Strategic Highway Research Program–Long-Term Pavement Performance Information Management System

WILLIAM O. HADLEY, CHARLIE COPELAND, AND SHAHED ROWSHAN

A major impediment to past pavement research efforts was the lack of availability of and access to comprehensive diverse, yet consistent, traffic, materials, structural, and climatic data for various pavement types. One of the basic objectives or goals of the Strategic Highway Research Program (SHRP) was to establish a national pavement performance data base in which to store all of the data being collected or generated, or both, under the Long-Term Pavement Performance (LTPP) Program. Under this program researchers could pursue long-term pavement performance studies by accessing all sorts of data with confidence. The quality assurance approach used in the SHRP-LTPP Information Management System to verify the accuracy and correctness of the data it receives and stores before releasing the data to the public is presented. In addition, the levels of data release are presented and the information sources of benefit to the user are identified.

A major impediment to past pavement research efforts was the lack of availability of and access to comprehensive diverse, yet consistent, traffic, materials, structural, and climatic data for various pavement types. There is no doubt that data sets containing variable and inconsistent data make it extremely risky to develop inferential conclusions. Because of the historic problems with data comprehensiveness, quality, and consistency, it is of strategic importance to develop a national data base that can overcome these shortcomings and allow researchers to pursue long-term pavement performance (LTPP) studies by accessing all sorts of data sources with confidence.

LTPP AND INFORMATION MANAGEMENT SYSTEMS

One of the basic objectives or goals of the Strategic Highway Research Program (SHRP) was to establish a national pavement performance data base (NPPDB) in which to store all of the data being collected or generated, or both, under the LTPP Program (*1*). The type of data collected in the LTPP Program and stored in NPPDB include the following:

- Inventory (as built)
- Materials testing
- Profile
- Deflection [falling weight deflectometer (FWD)]
- Cross profile
- Distress
- Friction
- Maintenance
- Rehabilitation
- Climate
- Traffic

The Information Management System (IMS) developed in the SHRP-LTPP Program to service NPPDB is composed of five nodes—the central node and four regional nodes. The National Information Management System (NIMS) is the central node (*2*), which is composed of the hardware and software systems that were assembled to house NPPDB. This system is administered by and resides at TRB. The four regional nodes are represented by the Regional Information Management Systems (RIMS). The data generally are checked and entered at the RIMS by the four regional coordination office contractor personnel under the direction of a SHRP regional engineer. Periodic uploads are made from RIMS to NIMS at TRB.

A critical function of IMS is to verify and validate the accuracy and correctness of the data it receives and stores before releasing the data to the public for review, compilation, analysis and research. The NPPDB data must pass a number of quality assurance (QA) checks before being released to the public from NIMS. These checks verify the presence, reasonableness, and validity of the data. The procedures for data checks, as well as data uploads to NIMS, are critical elements in the SHRP-LTPP IMS. Data uploads include newly acquired data as well as updated or revised data that were previously submitted.

DESCRIPTION OF REGIONS

The four SHRP regions were selected primarily on the basis of climatic and jurisdictional considerations (*2*). The North Atlantic region corresponds to the wet-freeze AASHTO classification, whereas the southern region is situated in primarily a wet-nonfreeze zone. The north central region is predominantly wet-freeze, whereas the western region contains both dry-freeze and dry-nonfreeze. The regions were adjusted to correspond to state boundaries as illustrated in Figure 1 (*3*).

Four regional offices were established to coordinate and communicate SHRP-LTPP related activities across the United States and Canada. Each region includes a group of states or provinces, or both, in their jurisdiction, with test sections located throughout the defined boundaries. The regional centers then operate as central data collection and validation centers for pavement section data. Inventory, maintenance, rehabilitation, and traffic data are

W. O. Hadley, Hadley and Associates, 12402 Mercury Lane, Austin, Tex. 78727. C. Copeland, Texas Research and Development Foundation, 2602 Dellana Lane, Austin, Tex. 78746. S. Rowshan, FHWA, LTPP Division, 6300 Georgetown Pike, McLean, Va. 22101.

