Development of Quality Assurance and Control Procedures for Network Level Contract Pavement Surface Condition Surveys

Shawn Landers, Michael Robson and Lynne Cowe Falls

Shawn Landers
British Columbia Ministry of Transportation & Highways
4C-940 Blanshard Street
Victoria, British Columbia, V8W 9T5
Canada
250-387-7631
shawn.landers@gems3.gov.bc.ca

Michael Robson
Stantec Consulting Ltd.
152 Main Street
Cambridge, Ontario, N1R 6R1
Canada
519-622-3005
mrobson@stantec.ca

Lynne Cowe-Falls
University of Calgary
(formerly with Stantec Consulting Ltd.)
2500 University Dr. N.W.
Calgary, Alberta, T2N 1N4
Canada
403-220-5821
cowefalls@ucalgary.ca

Pavement condition data is the most fundamental component of a pavement management system. It is used for monitoring the condition of the paved highways and forms the basis for predicting future performance for network needs analyses and rehabilitation programming. Like many agencies, the British Columbia Ministry of Transportation and Highways (BCMoTH) contracts out the collection of pavement surface condition data to multiple contractors. Because the Ministry is committed to open contracting and the need to collect accurate and repeatable data, comprehensive QA procedures comprised of multiple levels of field-testing have been implemented. Working with one of the data collection contractors and recognizing the importance that quality control (QC) plays in field operations and data processing, expanded QC procedures were developed to improve the quality of the final output for the Ministry. This paper outlines the QA and QC procedures developed and implemented by a highway agency and contractor in a collaborative effort to improve the quality and integrity of network level pavement condition surveys.

INTRODUCTION

Pavement condition data is the most fundamental component of a pavement management system. It is used for monitoring the condition of the paved highways and forms the basis for predicting future performance for network needs analyses, rehabilitation programming and budget preparation. Hence, given the importance of its application, it is paramount that the data be collected accurately and is repeatable from year to year.

While some agencies continue to conduct their own condition surveys, many agencies have moved to contracting this work out. These contractors use multi-function pavement evaluation vehicles that employ automated/semi-automated equipment for measuring surface distress, rutting, roughness, video images, etc. Some of the more commonly used vehicles/systems in North America include the ARAN, ARIA, LVS, PAVUE, PASCO, RT 3000 and VIV (1).

The level of accuracy and repeatability achieved from contract condition surveys, is driven by the Quality Assurance (QA) and Quality Control (QC) procedures followed by the agency and contractor. The QA focuses on testing and verification processes that are carried out by the agency to ensure that the contractor is fully qualified to perform the work that includes continued monitoring as the work progresses. This is complimented by the contractors internal QC procedures that are followed during the field surveys and data processing to ensure data integrity.

This paper outlines the QA and QC procedures developed and implemented by both a highway agency and contractor in a collaborative effort to improve the quality and integrity of network level pavement condition surveys.

AGENCY QUALITY ASSURANCE

Since 1993, BCMoTH has contracted out over 40,000 lane-km of automated network level, pavement surface condition surveys on its main highway network. The surveys include surface distress ratings, rut depth and roughness measurements in both wheel-path and video-logs of the right-of-way.

Because the Ministry is committed to open contracting, QA plays a critical role in ensuring that the data is accurately collected and repeatable from year to year. The Ministry has developed and implemented comprehensive QA procedures that consist of multiple levels of field-testing (2).

BCMoTH has worked closely with its contractors in an open effort to ensure the testing is practical and representative of the intended end use of the data for pavement management. Both of these factors which are inter-related played a key role in the development of the QA procedures. The entire methodology is dependent upon the contractor having real time processing capabilities with on-board computers to not only address and resolve any processing issues arising from the initial calibration process, but also to enable rapid response during the production QA. At any time, during the production

Practicality was important for two reasons, firstly the process must provide a realistic test of the contractor's capabilities and, secondly, the process should not present a huge burden to Ministry personnel to implement and monitor. A data QA program that cannot be effectively implemented provides little value in terms of agency understanding and thereby erodes the level of accuracy and acceptance of the survey results.

