injury
- Requirement changes, including waivers and deviations
- System nonconformance and problems caused by human error
- Rework expenditures

E.7.4 Data Sources for Programmatic Trend Analysis

The data sources for programmatic trend analyses are more varied than for other types of trend analysis. In most cases, program or project offices maintain databases that provide appropriate data or have the potential to yield management performance indicators with minimal modification.

Excellent data sources for programmatic trend analysis may be found in: budget planning and expenditure reports, program or project schedules, quality assurance records, test and development status reports, inventory records, equipment problem reports, contractual acceptance records, shipping and receiving reports, work authorization documents, workforce status reports, resource utilization records, safety reports, and management information centers.

E.7.5 Procedures for Programmatic Trend Analysis

E.7.5.1. Standard Data

Each program or project should compile data and reports as discussed in section E.7.6 and generate supporting figures and graphs. Programs and projects should maintain the list of elements for which they will supply programmatic data, ensure the validity of the data provided for programmatic trend analyses, develop required analytical techniques and controls, and determine the structure for project data collection, maintenance, and reporting.

Data should be made available to program management and displayed on a separate chart for each programmatic indicator selected either for trend analysis or in aggregate data reports. If work unit codes are defined for the program, they may be used to identify or reference subsystems in an element.

Each chart should display an historical trend of substantiated data on the programmatic indicator(s) being measured along with the realistic control limits established for that indicator by the responsible program or project. When an adverse trend has been identified (whether apparent or not from the summary trend information) or a control limit has been exceeded as a result of a trend, an analysis of that trend should be conducted. Each program or project should accumulate data on programmatic indicators through completion and closeout.

E.7.5.2. Parameters

Suggested programmatic trend analysis indicators (section E.7.6) include:

- Workforce strength
- Schedule changes per month
- Overtime usage per month
- Incidents per month
- Requirement changes per month

However, programs may use other indicators. The appropriate program or project should define the indicator(s) to be used. Policy, work standards, or directives may set parametric limits.

E.7.6 Reporting

E.7.6.1. Format

Programmatic trend analysis should be prepared with sufficient detail to assist management in identifying problems and taking appropriate action. The minimum content and format for the reports are defined in this section. Reporting should highlight high risk and problem areas to aid in identifying needed improvements and program progress and health.

The following list of suggested programmatic trend analysis indicators may be expanded or modified as the program or project and programmatic trend analysis matures: workforce strength, schedule changes per month, overtime usage per month, incidents per month, and requirement changes per month. Other indicators may be tracked and maintained by the programs and projects at management’s discretion.

a. Workforce Strength. The number of personnel assigned to the program or project should be reported each month (Figure E.7-1) through the program management information system. A history of the number of personnel assigned to each program should be included in a graphical report of overall personnel totals by month. Additional charts (Figure E.7-2) should show personnel totals by discipline and by percent change of individuals. Trends of changes in personnel assigned by total and by disciplines should be compared with an overall average change rate to determine if unusual turnover is reflected. At least 12 months should be reflected in each monthly report

