Performance Evaluation of Sensor- and Image-Based Technologies for Automated Pavement Condition Surveys

Renato A. C. CAPURUÇO, MASc, PhD Candidate
University of Waterloo - Department of Civil Engineering
200 University Avenue West
Waterloo, Ontario, Canada N2L 3G1
Phone: +1-519-888-4567 ext. 3869
Fax: +1-519-888-6197
Email: rcapu@uwaterloo.ca

Susan L. TIGHE, PhD, PEng - Canada Research Chair in Pavement and Infrastructure Management
University of Waterloo - Department of Civil Engineering
200 University Avenue West
Waterloo, Ontario, Canada N2L 3G1
Phone: +1-519-888-4567 ext. 3152
Fax: +1-519-888-6197
Email: sltighe@uwaterloo.ca

WORD COUNT: 5024 + 4 TABLES + 2 FIGURES = 6459

ABSTRACT

As pavement condition assesment companies are competing to provide innovative and better automatic analysis and diagnosis software each day, characterzation of industry as a whole is still sparse and difficult to make. In fact, because software and handling procedures are proprietary, each different vendor has its own technology to detect, classify, and quantify surface distresses through automated techniques.

In this paper, a preliminary investigation was conducted to determine significant differences or similarities between sensor- and image-based data, as part of a research effort sponsored by the Ministry of Transportation of Ontario, Canada. To achieve this goal, a simple four-step methodology was defined. First, a data management plan for consistent data compilation was put in place to allow efficient data manipulation. Second, a suitable set of similar distresses was selected as response variables of interest to design and conduct statistical experiments. Then, advanced analysis of variance was carried out to allow statistical data comparisons among companies and both automated technologies. Finally, discussions of the results and reccomendations were drawn.

Overall, this work shows no significant differences among contractors' measurements using sensor-based equipment. On the other hand, there are significant differences among measurements taken using digital image-based technology. The implications of such outcome are discussed in details and specifics about the implementation and development of the methodology are also provided in order to encourage practioners to benefit from the suggested preliminary evaluation approach. In a broader perspective, this paper provides an opportunity for road agencies to revisit selection decisions regarding the acceptance of pavement data collected by a wide range of contractors.

INTRODUCTION

State and local highway agencies are faced with the problem of collecting more data, with improved quality, in a safer manner, and at a lower cost. Recent advances in computer technology, digital pavement imaging, and digital imaging processing provide greatly enhanced methods to collect and interpret the required information. Nowadays, highway agencies are employing a wide variety of automated practices to collect, store, analyse, and dissiminate information on pavement condition.

Over the time, a number of pavement related information has been gathered, such as surface distress, roughness, structural capacity, and friction. As mentioned in *(1)*, there are basically two types of technologies to conduct such gathering; sensor- and image-based technologies. Sensor-based technology collects road profile data by instrumented vehicles using accelerometers and at least one type of three types of non-contact sensors; lasers, acoustic or infrared. Then, the measured profile is converted into a useful index of ride quality using mathematically standardized procedures. Some roughness-related indexes are IRI, HRI, and FRI, as most are described in *(2)*, *(3)* and *(4)*, respectively.

Image-based technology involves the identification of pavement surface distresses (e.g., cracking, patching, etc.) by the means of photographing, videotaping, and digital image capturing. The three methdologies consist of capturing pavement images on high-resolution photo logging, videotapes and digital cameras devices, and reduction of distress data through review of the images in a workstation. Reduction of distress data from image-based practices using workstations involves some sort of manual, semi-automatic, and automatic (or, a combination of all) reviews of the images to classify and quantify distress.

Nevertheless, sensitivity and calibration may affect equipment reliability, precision, accuracy of data, and overall long-term performance of devices, as mentioned in *(5)*. Therefore, mandatory certification of high-speed equipment is perceived as necessary for ensuring equipment accuracy. This certification requires the development of procedures for testing and evaluation through comparisons to reference measurements or testing with simulated (knows) inputs. Data from verification sections, when used, must be processed in a manner that directly evaluates the ability of a device to measure specific pavement distress or performance data.

As more and more post-processing techniques at later stages of pavement data collection are becoming the preferred method of pavement analysis, there are many other factors explaining the expected variability in the measured road data. For instance, in terms of roughness, this inconsistency can be influenced by operation conditions, driver skill, and interval in which data is collected, but not by the processing algorithm itself, which is standard *(6)*. So, given an equal road profile to different companies, the index computed from this necessarily will be the same for all.