Similarly, the scope of the QA procedures was driven by the intended end use of the data. This is an important distinction and can sometimes be overlooked in the effort to collect accurate data. In the Ministry's case, automated pavement surface condition data is collected with the clear understanding that it is to be used primarily for network level, pavement condition analyses. Hence, the degree of data accuracy and field-testing required is dictated by this fact.

The Ministry's quality assurance program is divided into two phases: initial quality assurance where the contractors' methods and equipment are initially calibrated and, production quality assurance where the survey is monitored to ensure continuing compliance.

Initial Quality Assurance

The initial QA tests have two objectives. The first objective is to achieve concordance between the contractor and the Ministry distress nomenclature and to develop agreement on the severity and extent levels described in the BCMoTH Surface Distress Rating System. The second objective is to test field safety procedures and to determine that the instrumentation suite is operating properly. All tests must be passed before authorization is given to begin the production surveys. It involves the following tasks, requiring 2-3 days to complete.

Control Site Selection

Four control sites, located central to the production survey routing are used for the initial QA. The sites are 750 metres in length and include a 250-meter lead-in, complete with markings so that the contractor can better position their vehicle during the test runs. It is preferable for the sites to be located on divided highway sections to facilitate field-testing with minimal traffic disruption and to ensure higher crew safety.

Taken collectively, the control sites provide a representative sample of pavement condition and exhibit the distress types that the contractor will encounter during the surveys. Prior year's survey data, combined with field reconnaissance, are used to select the sites and, in many cases, the same control sites are used for subsequent QA testing if there has been no rehabilitation in the intervening year.

If possible, the sites are located within close proximity to one another to better reflect production-operating conditions. The sites are also located on both on hot-mix asphalt and surface treated pavements with the latter being particularly important given challenge of crack recognition on a different surface texture.

Advance Manual Surveys

Ministry personnel as depicted in Figure 1, conduct manual surface distress, roughness and rut depth surveys at the control sites.

Class I Roughness Survey | Manual Distress Survey | Rut Depth Survey

FIGURE 1 Advance manual surveys at initial QA site.

The manual surface distress surveys are performed according to the Ministry's rating manual (3) including crack mapping for each 50 metro segment. The rut depth manual survey involves taking manual transverse profile measurements in each wheel path at 10 metro intervals using a rut-measuring gauge. A closed loop survey is conducted using the Digital Profilite 300, which is a Class 1 profiler to first establish the true longitudinal profile and the IRI is calculated with its PFLVIEW software program (4).

On-Site Review

The purpose of this exercise is to ensure that the contractor is well versed with the Ministry's surface distress rating methodology regardless of whether a windshield or video-based is being conducted.

Surface Distress Rating Testing

To test the contractor ability to accurately and repeatedly rate pavement distress, using a windshield method according to the Ministry's rating system, the control site is traversed five times with distress ratings collected at 50m intervals. The contractor completes an automated survey over the first site, summarizes the ratings and walks the site with Ministry staff comparing the two rating sets. The ensuing discussions assist in resolving ambiguities.

For video based rating methods, the process is modified, since the rating is performed following data collection. The QA testing for surface distress rating using a SVHS video based system follows the same principles as outlined for the windshield survey, however, unlike the windshield method, whereby the rating is done during the survey, video rating is generally done following the surveys back at the Contractor office. During the initial QA, the contractor is required to video record the four test sites five times and then rate the start-up site, generate a report showing the severity/density ratings for each distress type (as shown previously) and fax the report to the Ministry representative for review. The multiple recordings and ratings are used to test accuracy and repeatability of the process

For both techniques, two criteria are used to assess the contractors surface distress rating ability. The first criteria uses a composite distress score (Pavement Distress Index – PDI) calculated with the PAVER model (5) for the test site. While considered an aggregate evaluation, it provides an acceptable means to compare the relative magnitude of difference between the Contractors surveys and the manual surveys. It also provides a measure with which to assess repeatability as the results are calculated for each of the five passes. The second criteria, uses the severity and density rating totals for each distress type in the test site. Each distress is rated for three levels of severity and five levels of extent, which in detail form is used to assess the rating accuracy and to serve as a diagnostic tool to highlight discrepancies. Each distress severity/extent combination is entered as a keystroke into a spreadsheet and analyzed for three cases: is a particular distress type is being rated too severely, at a different density level or, missed altogether. A spreadsheet program was written to automate the both testing verifications.