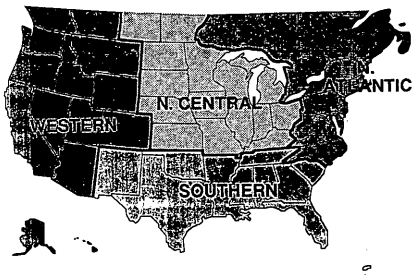


FIGURE 1 SHRP-LTPP regional boundaries.

collected at the state level and are then sent to the appropriate regional center. The regional centers receive these data from the states and collect test and monitoring data on the pavement sections. All data collected are entered in the RIMS through a menu-driven system or are loaded by programs reading data from machine-readable media. Quality checks are incorporated into all update programs, and reports are designed to provide additional checks. After verification, these data are transferred to NIMS.

The IMS functions performed at the regions involve primarily data collection, data validation, and data entry. This is an essential IMS element because the regional staff have a working relationship with all of the data providers and the technical expertise to judge data quality.

INFORMATION SOURCES

Several sources of information are available to describe in detail the data housed within the LTPP data base and how it was collected for the IMS. The SHRP-LTPP data collection guide (DCG) (*4*) is the main source of data collection sheets and instructions on data collection for the LTPP Program. Detailed DCGs for data gathered by SHRP contractors have been developed for the materials sampling and testing and some of the activities of the monitoring program. A schema report from IMS provides the data structure as it is implemented in the relational data base management system (ORACLE) and illustrates the data tables and the fields (or data elements) contained within those tables. The schema also identifies the key (index) fields and the data types associated with each field. The data dictionary report reference from IMS is a more thorough description of each of the fields (or data elements) and various items of interest about each of the fields.

DATA COLLECTION GUIDE

The primary purpose of the DCG (*4*) is to provide a uniform basis for data collection during long-term monitoring of the performance of pavement test sections under the LTPP study. Data items considered to be of high priority for achieving the goals of the LTPP Program are identified, but other data items that are desirable for inclusion in the NPPDB for other purposes are also included. Particular emphasis has been given to the collection of those data items considered essential to long-term pavement performance to ensure that crucial data will be available in NPPDB when it is utilized in the future for the development of pavement performance models.

The DCG was initially developed for use with the General Pavement Study (GPS) sections, but many of the DCG sheets are also used directly with the Specific Pavement Study (SPS) sections. Additional data sheets and tables have been designed and used to record data collected from the SPS project sections.

SCHEMA REPORT

The schema (*2*) is used in NIMS to define the various tables (categories of data) and fields (individual pieces of data) and to identify how they are to be stored in the IMS data base. Each of the data modules is composed of numerous tables (encompassing one or more data sheets) and represents a collection of information about a specific item, for example, the location of all LTPP pavement sections by state, elevation, and coordinates. Each table in turn is a collection of records that contain data about a specific pavement section. Each record is made up of individual fields that represent the smallest piece of information in the data base. The schema defines within the data base the fields that belong together as a record, the records that reside in a certain table, and the tables that compose a specific data module.

DATA DICTIONARY REPORT

The data dictionary is a supplemental report (*2*) that describes for the IMS users the various fields or data elements contained within each table. The data dictionary entry identifies the origin of the data (i.e., what data sheet and time) and presents a brief description of the field (data type), data ranges, and associated information.

The rules associated with the IMS data dictionary determine the amount and type of data that may be input in each field. For example, the data dictionary defines the length of a field, the type of data to be entered (e.g., numeric, alphabetic, date) and the acceptable ranges for the data (e.g., a positive number from 1 to100).

DATA TYPES, ELEMENTS, AND SOURCES

NIMS is the central repository for all LTPP data. All requests for LTPP information or data files from the user community are processed at NIMS, which consists of data uploaded from the four regional centers along with data entered directly at the national center. The data processed directly at NIMS includes the environmental data and all administrative data (e.g., information for new pavement sections, experiment assignments, and code tables). Each region is responsible only for the data on the SHRP pavement sections located within its assigned states; therefore, there is no overlap between states of the data collected. The procedures for the transfer of information are described in the SHRP Programmer's Reference Manual (*5*) and in the LTPP NIMS and RIMS User Manuals (*6,7*).

