

SAND95-2015C
The Systems Prioritization Method (SPM)
CD-ROM Demonstration for Waste Management '96

by

Curtis L. Harris
Los Alamos Technical Associates
Albuquerque, NM

D. M. Boak, N. H. Prindle
Sandia National Laboratories
Albuquerque, NM

Walt Beyeler
Science Applications, International Corp.
Albuquerque, NM

RECEIVED
MAR 22 1996
OSTI

ABSTRACT

In March 1994, the Department of Energy Carlsbad Area Office (DOE/CAO) implemented a performance-based planning method to assist in prioritization within the Waste Isolation Pilot Plant (WIPP). Probabilistic performance calculations were required for the Systems Prioritization Method (SPM) and roughly 46,700 combinations of activities were analyzed, generating a large volume of information to be documented, analyzed, and communicated. A self-contained information management system consisting of a relational database on a 600-megabyte CD-ROM was built to meet this need. The CD-ROM was used to store performance assessment results, data analysis and visualization tools, information about the activities, electronic copies of 40 CFR 191 and 40 CFR 268, technical reference papers, and the final SPM report. Copies of the CD-ROM were distributed to interested members of the public, WIPP participants, and the Environmental Protection Agency (EPA).

INTRODUCTION

In March 1994, DOE/CAO implemented a performance-based planning method to assist in prioritization within the WIPP project with respect to applicable EPA long-term performance requirements in 40 CFR 191.13(a) and 40 CFR 268.6.¹⁻⁵ SPM was designed to define the most cost-effective combinations of scientific investigations, engineered alternatives, and waste acceptance criteria to support the WIPP compliance certification application. Probabilistic performance calculations were required and roughly 46,700 combinations of activities were analyzed in the second phase of SPM (SPM-2). A self-contained information management system consisting of a relational database on a 600-megabyte CD-ROM was built to meet this need.

MASTER

THE SPM CD-ROM INFORMATION SYSTEM

In order to make this very large volume of data available and meaningful to a wide audience, SPM-2 results were published on a CD-ROM in the form of a self-contained information system. The CD-ROM operates on WINDOWS™ platforms. To minimize incremental costs of its use, all software (except WINDOWS™) is contained on the CD-ROM in a form that can be reproduced without violating copyright protection. A relational database on a 600-megabyte CD-ROM is used to store performance assessment results, data analysis and visualization tools, information about the activities, electronic copies of 40 CFR 191 and 40 CFR 268, technical reference papers, and the final SPM report (Figure 1). Information contained on the CD-ROM would, if printed out, create a stack of paper over three hundred feet tall.

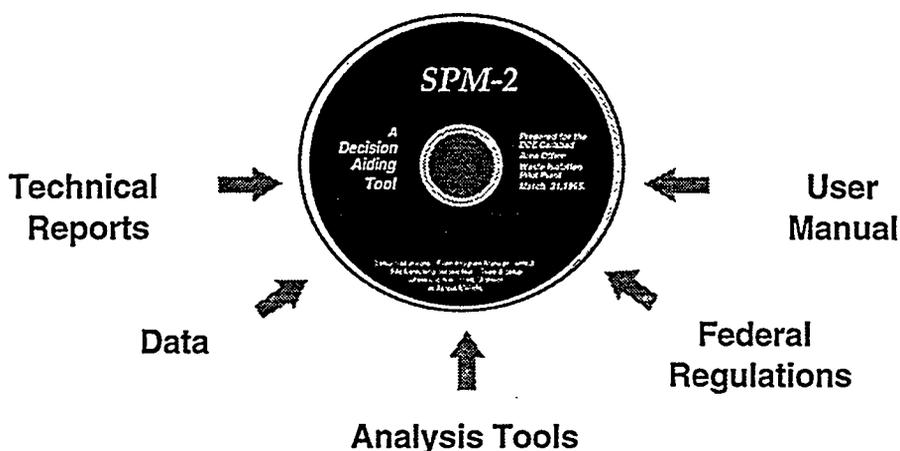


Figure 1. Types of information contained on the SPM CD-ROM.

Copies of the CD-ROM were distributed to interested members of the public, WIPP participants, and the EPA. To make the system as user-friendly as possible, the database uses a point-and-click-structure and all software needed to run the database is contained on the CD-ROM. In addition, all files are hyperlinked so that a user can explore regulatory requirements—such as human intrusion—from a point of origin in the Federal Regulations through their treatment in technical references. Over fifty online help panels provide introductory material on such topics as the WIPP mission, transuranic waste, performance assessment, and a step-by-step explanation of the SPM calculations. The CD-ROM also contains an online software manual and hard copy reports can be generated for each activity.