The distress ratings are performed in accordance with the BCMoTH Pavement Surface Condition Rating Manual, which was developed in house, based largely on the TAC and SHRP rating methodologies. The manual is used for both project and network level evaluations.

When using or interpreting the distress data, an understanding of how the data is collected is essential. This ensures that it is used in the proper context, recognizing its applications and limitations. The key point that must be remembered is that the objective of the RPMS surveys is to obtain performance data that is sufficiently accurate, representative and consistent for network level analyses. The data is used for monitoring network performance, identifying deficiencies, determining needs, estimating funding requirements, flagging specific problem areas for further investigation and establishing corporate performance measures.

Additional factors that can contribute to the accuracy achieved include:

- the subjectivity factor in visual ratings,
- the speed of the survey vehicle (which usually travels at 70 km/h),
- the angle of the sun,
- time of year (surveys are conducted in the summer when cracks are less problematic and sometimes less visible due to expansion of the asphalt and,
- the potential for human error due to nature of the work.

In terms of subjectivity and the rating methodology, experience has shown that slight severity cracking is very difficult to pick up accurately during an automated survey due to crack width (<5 mm) and in particular at the lower density levels. In practical terms, missing or underrating low severity ratings, was not found to have a significant impact on the PDI comparisons as low severity distress types are not weighted significantly in the PDI model. As well, the Ministry has a fairly aggressive crack sealing program with focus on sealing cracks in the early stages before they become problematic. The angle of the sun is perhaps the most problematic aspect of visually rated surveys and the contractor is required to follow specified travel directions depending upon the time of day. Transverse crack severity rating is another area where problems can sometimes arise with visual ratings with the crack width being perpendicular to the vehicle travel path and this requires specific attention during the early QA testing.

The Ministry is expanding its distress rating QA testing in the summer of 2001 to incorporate a Cohen's weighted kappa statistic analysis, which has been found to be an effective technique to observe the variability between multiple distress datasets. The Ministry has expanded its distress rating QA testing to incorporate a Cohen's weighted kappa statistic analysis, which has been found to be an effective technique to observe the variability between multiple distress datasets. A software program has been developed to automate the kappa analysis to calculate the agreement between the manual and automated raters, and to evaluate the amount of agreement beyond chance. The kappa statistic also allows the introduction of weightings for different distress types, severity and density depending upon agency practices. The kappa statistic analysis will be evaluated in the summer of 2001 during the Ministry's annual pavement surface condition surveys.

Roughness Testing

The roughness testing consists of validating the Contractor's automated surveying equipment by field comparisons to the known longitudinal profile at each test site. The survey vehicle completes a series of five runs over each site in order to assess both accuracy and repeatability. The International Roughness Index (IRI) values for each wheel path are generated and compared to the manual values for each as per the acceptance criteria presented in Table 1.

Rut Depth Testing

Because rut depth measurements are fully automated using a multi-laser rut bar, the rut depth QA tests are designed to validate the Contractor's automated surveying equipment by field comparisons to known transverse profiles. The survey vehicle completes five runs over the site to measure the accuracy and repeatability of the rut bar measurements. The average rut depth value is calculated for each wheel-path and compared to the manual values as per the acceptance criteria presented in Table 1.

Acceptance Criteria

The QA acceptance criteria for the accuracy and repeatability of the surface distress, roughness and rut depth measurements were developed on the basis of Ministry experience and are presented in Table 1.