LTPP IMS

IMS QA Process

The QA concept of data checks is presented graphically in Figure 2. This QA process is necessary to provide researchers with con-

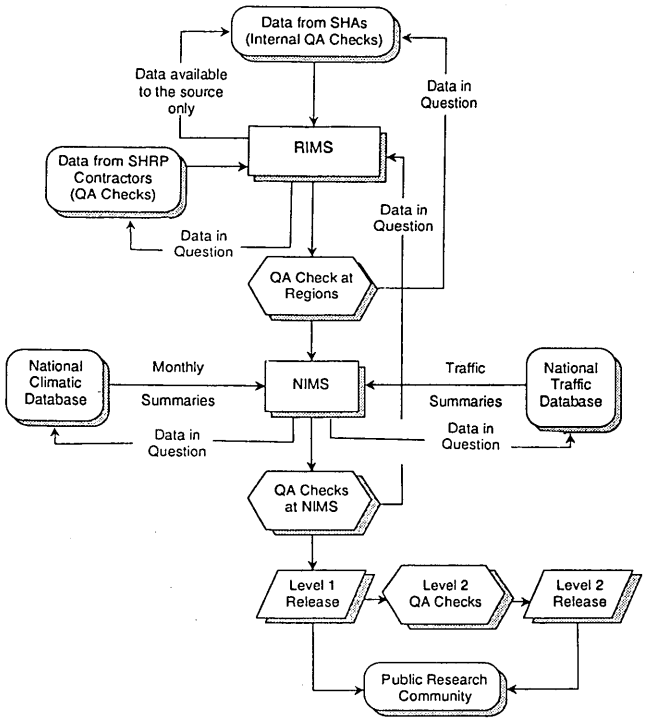


FIGURE 2 Data flow in the LTPP IMS.

fidence that the data are reliable and that their findings and recommendations are based on quality data.

The components of the IMS QA plan are performed in the following sequence:

1. Documentation of data collection procedures for each module in IMS to ensure that data are collected in similar format, types, and conditions, and so on;
2. Regional review of all input at RIMS to identify obvious data collection and data entry errors;
3. Internal checks at NIMS to identify data entry problems or errors; and
4. Execution of the format IMS QA software programs. This component involves nine categories of QA checks defined within two release levels.

IMS Data Entry Checks

Entry checks programming in the IMS include mandatory, logic, range, and data verification checks (*8*). The mandatory checks involve checks for non-null entries in all key fields and other designated fields. RIMS will require entry in these positions or will invoke an audible warning and message that data are required in the field.

Logic checks are also introduced in IMS and are included to ensure data compatibility across tables. The following is an example of a logic check: The "minimum data value $\leq$ mean $\leq$ maximum" for a given parameter.

Range checks are enforced to ensure that numeric field values fall within a defined value. Both absolute (i.e., theoretically possible range) and warning (i.e., practical range) limits are used.

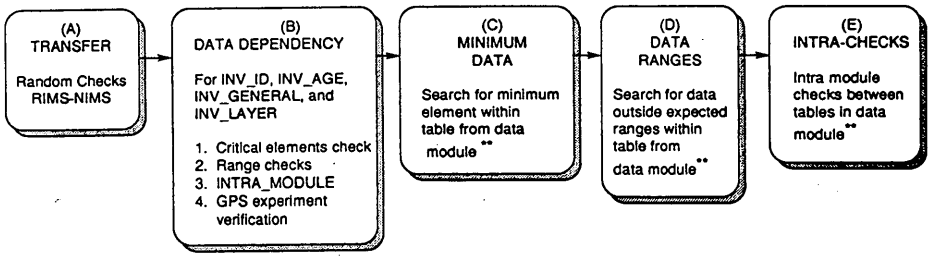
Verification checks are instituted systemwide in IMS to verify that the SHRP-LTPP sections have been authorized for the LTPP Program before any data from that section can be entered in IMS.

Level 1: Section Release

The first release level is a section-by-section release process involving five individual QA/quality control (QC) checks defined as Checks A through E (*3*). It should be noted that the QA/QC checks are conducted on the individual tables within IMS and not on the SHRP sections as a whole. In this process, the Level 1 release could allow some data to be released for a section (e.g., friction results), whereas while other section data that fail to meet the checks would not achieve the release status (e.g., climate). Each table includes a large number of individual data elements.

These Level 1 data checks are structured to ensure quality data within a particular SHRP section but do not address QA/QC requirements between sections, states, and regions. These more sophisticated checks are required at the next release level. The Level 1 release QA/QC checks are presented in Figure 3 and involve the following activities:

- Check A: Random checks to ensure correct RIMS-NIMS upload exchange;
- Check B: Data dependency checks to ensure that basic, essential section information is recorded in NIMS (e.g., location and elevation);



** Except INV_ID, INV_AGE, INV_GENERAL, and INV_LAYER tables in Inventory module

FIGURE 3 IMS Level 1 (section) release quality check.