On the other hand, in terms of imaging processing, lighting, moisture and shadowing conditions can reduce the ability to analyse variations related to pavement features. In special proprietary algorithms used to classify and quantify the distress might play a role in such identification's capability. Although various automated data collection technologies in current use by highway agencies are inducing a number of research projects and studies, none has been identified that could provide a simple and practical way to address the issue of how different companies are measuring distresses. As a matter of fact, there seems to be no clear delineation between what is the responsibility of the data collector (agency or vendor) and what is the responsibility of the buyer or user of the data collected, in terms of assessing acceptability of the data provided.

To shed more light into the black-box image and practical performance issues of such sensor- and image-based technologies, a four step methodology is outlined in the flow chart shown in FIGURE 1 and explained in greater detailed in the sections that follow. The novelty of this methodology relies in the use of an innovative data management plan and improved statistical experiments to provide easy, practical and accurate comparative analysis.

SCOPE AND OBJECTIVE

For many years, several road authorities and researchers have described in detail the major pavement distresses and the status of distress analysis technologies. In Canada, the Ministry of Transportation of Ontario has setup a research project to determine the applicability of high-speed pavement distress data collection as a replacement for the existing manually based method of visual condition surveys, which initially feed the pavement management systems.

In this paper, automated pavement condition data from three different companies (coded as companies A, B and C), captured in thirty-seven sections across the road network of Ontario, was used to conduct an advanced statistical experiment based on a complete randomized block design and paired comparision to show whether or not significant differences exist between companies. In particular, the issue whether or not the performance of sensor-based technology is significant similar or different, as compared to image-based technology's outcomes employed by each company, was examined.

Following a four-step methodology, data from the companies were first merged together using a innovative data warehouse-based approach, to enforce data conformity and allow easy and fast data retrieval. Then, a suitable subset of sensor- and image-based data was selected from the list of available condition data to allow equal comparisions within each group. Third, using repeated measurement-based designs, hypothesis-testing and confidence interval estimation procedures were performed to provide guidelines as to the reliability and validity of the results. Then, the results are discussed and reccomendations made.

DATA MANAGEMENT PLAN

A data management plan was deemed necessary due to the diversity of pavement data received from the participant companies. Thirty-seven sections of different pavement types (i.e., 16 AC, 9 PCC, 6 COM, and 6 ST) in Southern Ontario have been identified for automated inspection. Within each section, a length of pavement was identified for evaluation and marked accordingly in the field. Three repeat measurements were taken in most sections, and results for all sections were reported at 100 meter interval, using a metric system. Each company provided their own dataset. For instance, company A reported a total of 25 different types of evaluations (5 sensor-based and 20 image-based), company B collected a total of 22 types (3 sensor-based and 19 image-based), and company C collected a total of 39 types (4 sensor-based and 35 image-based).

Based on the shear volume of data received, the total number of evaluations (eq. 1) easily surpassed the mark of *one hundred thousand* unique pieces of information. Besides, due to the fact that each company provided data on their own proprietary nomenclature and file format, data inconsistency was high and productive and expediting data manipulation was made impractical.

$$N = \sum_{i=1}^{37} \sum_{j=1}^{g} \sum_{k=1}^{d} S_i \times G_j \times D_k \quad \text{, where} \qquad (1)$$

N = *Total number of evaluations*
S_i = *Section "i" to evaluate*
G_j = *Segment "j" in a given section "i"*
D_k = *Distress "k" in a given segment "j", in a given section "i"*
g = *Number of segments in a given section*
d = *Number of distresses of interest in a given segment*

This implies that efficient manipulation and easy access to this data and information required to make the analysis can not be carried out using a standard transactional database formats. Originally, it was thought this could be used. Another approach was needed and a data warehouse technology was employed to overcome this problem.

Data Warehousing Technology

OLAP, or Online Analytical Processing is a technology that was termed as such by Codd *(7)*, who invented the relational database model. OLAP was used originally as a buzzword to differentiate it from OLTP (On-Line Transaction Processing). The "T" was replaced by "A" to emphasize the Analytical capabilities of the new technology as opposed to the transactional capabilities of the relational database technology. OLTP databases are designed to maintain atomicity, consistency and integrity (e.g., the "ACID" tests) and are ideal for transaction enviroments. Since OLAP databases (i.e., data warehouse) are not as often updated, these constraints are relaxed, they have a better time referenced, subject-oriented, non-volatile (read only) and integrated approach. These features make it ideal for analytical environment.