TABLE 1 Acceptance Criteria for Initial QA Testing

Parameter	Surface Distress	Roughness	Rut Depth
Measure	PDI (0-10scale)	IRI	mm
Survey Interval	50 m	100 m	10 m
Calculation	500 m average	500 m average	500 m average
Unit	lane	Each wheelpath	Each wheelpath
Accuracy	+/- 1 PDI value of manual survey	10% of Class I profile survey	+/- 3 mm of manual survey
Repeatability	+/- 1 std deviation of the PDI for 5 runs	0.1 m/km std deviation for 5 runs	+/- 3 mm std deviation for 5 runs

Should the contractor fail to meet the criteria for acceptance, it is their responsibility to provide remedy until such time that the acceptance criteria is met and the Ministry is satisfied. This may necessitate, but not be limited to additional on-site discussions, repeat of automated surveys, equipment repairs/modifications, retraining and/or replacement of rating staff. Once the contractor has satisfactorily completed the Initial QA testing as described herein, they are authorized to begin the production surveys.

Production Survey Quality Assurance

The accuracy of the contractor's surface condition rating/equipment is also closely monitored during the production surveys. This QA is done using blind sites that are situated along various highways throughout the province where surveys are to be conducted. The blind sites are manually surveyed in advance and of unknown location to the contractor.

Blind Site Locations

The number and general location of the blind sites are based on the contract quantities and the contractors routing schedule. Blind site tests are generally scheduled once every three days during the surveys.

Similar to the initial QA, blind sites are 750 metres in length, including a 250-meter lead-in and exhibit a representative sample of distress types and range in pavement deterioration that the contractor will encounter during the surveys. The sites are designed to not only maintain data accuracy, but are useful in identifying any deterioration or creeping bias in the data collection process. Based on the highways to be surveyed, the sites are also located on hot-mix asphalt and surface treated pavements accordingly.

Advance Manual Surveys

Manual surface distress, roughness and rut depth surveys are conducted at each blind site as described previously, approximately three weeks prior to the production surveys.

A single rater is used for all of the blind sites to provide a consistent benchmark for the manual surface distress surveys. As part of this work, the rater is also required to drive over the site to verify the ratings from a driven perspective and record the surface condition on video to assist later in the QA evaluations.

Monitoring Process

During the production surveys, there are three levels of monitoring performed as part of the QA that includes the use of blind sites and re-testing the initial QA sites.

Each day, the contractor is required to contact and update the Ministry as to their progress. At this time, the contractor is informed that they have passed over a site on the previous day and provided with the site location kilometre chainages. The survey ratings and measurements are immediately submitted by email or fax (See Figure 2) and compared to the manual surveys as per the acceptance criteria. The contractor is instructed to stand down until further notice.

Chainage	LWT*		LJC		PEC		TC		MLC		AC		POT		DST		BLD		IRI		RUT	
(m)	S	D	S	D	S	D	S	D	S	D	S	D	S	D	S	D	S	D	R	L	R	L
94.75	2	3					2	3											2.11	2.23	10	6
94.80	2	3					2	3											2.12	2.25	11	7
94.85	2	3					2	3							2	2			2.13	2.23	9	6
94.90	3	5					2	3							2	2			2.16	2.26	9	5
94.95	3	5					2	3											2.20	2.28	7	5
95.00	3	5					2	3											2.18	2.27	8	6
95.05	3	5					2	3					2	2					2.15	2.25	9	6
95.10	3	5									1	2							2.17	2.23	12	5
95.15	3	5			3	3					1	2							2.11	2.24	10	5
95.20	2	5			3	3													2.17	2.28	9	6

* LWT – Longitudinal Wheel Track; LJC- Longitudinal Joint Crack; PEC- Pavement Edge Crack; TC- Transverse Crack; MLC-Mid Lane Crack; AC-Alligator Crack; POT-Pothole; DST-Distortion; BLD-Bleeding; IRI- International Roughness Index; RUT – Rut Depth

FIGURE 2 Sample survey blind site submission. (S – Severity: D – Density: Numbers refer to value 1-3 for S and 1-5 for D).

As an additional measure, if the production surveys are more than one month in duration, the contractor is required to rerun the four initial QA sites, subject to the acceptance criteria outlined previously (6). The purpose of this latter requirement is to analyze variance between successive runs of the calibration site to determine in there is any creeping bias within each rating crew (for example, are they ceasing to rate moderate longitudinal wheel track cracking consistently).