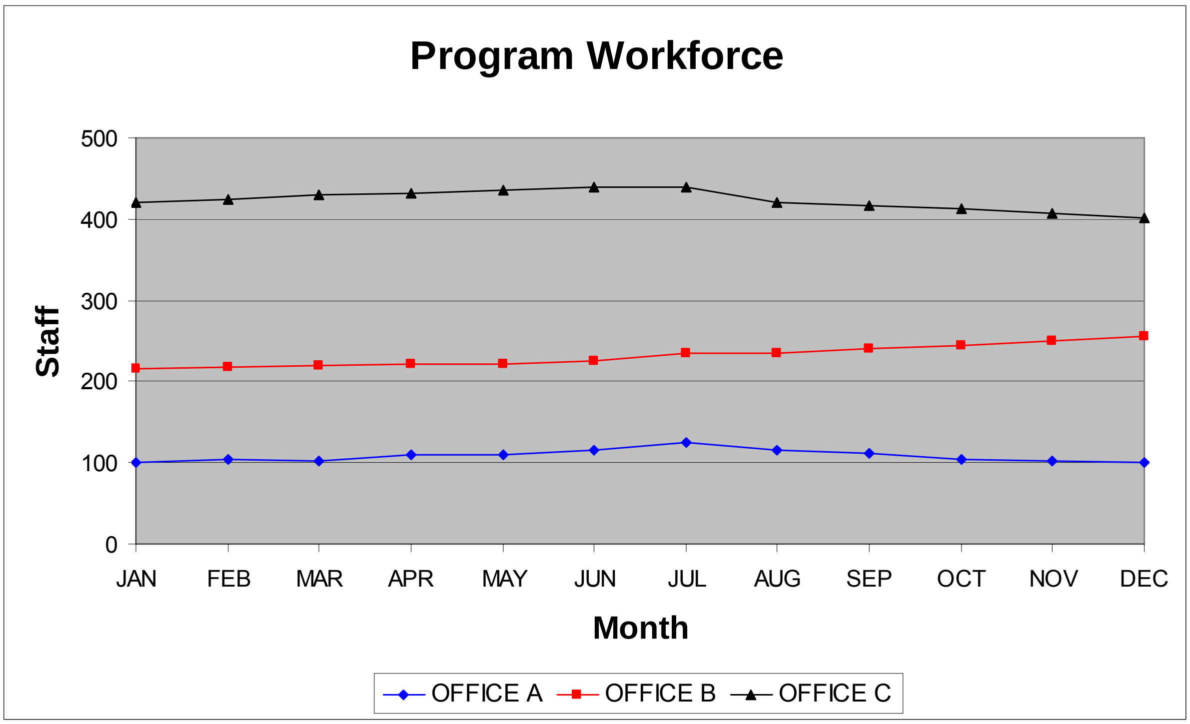


Figure E.7-1: Program Workforce

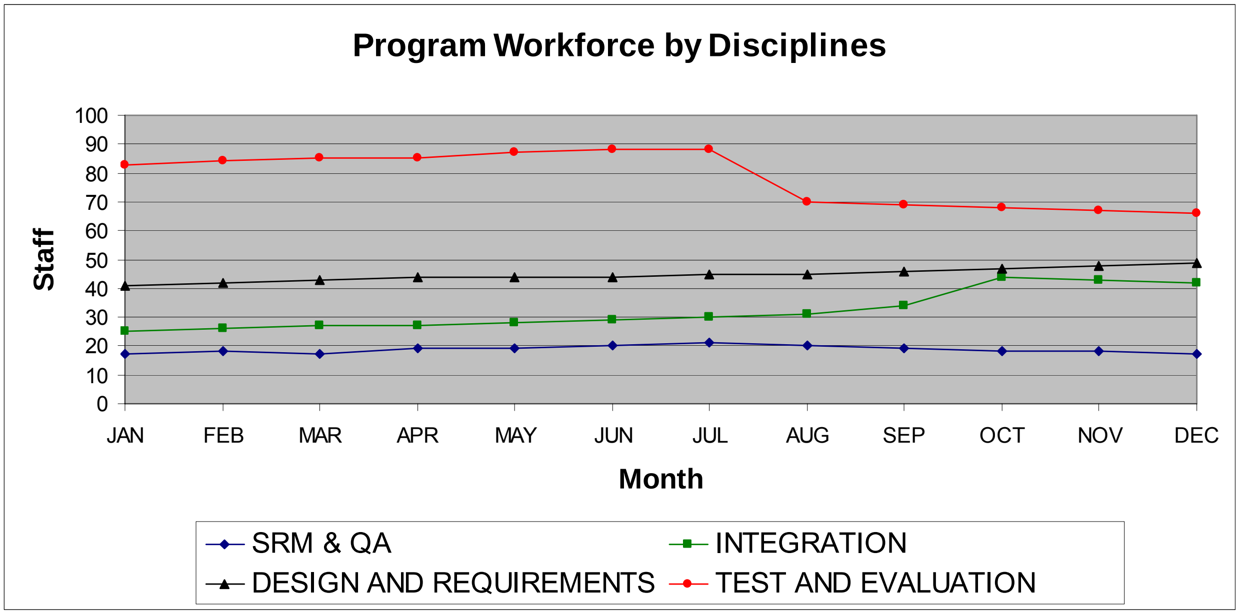


Figure E.7-2: Program Workforce by Disciplines

b. Schedule Changes per Month. This report (Figure E.7-3) should detail the schedule deviations per month for the past 12 months, including total number of schedule deviations and the average amount of monthly deviation. When a schedule for a particular activity or milestone is changed two or more times, the affected activity should be highlighted and explained in the monthly report