OLAP is also a technology that provides a single store of data held within the OLAP ‘Cube’, where data and information can be easier accessed and manipulated by many users regardless of their query expertise. As the dimensionality and hierarchies map the fundamentals of the business (i.e., pavement condition evaluation), analyzing data is an intuitive process.

Data in the OLAP Cube is stored in an efficient manner specifically tailored to analysis. It is therefore possible to analyze data within reports on the fly, performing the so-called slice-and-dice operations with ‘Drilling Down’ or ‘Up’, pivoting, page-by and sorting capabilities to the underlying data which makes up the reported figure. For instance, the user is able to easily and quickly answer questions like

> *“Show me the length of longitudinal cracking, the number of potholes, the area of map cracking and roughness collected by company A for the first 500 hundred meters of Section 2 as compared to company B and C, within the 1st run”*

by simply dragging-and-dropping fields of data in a friendly analysis environment as opposed to write structured queries using programming languages, such as T-SQL and PL/SQL.

Desiging a data warehouse solution, which employs OLAP technology, involves three basic steps: Defining the logical structure of the database itself, developing routines and algorithms for acquisition processes of extracting, transforming and loading (ETL) data from assorted source systems into a single database, and developing a front-end application for easy data access and manipulation.

For brevity (and because this paper isn't about an in-depth description of the software development process) none of the system’s development processes are reproduced here in detail. However, a summary of the data warehousing components used in this research project is given in TABLE 1. A so-called star-schema was used to represent the dimension tables and the levels of data hierarchies. At the end, the developed OLAP-based data warehouse system, within the scope of the data management plan shown in FIGURE 2, allowed efficient and easier data manipulation, including exportation capabilities to more powerful statistical applications. This makes analysis more reliable and it is designed to avoid user error in operations involving data manipulation.

SELECTION OF RESPONSE VARIABLES

Two categories of data had to be selected, so statistical analysis could be performed. For the first category, sensor-based data, the ideal candidate variables were the ones related to roughness due to their great importance for road agencies worldwide. For the image-based category, 15 individual surface distresses are identified by severity and density levels to reflect the overall pavement condition for pavements in the province of Ontario, Canada. In fact, through a summary index called Distress Manisfestation Index, these distresses are weigheted based on the pavement type, and then used in the regional pavement mangement system. Thus, it was desirable to extract data from this universe of distresses only.

With the data ready for efficient data manipulation, the assembled OLAP database was examined through a fully customized front-end application featuring powerful data browser’s capabilities to determine its general data properties and to quickly and correctely identify possible data anomilies (i.e., outliers, missing or errouneous data). The data was “cleaned” and then sorted to allow for the characterization of the responses of interest.

After meticulous inspection of all received data, only a portion of it was found to be suitable for analysis due to higher degree of uniformity as compared to the rest. As such, two types of sensor-based evaluations (i.e, IRI for the right and left wheelpaths), plus three types of image-based (i.e., total length of low/medium/high severities for longitudinal and transverse cracking, and total number of potholes) were chosen among the numerous types of received evaluations to be the response variables of interest for the experiments to be conducted. This is explained next.

THE DESIGN OF EXPERIMENT

The choice of an experimental design depends on the objectives of the experiment and the number of factors to be investigated. In a broader sense, depending on the goals to achieve, four main categories of experiments maybe of interest: *Comparative* designs provides the ability to choose between alternatives; *Screening* designs helps to identify which factors/effects are important; *Response Surface* modeling is used to optimize a variable response or hit a target value; and *Regression* modeling is estimates the dependence of response variables.

The process of planning the experiment so appropriate data that can be analyzed by statistical methods will be collected, resulting in valid and objective conclusions follows three basic principles: *Replication* means the repetition of a basic experiment allowing the experimenter to obtain an estimate of the experimental error; *Randomization* assumes that both the allocation of the experimental material and the order in which the individual runs or trials of the experiment are to be performed are randomly determined; and *Blocking*, which is a design technique used to improve the precision of comparisons among factors of interest by reducing or eliminating the variability transmitted from nuisance factors, that is, factors that may influence the experimental response but may not be of direct interest.