Acceptance Criteria

The acceptance criteria, in Table 1, apply for the production survey QA as well. Upon receipt, the Ministry promptly reviews the survey data and the contractor is notified as to the results. If the contractor has satisfactorily completed a blind site QA test, they are authorized to continue the production surveys. Conversely, if the test results fail the criteria, the contractor is required to review the video-logs of the blind site in the case of distress data discrepancies and/or make equipment repairs/modifications and if necessary repeat the surveys from the time of last blind site test.

CONTRACTOR QUALITY CONTROL

Quality control is an important issue for data collection contractors. It enhances the integrity of the field surveys and data post processing, thereby reducing turnaround time to the client and ultimately improving the quality of the final product. Contractor quality control operates on three levels that include on-board confirmation/calibration tests, blind site verification and post processing.

QA/QC for the contractor is focussed in two areas: Data integrity and Data continuity. Data integrity relates to making sure that all data fields (regardless of origin) are complete, accurate and delivered to the client in a timely manner. To facilitate the post-processing, on-board verification and validation checks are done automatically to flag incorrect values (for example, has the rater keyed in bleeding and ravelling for the same section?).

Data continuity is concerned with two things: firstly, ensuring that the data is where it should be and is referenced correctly using the client's location referencing system and, secondly that there are no breaks in the data as a result of stop/start driving conditions. All the checks described below were expedited through custom software, which provided flags and logs

Key to the QA/QC process was continuous communication with the client and regular submission of data. This not only smoothed out the workload for the both the contractor and agency, but also significantly increased confidence in the production operation and ultimately provided a better end product.

On-Board Confirmation and Calibration QA/QC

The data control tools that have been developed, practiced and refined over the past 10 years facilitate excellent data management. Figure 3, for example presents the Calibration Screen from the calibration software module that produces IRI from the raw laser data according to any output interval specified.

Other typical on-board QC provisions that Stantec Consulting Ltd. has implemented for contract surveys include:

- Scope Confirmation - Finalize scope of surveys, confirm file formats, agency and in-house training to ensure clear understanding of survey protocol, procedures and communication
- Calibrations - verify event keyboard configuration, rut bar calibration, IRI calibration, distress survey training review and calibration to ensure internal systems accept only agency levels of distress, type, severity and extent, 10-sensor rut bar is within tolerance, roughness data is within tolerance and repeatable and survey crew, project manager, technical advisor and client representatives agree on pattern recognition and distress definitions.
- RT Logs - manual notification of survey issues for review and addressing by RT crew with day to day log of events and production sent to agency.
- Data Logger - on-board software to check invalid keystrokes and data to flag data errors for operator re/view
- Bench Marking - check measured section length versus landmark lengths which allows operator to ensure stop/starts match landmark survey before leaving a section and provides the agency with opportunity to address and react to variances.
- Tolerances - event locations can be compared to known landmark chainages within user defined tolerances so that the agency can set and then validate tolerance for event locations
- Duplicate Video Masters - simultaneous recording of two master videotapes on board vehicle, which generates 2 masters delivered instead of 1 master and copy.