One of the objectives of the CD-ROM information system is to promote efficiency (and quality) by providing as much relevant data as possible to support users' understanding of the SPM-2 data, results, and calculations. To that end, a graphical interface is provided to orient the user and help navigate through the relational database system. In addition to providing a contextual background for understanding the meaning of data displayed on any panel, four levels of explanatory information are provided.

The Data Visualizer

The heart of the SPM-2 information system is a data visualizer containing performance measures (cost, duration, and probability of demonstrating compliance (PDC)) for roughly 46,700 programmatic options (activity sets). The data visualizer is implemented using a commercially available off-the-shelf database management system for data storage and retrieval, computations, data display, and graphical user interface.

The database management system provides interactive functions for selecting groups of activity sets for analysis and visualization based on performance measure ranges, presence or absence of particular subsets of activities in the activity sets of interest, or both. The visualization system provides six views of selected activity set groups: two-dimensional scatter charts with three-axis pairs (cost-PDC, duration-PDC, cost-duration); a three dimensional view plotting maximal PDC for activity sets whose cost and duration values fall within user defined ranges; and a sorted list display showing activity composition, PDC, cost, and duration for each selected activity set. The system interactively displays details of activities composing each activity set and the complementary cumulative distribution functions (CCDFs) corresponding to the selected activity set outcome. The database component of the system is built of more than 200 display panels, forms, tables, queries, reports, and macros.

In a database, numbers standing alone without annotation or explanation are without meaning to many, if not most, users. Yet most conventional uses of databases provide nothing more than numbers. This often means that the user may have to expend significant effort on finding documentation for the numbers provided in a database in order to adequately understand their meaning. To the extent that effort expended on finding explanations for the meaning of numbers in a database is effort that could be directed to using the numbers productively, the lack of explanation conveniently at hand while using the database represents an element of inefficiency.

Metadata

Within the visualizer display panel, titles, table column headings, table row labels, and graph legends are supported with metadata: information that defines, explains, and relates each of these data entities to WIPP, to the SPM-2 process, and to the scientific and engineering activities whose consideration produced the SPM-2 results. As a minimum, text provides brief definitions and explanations. These brief annotations are linked to a detailed online overview, which is internally hyper-linked to pop-up definitions, to related topics, and to a glossary of terms found in the system. Annotations are linked to the online electronic document library where detailed technical references for the WIPP project are provided.

In addition to the passive metadata, items of particular importance are supported by active-interactive panels. For example, one of the three principal display panels in the visualizer, the "Activity Set Composition Panel," is an interactive table. Its rows correspond to individual activity sets. It has two sets of columns. The first set of columns correspond to technical program areas and the second set correspond to the performance measures: PDC, cost, and duration. Clicking on a column heading activates a pop-up panel that displays a "passive" text definition of the technical program area associated with that column; this definition is hyper-linked to an appropriate destination in the electronic document library. But if the activity set number that labels a row is clicked, an interactive panel is activated. This panel expands the corresponding row of

the "Activity Set Composition Panel," displaying detailed information about every activity in the set. This panel computes and displays the PDC for any combination of outcomes for the set's component activities the user selects and upon the user's request will access the necessary data and generate the CCDF chart for the selected outcome combination. This interactive functionality allows the user to exercise one of the most conceptually elusive computations in the SPM-2 analytical process to gain a concrete grasp of the mechanics of the operation, providing an invaluable complement to the detailed explanation of the computation contained in the overview.

Both passive annotations and interactive panels are an integral part of the database structure along with the SPM-2 results. Every row, every column, every cell has an integrated information component which can be used for the appropriate level and detail of definition, explanation, or demonstration.

The Navigator

The navigator (Figure 2), which acts as the user's portal to the system, provides point-and-click access to all of the other components of the system. The navigator is implemented using an off-the-shelf hyper-text/full text search system. Navigation to predefined points in the textual portions of the system is accomplished by HYPERLINK™. Ad-hoc navigation is supported by a sophisticated full text search system that provides virtually instantaneous shift of focus to locations of topical interest in the text defined by the user. Topical interest can be formulated as words, phrases, and character strings (any of which may include fixed or variable length "wild cards" used to ignore irrelevant variations in textual expression of an idea of interest). These elemental topical recognition cues can be combined with both Boolean (and, or, not) and proximity (e.g., these two words within so many characters or so many words of each other) operators to express complex and precise search specification. The query formulation dialog is designed to be intuitive and easy to use.