A quick search on the literature *(8)* shows several types of experiments for each of the categories and principles mentioned. In this investigation, however, the randomized block design (RBD) approach was chosen to provide the best outcome, followed by paired comparision testing to further delve into the achieved results. The RBD design strategy improves the accuracy of the comparisons among companies by eliminating the variability among sections. These experiments are illustraded and explained in detail next.

Complete Randomized Block Design

For complete RBD, there is one factor or variable that is of primary interest. However, there are also several other nuisance factors. Nuisance factors are those that may affect the measured result, but are not of primary interest. When the nuiasnce source of variability is known and controllable, a design technique called blocking can be used to systematically eliminate its effect on the statistical comparisions among treatments. Blocking is an extremely important design technique, and its use in the pavement industry is not common.

It is anticipated that each one of the 37 sections will perform differently from each other. For that reason, it is desirable to make the experimental error as small as possible; that is, it is necessary to remove the variability between the 37 road sections from the experimental error. Thus, the previous five response variables chosen in the step two will be independently blocked by section, and a complete design will be achieved because each section contains all the treatments (companies). Within each block (section), it's assumed that the order in which each company conducts its 5 evaluation is randomly determined.

Hypothesis Testing for RBD

In a experiment involving the RCBD, testing the equality of the treatment means is equivalent of conducting testing in terms of treatment effects. Therefore, the hypotheses of interest can be written as follows:

$$H_o\text{: } \tau_1 = \tau_2 = \ldots = \tau_a = 0$$

$$H_1\text{: } \tau_i \neq 0 \text{ for at least one } i$$

In this investigation, the hypotheses testing procedures were summarized in a analysis of variance (ANOVA) table for each one of the five chosen types of data, that is, roughness for left and right wheelpaths, potholes, longitudinal and transverse cracking.

Additional testing was further conducted for each one of the degrees of severities, (i.e., low, medium and high) for both types of cracking to closer identify the main source of variablility. All analyses used a confidence interval of 95%. An example of the analysis of variance for high severity longitudinal cracking is shown in TABLE 2 (a). Using $\alpha = 0.05$, the critical value is $F_{0.05, 2, 72} = 3.124$. Because 3.168 > 3.124, it is concluded the type of company affects the high severity longitudinal cracking mean reading. Also, the sections (blocks) don't seem to differ significantly, because the mean square for the blocks is not large relative to error.

It is interesting to observe the results that would have been obtained if a RBD was not utilized. Supppose all 37 sections were randomly assigned to companies A, B and C, and (by chance) the same design resulted as the data stored in the data warehousing. The incorrect analysis of these data as a completely randomized single-factor design is shown in TABLE 2 (b). Because ($F_o = 3.054$) < ($F_{0.05, 2, 108} = 3.080$), the hypotheses of equal mean high severity longitudinal cracking measurements from the three companies cannot be rejected. Thus, the randomized block design reduces the amount of noise in the data sufficiently for differences among the three companies to be detected.

A summary of each one of the sensor-based and image-based ANOVA tables is presented in TABLE 3, where the critical value of $F_{0.05, 2, 72} = 3.124$ was compared to each one of the $F_{observed}$. As seen, the review of the results of each testing is given in in terms of Yes (reject H_o) and No (fail to reject H_o).

Paired Comparison Design

A special case of the general RBD called paired comparison design (PCD) was conducted to determine the relative performance of matching companies, for all selected set of pavement data. In this case, two observations of the same type of data (one for each pair of companies) should be more similar to one another than to the rest of the data.

The paired test statistic is the difference between the paired observations, which is symbolized by *d* (for difference). The average difference has the same value as the difference between the means of the two samples. The mean of the differences is the same as the difference between the means ($\mu_d = \mu_1 - \mu_2$). One can also calculate the standard deviation of *d* in the normal way using *d* as the observations. The sample size (*n*) is simply the number of paired observations.