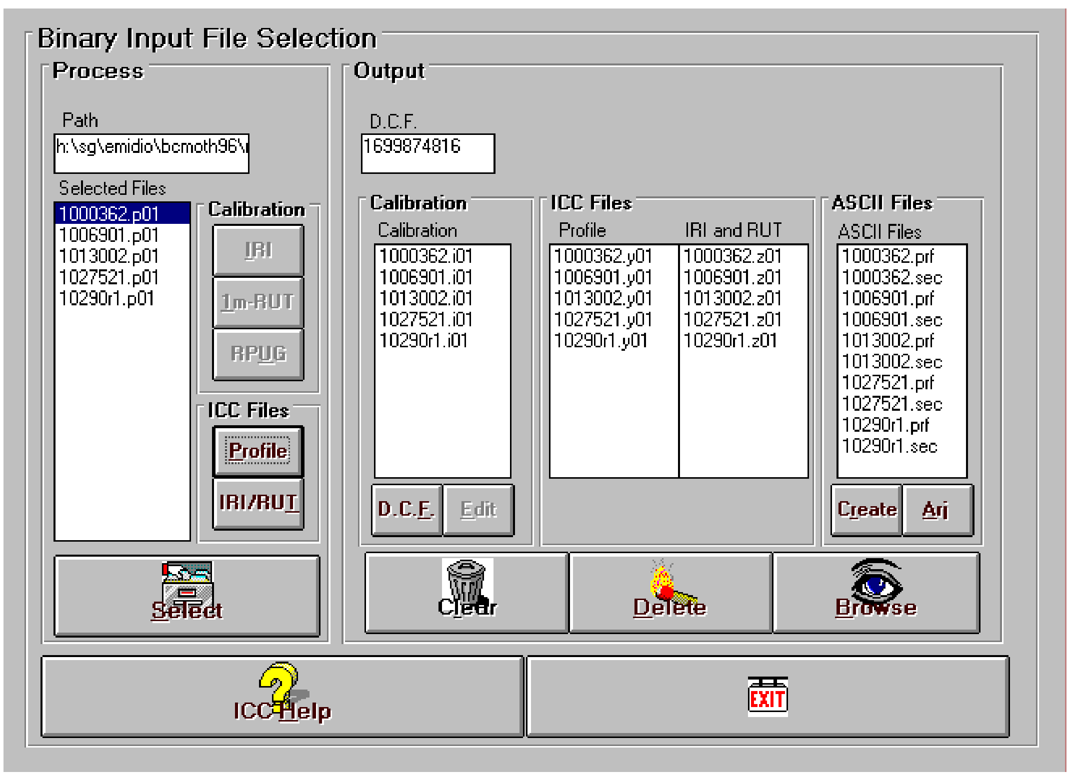

FIGURE 3 On-board data QA screen.

On-Board Edit

At the end of each data collection day, or prior to leaving an established testing zone, the data collected is processed by the operators in the field as follows:

- A manual edit using daily data collection log notes is performed.
- Section ID, Length and stop - start mile points are verified.
- Electronic validation using on-board software is performed to ensure all data fields contain correct numeric data and descriptions.

The ASCII column of the Calibration Screen (at the right hand side of the screen), shown in Figure 3 illustrates the editing capabilities. A comparison of all features, by location and description, is conducted and a log produced that reports any discrepancies between field and section list. The software has the capability to initiate an edit (see icon lower left) which could adjust or reconcile all location points. Conversely a manual edit can be invoked if the user is instructed to adjust selected locations or discrepancies such as the description for any specific feature.

Blind Sites

The data control software enables the operators to immediately produce blind site outputs which is critical given the 'stand down' requirement during verification by B.C. personnel. A module of the processing software is used to convert field data from binary to ASCII format. The "E" file database is accessed through the software and annotated to produce an output equal to the requested blind site.

The Blind Site Reporting Example shown in Figure 4 illustrates this capability. Within the example, a report on a section, say kilometre 0.000 to 0.09 could be produced quickly. These tools for blind site testing and review of results are simple and fast to use in order to facilitate this testing.