Search time is invariant with respect to query complexity and sub-linear by several orders with respect to the magnitude of text searched. This aspect of system performance is of importance quite beyond the dimension of mechanical efficiency. When pursuing an inquiry regarding the meaning of a particular data set, the more rapidly the system can respond, the more questions can be resolved within a unit of time.

The system provides a complete history of the user's trail through the text (whether by hyperjump or by topical search), displaying a list of all locations visited. The user can click on any item in the history and automatically return to that point in the text. A back-track function allows the user to retrace the trail one step at a time. These functions are an invaluable aide to understanding by allowing the user to leave a particular context to follow explanations and related information and then return with little effort to the original context armed with additional relevant information to augment understanding of the original context.

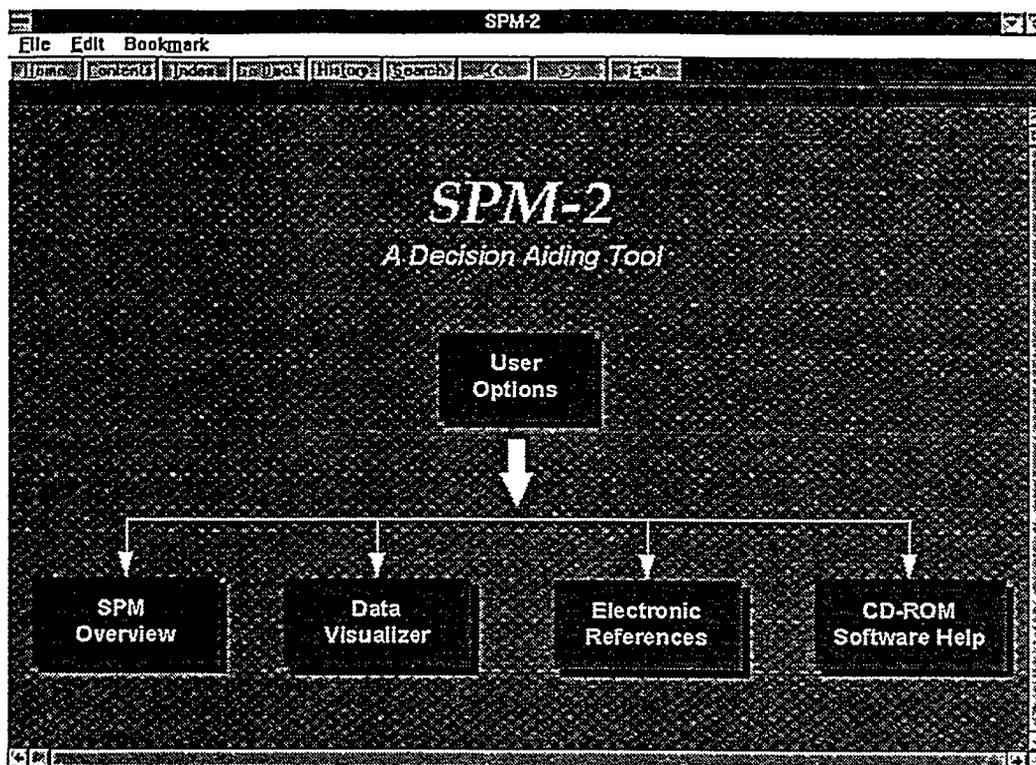


Figure 2. The SPM-2 CD-ROM navigator panel.

Online Context-Sensitive HELP

With over eighty hyper-linked topics, online context-sensitive HELP contains explanations of how to use the system, descriptions of each display control (buttons, check boxes, etc.), and definitions of displayed data items. Brief definitions of SPM-2 terminology are also provided in the software HELP subsystem.

Many of the help panels are graphical and interactive. For example, a multi-level hyper-linked user structure diagram shows the relationships among the user interface panels. Clicking on a block in the structure diagram that represents a particular user interface panel hyperlinks to an image of the panel itself. The panel image, in turn is hyperlinked to explanations of the meaning of data displayed on the panel and instructions on how to use it.

Overview

Definitions of SPM-2 terminology as well as detailed descriptions of WIPP and the SPM-2 project and its objectives, methods, and computational procedures are contained in the *SPM-2 Overview*. This electronic document is internally hyper-linked, and hyper-linked to the database/visualizer system and to the electronic document library described below. The *Overview* is also full text content searchable.

Electronic Document Library

The fourth level of explanatory data, the electronic document library (EDL), is also internally hyper-linked and can be content accessed via full text search. The EDL contains the SPM-2 Final Report, the SPM-1 Final Report, the technical position papers which led

to the computation of the data contained in the database/visualizer, a set of general references, and the full text for 40 CFR 191 and 40 CFR 268.6. The general references, which consist of SPM-2 charter documents and other key instructions establishing the scope and direction of SPM-2, are included as electronic page images (containing signatures, etc.) and as electronic text documents that can be full text searched.