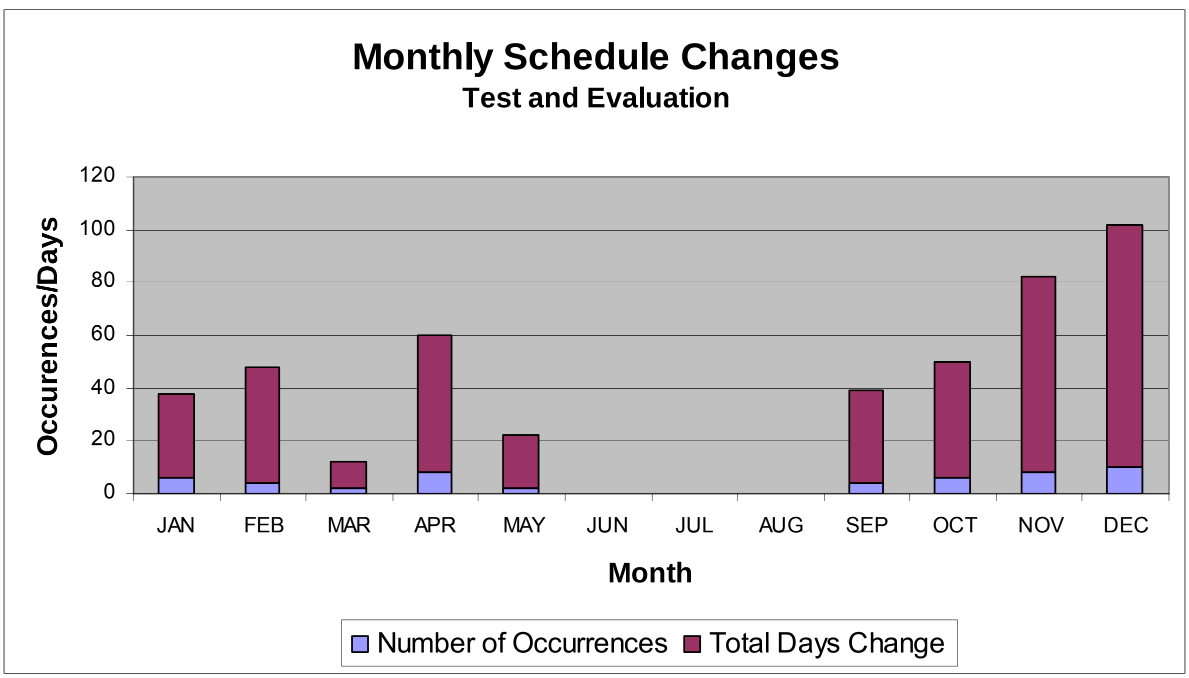


Figure E.7-3: Monthly Schedule Changes

c. Overtime Usage Per Month. This report (Figure E.7-4) should track the total amount of overtime beyond a 40-hour workweek