There are two reasons for using a paired design rather than a regular comparison of sample means: reduction of bias and/or increased precision. Both reasons may be true at once. The difference in the outcome usually lies in the standard error of *d* being smaller than the standard error of the difference between the means (although the degrees of freedom is usually greater in the second case)

Hypothesis Testing for PCD

In a experiment involving the PCD, it is possible to compare the performance of a pair of companies using a *t-distribution*, for each type of pavement data used in the previous experiment. Thus, the hypotheses of interest can be written as follows:

$$H_o\text{: } \mu_d = 0$$

$$H_1\text{: } \mu_d \neq 0$$

The test statistics for this hypothesis are summarized in TABLE 4. It shows that for a confidence level of $\alpha = 0.05$ and 36 degrees of freedom (37 road sections minus 1), each computed (observed) *t-test* statistic value of a matching pair of companies A, B and C was compared to ($t_{0.025, 36} = 2.028$). Then, decisions were made either to reject (Yes - there is evidence to indicate that two companies produce different readings) or fail to reject (No – there is no evidence to indicate that two companies produce different readings) the null hypothesis.

DISCUSSION

The results of the blocking design, as summarized in TABLE 3, shows that while there seems no gaps among companies in roughness (IRI) testing there is significant ones for both crackings and pothole measurements. For the image-based measurements, significant differences were found for all types of distress and severities (greater for transverse cracking as compared to longitudinal cracking), except for two instances of severites in the pothole type of distress. Further investigation on these two instances revealed that for the "low" severity occurrence, 98 out of 111 measurements were zeros. This might have also contaminated the results when "all" severities were considered together, which explains the different results from the rest of image-based technology data.

A reasonable explanation for the discrepancy of the results for each technique in the blocking design would be the fact that automated roughness testing (sensor-based) is based on standards procedures and practices *(6)*, used by all companies. In fact, as being a basic measure of the pavement surface from a ride quality point-of-view, derived from measurements of the true road profile, converting profile features into a useful index of ride quality has been done using standard algorithms. Using a particular algorithm for all calculations reduces varialbility, as differently observed in surface distresses testing, where each company uses some sort of non-standardized resources and procedures to detect, classify, and quantify them.

The results of the paired comparison experiment, as summarized in TABLE 4, shows similar results. A company-to-company comparision showed no significant difference for automated roughness testing, except in one instance. That maybe caused by bad equipment calibration. Again in this case, each company is using the same type of equipement and algorithms to process profile data, culminating in equally significant similar outcomes. However, when considering image-based technology, significant differences were found for almost all matching companies, that is, none of the companies performed equally (greater difference for transverse cracking as compared to longitudinal cracking, as well). This might be because each company relies on distinct software applications and different worker's training and skills for image processing. For the cases where there was no evidence to indicate differences, that was mainly caused by "zero" readings, which interfered with the comparisions.

RECOMENDATIONS

Based on the findings of this study, a more in-depth investigation with all image-based data is to be performed to check for significant diferences or similarities within companies, for each type of pavement data. If significant diferences still persists, that might indicate that some improvements of the current data collecting and analyses procedures for may be worthwhile to further examination. This can come in several ways. For instance, by improving the quality of image capturing (device related problem), providing more training to data analysists (process related problem), or enhancing / proof-checking distress identification-oriented algorithm's routines (software related problem).

Additionally, this might provide an opportunity for agencies to introduce new and/or stronger stardadization practices for image-based technique. For instance, in order to better assess accuracy, repeatability and possible biases within companies and types of pavement image-based analysis using proprietary processing software and in-house procedures, a direct comparison between manual and automated image-based distresses would be necessary to establish tolerance ranges.

SUMMARY AND CONCLUDING REMARKS

In this paper, an investigation was conducted to identify similarilies or differences between sensor- and image-based technologies, in regards to automated pavement condition surveys. Pavement data from the 37 sections from the road network of the province of Ontatio, Canada, provided by three different contractors was (i) compiled using an innovative data warehousing approach to allow efficient data manipulation, (ii) statistically analysed using a complete randomized block design and a paired comparison experiment, and (iii) the results were discussed. At a confidence level of 95%, it was shown that sensor-based data (i.e., Roughness) are statistically the same for all companies, regardless the underlined section. However, image-based data (i.e., total length of low, medium and high severity longitudinal and transverse cracking, and total number of potholes) are statistically different, and no company statistically matched each other's results.

Accordingly, recommendations were made in regards to identifying possible sources for the encountered problems. These might be related to three main sources (i.e., contractor's in-house processes-, devices-, or software-related causes), which must be addressed by road authorities and industry to assure acceptable and reliable quality for automated image-based condition surveys. As shown in this paper, if existing gaps of automated image-based techniques are not correctely reconciled, effectiveness of pavement management systems will be compromised, contributing to unreliable management decisions.