EVOLUTION AND USES OF THE SPM-2 CD-ROM

The magnitude of the data and results suggested using a database to organize the results in a manageable form. A visualizer was included to enhance the value and utility of the data. The estimated size of the database suggested that a high capacity, high density storage medium would be needed. At about the same time as the storage medium became a consideration, DOE direction to make the SPM-2 data and results available to the public and the consequent need for distribution of many copies introduced the requirement for a low cost storage medium. The CD-ROM was the obvious choice. Release to the public also led to development of the navigator, metadata, and all of the other supporting information and tools that resulted in self-contained form of the final system.

Copies of the CD-ROM have been distributed to interested members of the public, WIPP participants, and the EPA. The system allows project members to understand parts of the project on which they had not participated. The supporting explanatory material in the system may be of interest to a wider audience than the results and data and may continue to be of value well after the original purpose of the system has been served.

CONCLUSIONS

The SPM performance-based decision methodology proved to be useful in structuring a large, complex decision and focusing attention on important portions of the problem.

A self-contained information system to support a specific task is made possible in part by the availability of information in electromagnetic form and by contemporary information system tools; like the personal computer, large capacity data storage media, fast processors, large ram memory, low cost database management systems, HYPERTEXT and search engines, etc. The CD-ROM is a key element for distribution of self-contained information systems. The SPM-2 self-contained information system could not have been economically feasible as recently as three years ago. Just since the SPM-2 CD-ROM was released, there have been substantial improvements in all of the software and hardware elements used in its creation. Later this year CD-ROMs with a capacity of 6 gigabytes will be commercially available. A technology for manufacture of 100 gigabyte CD-ROMs is nearing maturity and is expected to be brought into manufacture within the next year or two. Multispectral approaches to recording and reading optical data are expected to yield CD-ROM systems with even higher capacities within the next five to ten years. The CD-ROM's cost performance is superior to any paper form or even any microform. When large numbers of copies are required, the per disk cost of the CD-ROM is small. The economies of distribution via CD-ROM are so compelling that there should be a conversion of many publication and record keeping requirements to this medium.

REFERENCES

1. N.H. PRINDLE, F.T. MENDENHALL, D.M. BOAK, W. BEYELER, D. RUDEEN, R.C. LINCOLN, K. TRAUTH, D.R. ANDERSON, M. MARIETTA, and J. HELTON, "The Second Iteration of the Systems Prioritization Method: A Systems Prioritization and Decision-Aiding Tool for the Waste Isolation Pilot Plant. Volume 1: Synopsis of Method and Results," SAND95-2017/1, Albuquerque, NM: Sandia National Laboratories (1996a).
2. N.H. PRINDLE, F.T. MENDENHALL, W. BEYELER, K. TRAUTH, S. HORA, D. RUDEEN, and D. BOAK, "The Second Iteration of the Systems Prioritization Method: A Systems Prioritization and Decision-Aiding Tool for the Waste Isolation Pilot Plant. Volume II: Summary of Technical Input and Model Implementation," SAND95-2017/2, Albuquerque, NM: Sandia National Laboratories (1996b).
3. N.H. PRINDLE, D. M. BOAK, R. F. WEINER, W. BEYELER, S. HORA, M. G. MARIETTA, J. C. HELTON, D. RUDEEN, H. JOW, and M. TIERNEY, "The Second Iteration of the Systems Prioritization Method: A Systems Prioritization and Decision-Aiding Tool for the Waste Isolation Pilot Plant. Volume III: Analysis for Final Programmatic Recommendations," SAND95-2017/3, Albuquerque, NM: Sandia National Laboratories (1996c).
4. WIPP PA (Performance Assessment) Department, "Preliminary Performance Assessment for the Waste Isolation Pilot Plant, December 1992. Volume 4: Uncertainty and Sensitivity Analyses for 40 CFR 191, Subpart B," SAND92-0700/4, Albuquerque, NM: Sandia National Laboratories (1993).
5. D.M. BOAK, N. H. PRINDLE, R. A. BILLS, S. HORA, R. LINCOLN, F. MENDENHALL, and R. WEINER, "Summary of the Systems Prioritization Method (SPM): as a Decision-Aiding Tool for the Waste Isolation Pilot Plant," Waste Management '96, Tucson, AZ, February 25-29, 1996, SAND95-1998C, Albuquerque, NM: Sandia National Laboratories (1996d).

This work was supported by the United States Department of Energy under Contract DE-AC04-94AL85000.

DISCLAIMER

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