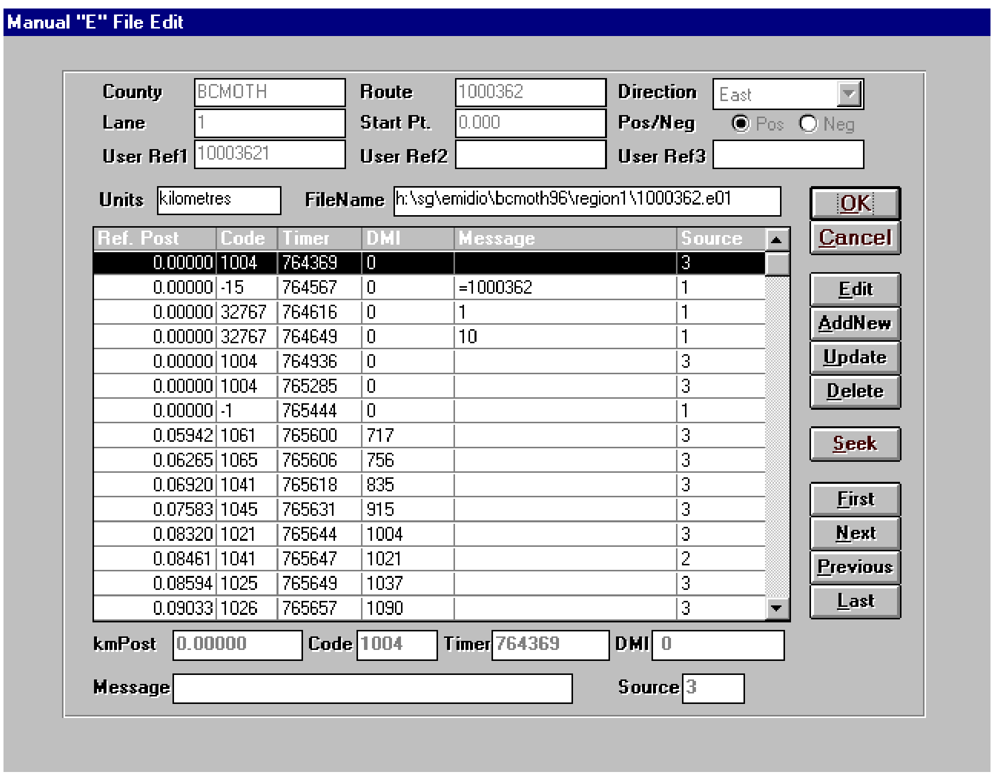

FIGURE 4 Blind site output screen.

Post Processing

The QC provisions for data processing are particularly important as it involves all tasks associated with the conversion, verification and production of the pavement condition data files for uploading into an agency's pavement management system. Stantec consulting Ltd. has developed in-house custom software program that have largely automated which was once a time consuming manual effort. This in turn has reduced data processing time and provides the agency with a its data sooner.

The data post processing QA encompasses conversion of binary files to database files, with full range statistical analyses conducted for data verification, digital conformance checks, hard search and data confirmation and acceptance testing procedures.

CONCLUSIONS

Like many agencies, BCMoTH contracts out the collection of pavement surface condition data. As the need for accurate and repeatable data is paramount, comprehensive QA procedures comprised of multiple levels of field-testing have been implemented to guide the surveys. The need for the testing to be both practical and representative of the intended end use of the data played a major role in defining the QA and is reflected strongly in the end product. To date, the QA survey procedures have served the Ministry well in guiding the collection of its network level, pavement condition data and will continue to evolve.

Quality control is an important issue for data collection contractors as it improves the quality of the final product by enhancing the integrity of the field surveys. Stantec Consulting Ltd. internal QC program has been developed, practiced and refined over the past 10 years in order to facilitate effective data management. The QC provisions encompass on-board confirmation, calibrations, editing, blind sites and post processing.

REFERENCES

1. Transportation Association of Canada, Pavement Design and Management Guide, Transportation Association of Canada, Ottawa, 1997.
2. British Columbia Ministry of Transportation and Highways, “Network Level Automated Pavement Surface Condition Surveys Quality Assurance Procedures”, Construction and Maintenance Branch, Victoria, 1999.
3. British Columbia Ministry of Transportation and Highways, Pavement Surface Condition Rating Manual, Geotechnical and Materials Engineering Branch, Victoria, 1994.
4. CSC Prolilair, “Digital Profilite 300 Operators Manual R1.3”, Vancouver, 1997.
5. Shahin, M.Y. and Kohn, S.D., "Development of a Pavement Condition Rating Procedure for Roads, Streets and Parking Lots", Construction Engineering Research Laboratory, U.S. Army Corps of Engineers, Technical Report M-268, July, 1979.
6. British Columbia Ministry of Transportation and Highways, “1999 RPMS Network Level Automated Pavement Surface Condition Surveys Quality Assurance Report”, Construction and Maintenance Branch, Victoria, 1999.