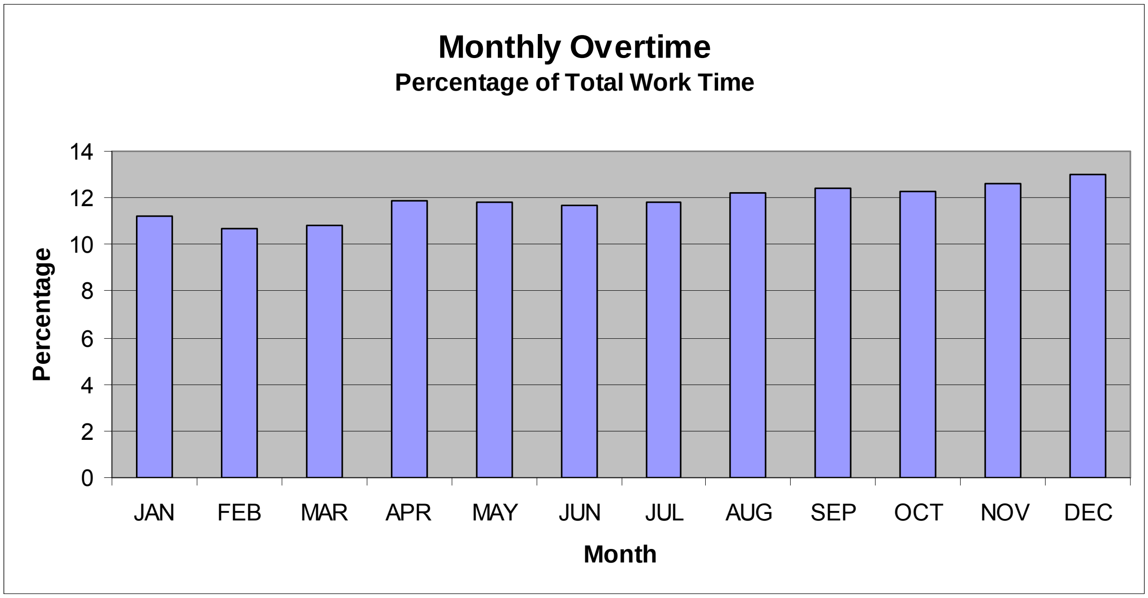


Figure E.7-4: Monthly Overtime

d. Incidents Per Month. This report (Figure E.7-5) should include the incidents per month for the preceding 12 months. The major elements of this report should be: damage, injuries, and major mishaps per A, B, or C category. Graphs should be presented to display the number of incidents and cost of each category, where applicable

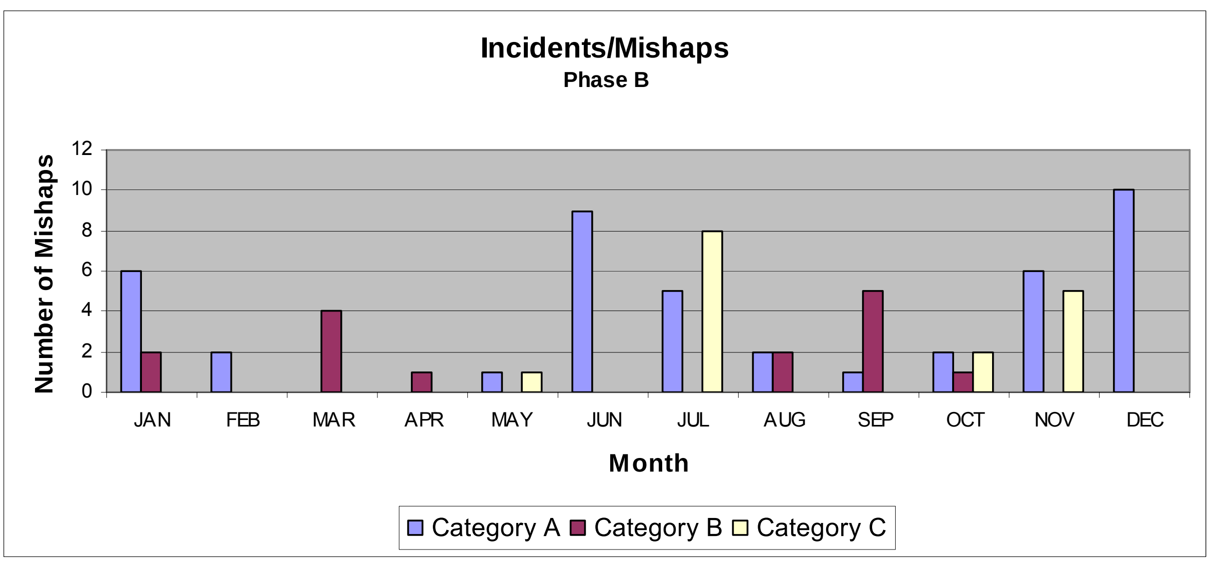


Figure E.7-5: Incidents and Mishaps

e. Requirement Changes Per Month. This report (Figure E.7-6) should show the number of changes to the top-level operational and maintenance requirements document per month for the last 12 months. It should delineate the number submitted versus the number approved, by major element. Waivers and exceptions, and the number of new requests, should be shown by month (Figure E.7-7)