REFERENCES

1. McGhee, Kenneth H., **Automated Pavement Distress Collection Techniques**. The National Cooperative Highway Research Program, NCHRP Project 20-5. Interim Report 34-04. Washington DC, December, 1980

2. Sayers, M.W., **Development, Implementation, and Application of the Reference Quarter-Car Simulation**. , Measuring Road Roughness and Its Effects on User Cost and Comfort, ASTM STP 884, American Society for Testing and Materials, Philadelphia, 1985 pp 25-47

3. Sayers, M.W., **Two Quarter-Car Models for Defining Road Roughness: IRI and HRI**. In Transportation Research Record 1215, TRB, National Research Council, Washington DC, 1989 pp 165-172

4. Capuruço, Renato A. C. et al, **A Full-Car Roughness Index as a Summary Roughness Statistic**. accepted for publication in the Transportation Research Record (TRR), Journal of the Transportation Research Board. National Research Council, Washington DC, 2005

5. Perera R.W. and Kohn, S.D., **Issues in Pavement Smoothness: A Summary Report**. National Cooperative Highway Research Program, Project No 20-51, Transportation Research Board, National Research Council, 2002

6. American Society for Testing and Materials, **Standard Practice for Computing International Roughness Index of Roads from Longitudinal Profile Measurements**, Designation E 1926-98, March 1999

7. E.F. Codd & Associates, **Providing OLAP (On-line Analytical Processing) to User-Analysts: An IT Mandate**, 1993.

8. Montgomery, Douglas C., **Design and Analysis of Experiments**., 5th edition. John Wiley & Sons. 2001

LIST OF TABLES

LIST OF FIGURES

TABLE 1: Data Warehousing Modeling Components

Component	**Description**
FACT-TABLE	A central table in the data warehouse that contains numerical measures and describes a specific event within a business, such as a bank transaction or product sale. • In this research project, the specific event of interest is the “distress evaluation”.
FACT	A row in a table in a data warehouse. A fact contains one or more numeric values that measure a data event, such as a sales transaction. • In this research project, the fact of interest is all given characteristics that allow unambiguously identification of a given evaluation.
MEASURE	A quantitative, numerical column in a fact table. Measures typically represent the values that are analyzed. • In this research project, the numeric value of interest is the “distress measurement”.
DIMENSION	A business entity containing descriptive, textual information, which helps define the hierarchy of the problem in terms that users understand. • In this research project, eleven dimensions were chosen to represent the linguistics of each evaluation. They are as follows: 1-Sections, 2-Pavement Types, 3-Categories, 4-Distresses , 5-Lane Locations, 6-Passes/Runs, 7-Units, 8-Metrics, 9-Companies, 10-Segments, and 11-Severities
DIMENSION TABLE	A table in a data warehouse that describes data in a fact table. • In this research project, fourteen tables were designed to describe data in the given fact table.

TABLE 2: RBD Experiment - ANOVA Tables for High Severity Longitudinal Cracking

Source of Variation	Sum of Squares	Df	Mean Square	Fo
TREATMENTS (COMPANIES)	1451933.032	2	725966.516	3.168
BLOCKS (SECTIONS)	9173177.459	36	254810.485	
ERROR	16497583.674	72	229133.107	
TOTAL	27122694.165	110		

(a) Correct Analysis

Source of Variation	Sum of Squares	Df	Mean Square	Fo
TREATMENTS (COMPANIES)	1451933.032	2	725966.516	3.054
ERROR	25670761.133	108	237692.233	
TOTAL	27122694.165	110		

(b) Incorrect Analysis

TABLE 3: RBD Experiments - Hypotheses Testing Results

Evaluation	Severity	$F_{observed}$	Significant Different ?
IRI	LWP	0.118	No
	RWP	0.061	No
LONGITUDINAL CRACKING	All	6.941	Yes
	Low	17.473	Yes
	Medium	12.854	Yes
	High	3.168	Yes
TRANSVERSE CRACKING	All	35.687	Yes
	Low	55.680	Yes
	Medium	3.576	Yes
	High	10.353	Yes
POTHOLE	All	1.828	No
	Low	0.978	No
	Medium	6.195	Yes
	High	3.787	Yes

Note 1: $\alpha = 0.05$; $Df_{treatment} = 2$; $Df_{error} = 72$; $F_{table} = 3.124$
Note 2: RWP – Right Wheelpath ; LWP – Left Wheelpath