- Check C: Minimum data search for critical elements (e.g., friction data should include skid number);
- Check D: Expanded range checks to identify data elements that fall outside an expected range; and
- Check E: Intramodular checks to verify the consistency of data within data modules.

The five checks (A through E) in the Level 1 release category are hierarchical in concept and must be conducted in succession, as indicated in Figures 3 and 4. Data elements that do not pass a particular QA check must be reviewed for confirmation or revision by SHRP regional personnel. In this concept, the data dependency checks (Check B) will not be processed until the RIMS-NIMS data check transfer (Check A) has been successfully completed. Similarly, Check E (intramodular) is not initiated until the range checks (Check D) have been successfully completed. After Check E has been conducted and the data within the particular IMS table have passed the check, that IMS table can be released for public use. Once records have passed through Check E, the data are available for a sectional release.

In May 1993 the Level 1 checks were defined and installed within NIMS. Four Level 1 data releases were completed using the checks. In the process the checks were reviewed, expanded, and revised as necessary.

Level 2: Experiment Release

A Level 2 IMS release is classified as an experimental release and includes QA checks across data modules (*9*), confirmation of GPS experiment and cell assignments (*10*), and statistical checks on the data and IMS tables within each designated GPS experiment (*11*). The successful completion of these checks means that the LTPP data would be available for a general experiment-by-experiment evaluation and analysis. The IMS Level 2 release QA/QC checks involve the following activities:

- Check F: Intermodular cross checks to verify existence and consistency of data for related categories;
- Check G: Experiment and cell assignment checks based on collected data;
- Check H: Various checks involving frequency distributions and bimodal and variance checks;
- Check I: Statistical checks for outliers, missing data, and completeness of experiment.

An example of the type of intermodular cross checks (QA/QC Check F) included in the QA program is presented in Figure 5. To assess FWD data at SHRP sites, it would be essential to have information on environment (temperature), materials (layer thicknesses and resilient modulus estimates), and depth to rigid layer. Similarly, an analysis of AASHTO performance (i.e., present serviceability index) would require information on roughness (profile), cracking and patching (distress), rutting, and surface material types. This check is in fact conducted for a specific SHRP-LTPP section but represents the type of checks that are performed across data modules. This check must be completed before being subjected to the experiment and cell assignment checks (or Check G).

The experiment and cell verification check (Check G) is essential for establishing the completeness of each GPS experiment matrix. As shown in Figure 6, the process is conducted for each SHRP section and involves

- Confirmation of the GPS experiment assignment,
- Confirmation of the cell assignment within the GPS experiment matrix, and
- Assessment of experiment completeness.

In essence, this IMS QA check is used to ensure appropriate GPS experiment assignment and to confirm that the distribution of LTPP sections within the experiment matrix is good enough to ensure unbiased data. This check must be successfully completed before Checks H and I are begun.

- The variation in data across regions, as well as within regions, will be analyzed as part of Check H to assess nonuniformity in variance distributions and to check for unusual occurrences or biases that may affect future analyses. Examples of this type of QA check are presented in Figure 7.

The final check before an IMS Level 2 release is shown in Figure 8 and involves statistical checks to identify missing and aberrant data and to confirm outliers. The process will include initial variance analyses at both the regional and national levels and preliminary regression analyses to investigate important factors and variability in materials, construction, or both.

Once the data and IMS tables have passed through Check I, the data are available for an experiment analysis release.

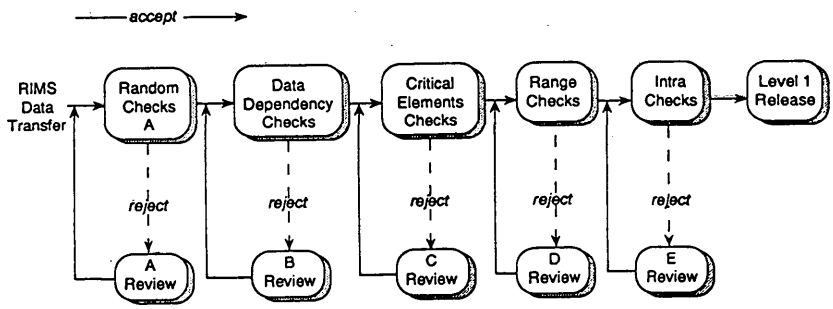


FIGURE 4 Data level advancement with quality control checks.