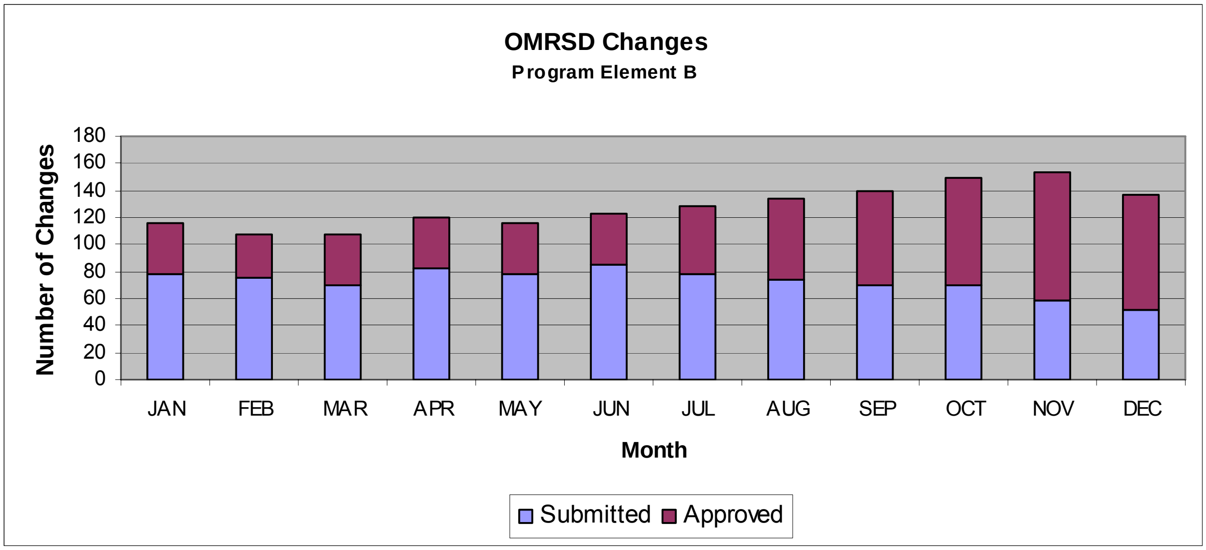


Figure E.7-6: OMRSD Changes, Program Element B

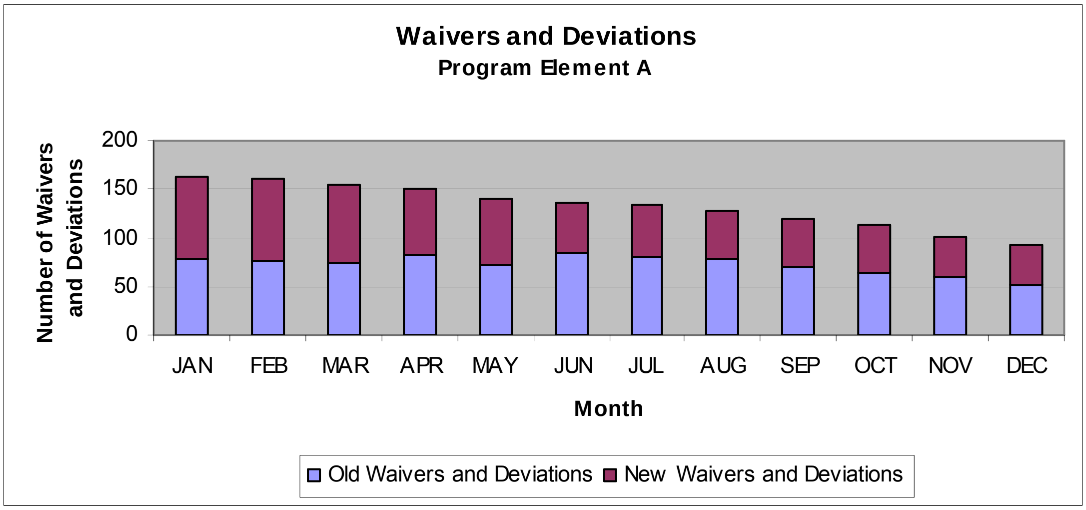


Figure E.7-7: Waivers and Deviations, Program Element A

E.7.6.2. Frequency

The cognizant program or project office should specify frequency of programmatic trend analysis.