TABLE 4: Summary of Paired Comparison Analysis

Evaluation	Companies	$t_{observed}$	Significant Different ?
IRI (LWP)	A – B	1.173	No
	B – C	0.213	No
	A – C	2.804	Yes
IRI (RWP)	A – B	0.622	No
	B – C	0.203	No
	A – C	1.077	No
LONGITUDINAL CRACKING (ALL)	A – B	0.948	No
	B – C	3.084	Yes
	A – C	3.549	Yes
TRANSVERSE CRACKING (ALL)	A – B	6.344	Yes
	B – C	4.407	Yes
	A – C	6.614	Yes
POTHOLE (ALL)	A – B	3.076	Yes
	B – C	0.225	No
	A – C	1.176	No

Note: α = 0.05 ; $\alpha/2$ = 0.025 $Df_{observations}$ = 36; t_{table} = 2.028

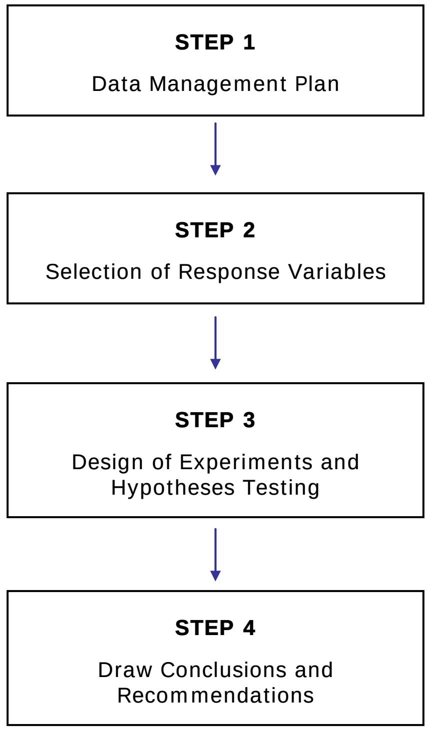

FIGURE 1 Flow Chart for Performance Evaluation of Sensor- and Image-Based Technologies

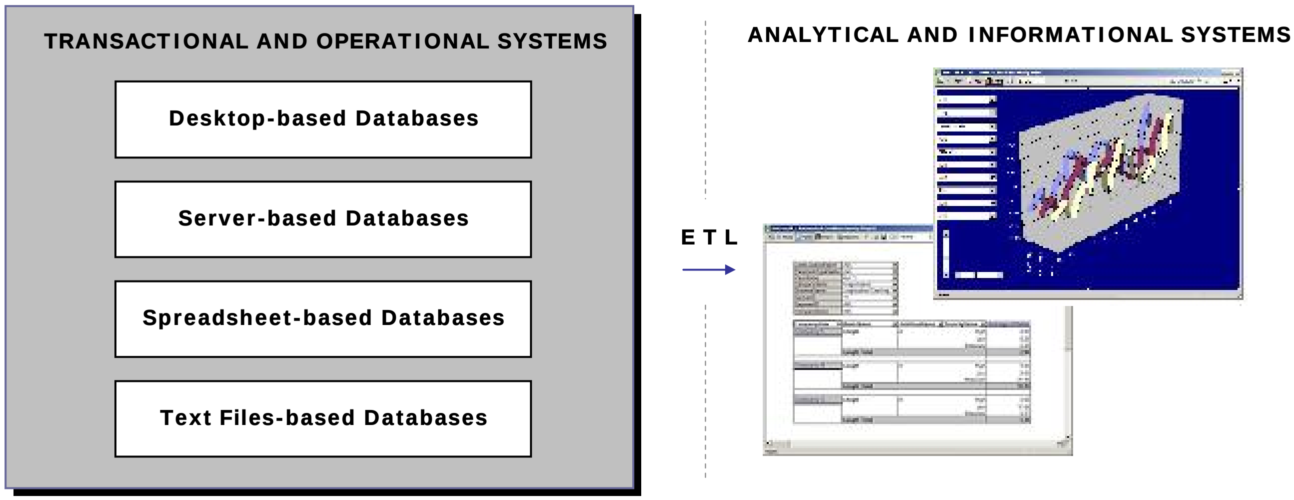

FIGURE 2 Data Management Plan: From Data to Information