INTER-MODULAR CROSS CHECKS
Behavior Considerations

NDT - FWD	---	Load deflection
Environment	---	Mean monthly temperature
Materials	---	Layer thickness
Materials	---	Depth to rigid layer
Materials	---	Mr values

INTER-MODULAR CROSS CHECKS
Performance Considerations

Profile	---	IRI
Distress	---	Cracking & patching
Rutting	---	Rut depth
Materials/ Inventory	---	Surface type

FIGURE 5 IMS Level 2: F checks.

EXPERIMENT / CELL VERIFICATION

GPS Experiment Confirmation

Experiment Cell Confirmation

Completeness Of Experiment

GPS Experiment Definitions

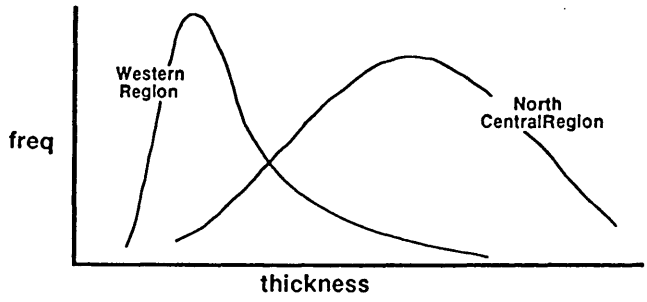
GPS-1	Asphalt Concrete (AC) on Granular Base
GPS-2	AC on Bound Base
GPS-3	Jointed Plain Concrete
GPS-4	Jointed Reinforced Concrete
GPS-5	Continuously Reinforced Concrete
GPS-6A	Existing AC Overlay on AC
GPS-6B	New AC Overlay on AC
GPS-7A	Existing AC Overlay on Portland Cement Concrete (PCC) Pavements
GPS-7B	New AC Overlay on (PCC) Pavements
GPS-9	Unbound PCC Overlays of PCC Pavements

FIGURE 6 IMS Level 2: G checks.

DATA AVAILABILITY

Data generally are made available to the public from NIMS after appropriate QA and QC checks have been concluded. To obtain LTPP data from IMS, requests must be made to the TRB IMS administrator using a completed LTPP IMS data request form (*3*). All data requests are processed at TRB by the IMS administrator.

FREQUENCY - VARIANCE
DATA DISTRIBUTION CHECKS



FREQUENCY - VARIANCE
BIMODAL DISTRIBUTION CHECKS

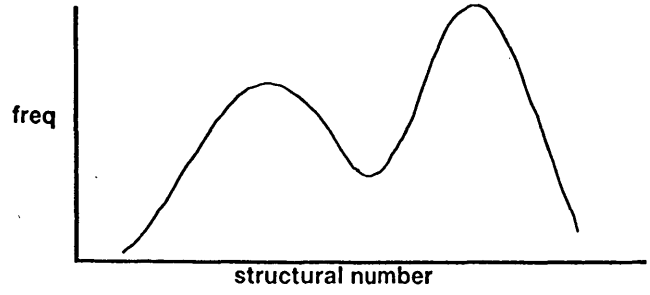


FIGURE 7 IMS Level 2: H checks.

STATISTICAL CHECKS / ANALYSES

Data Investigation
- Missing data
- Aberrant data
- Outliers

Initial Variance Analyses
- Regional
- National

Regression Analyses (Prelim)
- Important factors
- Materials/construction variability

FIGURE 8 IMS Level 2: I checks.

TABLE 1 Examples of Level 1 Releases

TABLE	RELEASE DATES JAN. 91	JULY 91	JAN 92	JULY 92
INV_ID	226	561	660	685
MON_SKID	95	416	560	744
MON_RUT_MASTER	-	2	355	902
REF_LAYER	-	118	296	495
MON_DEFL_MASTER	-	497	515	656
MNT_HIST	-	-	9	20
RHB_IMP	-	-	1	2
MON_PROFILE_MASTER	-	-	2860	3304
COMMENTS	-	913	2386	2800

In return, TRB will provide the requestor with a package containing the data on the requested media, a diskette containing significant portions of the Database Structure Manual (*12*) , and a notice describing major changes to the data base in the previous 6 months. The package will include a detailed LTPP schema and the LTPP data dictionary. The schema identifies the fields in each IMS table along with the columns where these data are available in flat ASCII files. The LTPP data dictionary contains a description of each field including the size, units, and expected ranges and identifies the table name where the field can be found.

The diskette also houses a report that indicates the codes used in the IMS and their associated descriptions. The use of codes provides more quality control by reducing the amount of data entry and storage. For example, comment codes were established for use in recording laboratory test results. Each numeric code corresponds to an individual comment relating to conditions that may affect the results (e.g., color, condition, insufficient size sample). The codes and their corresponding descriptions are provided on the diskette distribution with each completed request.

STATUS OF IMS RELEASES

By November 1992, four public data releases had been made, all at Level 1 involving only GPS data. The releases were completed at 6-month intervals because of the large volume of data inserted in RIMS during this start-up period. This can be seen in the amount of data that were released each time (Table 1). The table names (e.g., INV_ID) were selected to represent the status of the data module with which the table is associated.

The first data release was in January 1991 and was the initial trial of both the data release procedures and the QA/QC checking software. This initial release produced many expected anomalies involving missing inventory data. Because this information would never be available because it either was never collected originally or had been lost or destroyed over the years since the GPS section had been built, a comments table (*12*) was added to the IMS structure so that the regions could document what was missing and allow the data to pass through the QA/QC process without being permanently held at that level (and never being released). The initial release included 226 releasable sections (see INV_ID) and the only other module to successfully pass through Level 1 was friction (skid). For informational purposes, MON_RUT_MASTER is the table that contains the cross-profile data; REF_LAYER represents the materials and testing data module, which includes records for each pavement layer; MON_DEFL_MASTER represents the FWD data; MNT_HISTORY represents the maintenance data; RHB_IMP represents rehabilitation data; and MON_PROFILE_MASTER represents the profilometer data.

ACKNOWLEDGMENTS

The research in this paper was sponsored by SHRP, National Research Council, and appreciation is extended to the cooperative effort of the SHRP personnel.

REFERENCES

1. *Strategic Highway Research Program Research Plans.* Final Report. Strategic Highway Research Program, National Research Council, Washington, D.C., May 1986.
2. *Operations and Quality Assurance Manual.* Long-Term Pavement Performance Informational Management System, April 1990.
3. *Long-Term Pavement Performance Information Management Researchers' Guide.* Strategic Highway Research Program, National Research Council, Washington, D.C., July 1991.
4. *Data Collection Guide for Long-Term Pavement Performance Studies.* Operational Guide SHRP-LTPP-OG-001. Strategic Highway Research Program, National Research Council, Washington, D.C., Jan. 1990.
5. *Programmer's Reference Manual.* Long-Term Pavement Performance Information Management System, July 1989.
6. *NIMS Users Manual.* Long-Term Pavement Performance Information Management System, March 1990.
7. *RIMS Users Manual.* Long-Term Pavement Performance Information Management System, Jan. 1990.

8. *LTPP IMS Data Quality Assurance Checks.* Draft. Long-Term Pavement Performance Information Management System, May 1992.
9. Hadley, W. O. *IMS Level 2: F Checks—Intermodal.* Information Management System (IMS) Message TRDF-134. Texas Research and Development Foundation, Sept. 29, 1992.
10. Hadley, W. O. *IMS Level 2: G Checks—Experiment and Cell Verification.* Information Management Systems (IMS) Message TRDF-88. Texas Research and Development Foundation, Sept. 23, 1991.
11. High, R., and W. O. Hadley, *IMS Level 2: Statistical Quality Control and Assurance Procedures.* Information Management System (IMS) Message TRDF-134a. Texas Research and Development Foundation, Sept. 11, 1992.
12. *Database Structure Reference Manual.* Long-Term Pavement Performance Information Management System, Jan. 1990.
13. Copeland, C. *Comments Table Review and Usage Recommendations for SHRP-LTPP Regions.* Information Management System (IMS) Message TRDF-56. Texas Research and Development Foundation, May 3, 1991.

The contents of this paper reflect the views of the authors, who are responsible for the findings and the accuracy of the data presented herein. The publication of this paper does not necessarily indicate approval or endorsement by the National Academy of Sciences, by FHWA, or by any state highway or transportation department of the findings, opinions, conclusions, or recommendations either inferred or specifically expressed herein.

Publication of this paper sponsored by Committee on Pavement Monitoring, Evaluation, and Data Storage.