California’s Experience in Evaluating Automated Pavement Condition Survey Service Providers’ Technical Competence

Pengcheng Fu, PhD (Corresponding author)
Former Postdoctoral Researcher
University of California Pavement Research Center
Department of Civil and Environmental Engineering
University of California
One Shields Ave., Davis, CA, 95616
Tel.: (530) 752-5363   Email: pfu@

John T. Harvey, PhD, PE
Professor
University of California Pavement Research Center
Department of Civil and Environmental Engineering
University of California, Davis
One Shields Avenue, Davis, CA, 95616
Tel: (530) 754-6409   Email: jtharvey@ucdavis.edu

James N. Lee, PhD, PE
Pavement Management Systems Engineer
Pavement Management Program
California Department of Transportation
2389 Gateway Oaks Dr., Suite 200, Sacramento, CA 95833
Tel: (916) 274-6095   Email: james_n_lee@dot.ca.gov

Jeremy D. Lea, PhD, PE
Research Scientist
University of California Pavement Research Center
Department of Civil and Environmental Engineering
University of California, Davis
One Shields Avenue, Davis, CA, 95616
Tel.: (530) 752-4916   Email: jdlea@ucdavis.edu

Jon Lea
Senior Development Engineer
University of California Pavement Research Center
Department of Civil and Environmental Engineering
University of California, Davis
One Shields Avenue, Davis, CA, 95616
Tel.: (530) 752-3561   Email: jlea@ucdavis.edu

Word count: 7280 words (6280 in text, 4 figures). Initial submission on July 31, 2010
Abstract:

Most highway agencies contract their automated pavement condition survey (APCS), and therefore the service providers’ technical competence is critical to the success of the APCS project. However, evaluating APCS service providers presents unique challenges primarily because the technologies are being rapidly advanced by the service providers and the technical details are often outside the skill set of practicing pavement engineers and managers. This paper presents California’s experience gained in evaluating multiple service providers at a recent APCS demonstration conducted by the California Department of Transportation and technically supported by the University of California Pavement Research Center. The participating service providers’ data were assessed on an item-by-item basis instead of the conventional approach of aggregating different distresses and comparing a composite index. Problems and errors were identified for each data item, which will be used to help the winning service provider improve their data quality in the subsequent statewide survey. This paper also documents a number of technological and methodological innovations, including 1) use of several independent methods to improve the credibility of ground truth data, 2) synchronization of profile measurements and pavement images to strengthen the credibility of the determined ground truth, and 3) development of a comprehensive pavement condition information viewer using open standards. Detailed evaluation results of georeference information, rut depths, mean profile depths and roughness are also presented to demonstrate the evaluation procedure and innovations.
INTRODUCTION

Automated pavement condition surveys (APCS), defined as systems that include collection of pavement surface condition information at highway speeds and identification and quantification of distresses/conditions using a combination of software and human judgment, are being used by many highway agencies to replace traditional manual condition surveys where distresses are identified and quantified based on visual identification by humans in the field. The APCS data are used as the primary input into pavement management systems (PMS) for network level management purposes. The advantages of APCS, if implemented optimally, include enhanced crew and public safety, higher productivity, more thorough network coverage, better data objectivity and other intangible benefits compared to its manual counterpart. While a small number of agencies have committed serious levels of resources and time to in-house mastering of APCS technologies (1,2), most agencies choose to contract annual surveys to APCS service providers (e.g. 2,3), thereby outsourcing the cost of ongoing technology development and the risk of equipment and staffing problems, but taking on the problem of assuring that the data received meets quality requirements. An agency’s annual APCS data is often worth millions of dollars, and it is self-evident that the quality of APCS data is crucial to the pavement management systems and in turn to optimal budgeting and project prioritization by the agencies. Therefore, quality management of APCS contracts is an important issue, and numerous studies have been published on this matter (1,3-10).

In terms of quality management of APCS contracts, the attention of most highway agencies has been paid to quality assurance (QA) of data in the production stage (4,6,9), namely to ensure that the data delivered by service providers comply with the specifications and provisions in the contracts. On the other hand, since the service provider’s technical competence is crucial to the quality of data delivered, critical evaluation of the technical competence of potential service providers so that an informed decision can be made regarding the selection of the service provider is an important step towards the success of the agency’s APCS work. A Request for Proposal (RFP) is currently the most used procurement method for APCS contracting (2). A highway agency has to select a service provider from several qualified proposing service providers according to certain technical and financial criteria. Although each state highway agency typically awards hundreds of millions of dollars worth of contracts annually and assessing bidders’ qualifications is almost a routine operation, the evaluation of APCS service providers’ technical competence presents a number of outstanding challenges compared to procurement of more traditional pavement engineering services for the following reasons:

1. APCS collects and integrates a wide variety of data items using many different technologies, including imaging, sensing, positioning, data processing and data management. Many of these technologies are less than three years old, and are constantly and rapidly evolving. The technical knowledge to evaluate these data is outside the skill set of traditional pavement engineers and pavement managers.

2. This kind of technical evaluation usually involves comparison of data collected by service providers on a number of test sections to the “ground truth” established by the agency. APCS service providers are often the inventors and developers of the related technologies, and they often claim to be better equipped and more experienced than their clients. When disputes arise, service providers tend to
question the accuracy of the agency’s ground truth, and it is desirable to have direct and definitive evidence to support the agency’s comparison of the two sets of data.

3. The technical evaluation serves more purposes than merely selecting a service provider. Even the final winner of the APCS contract is not technically perfect. The agency not only wants to know how many mistakes each service provider has made in the evaluation and how severe these errors are, but the agency also needs to understand the nature and sources of these identified errors, so that the agency can work with the selected service providers to solve the problems. Additionally, information collected and lessons learned in the evaluation can greatly help the agency in QA during the production phase.

This paper presents some of the experience gained in the evaluation of APCS service providers for a comprehensive APCS project in California. The selected service provider will perform APCS in the fiscal years (July to June) of 2010-2011 and 2011-2012 for the California Department of Transportation (Caltrans, sometimes referred to as “the agency” in this paper). The evaluation was performed by Caltrans with technical support from the University of California Pavement Research Center (UCPRC). The goal of this paper is to share technological innovations made and lessons learned in this evaluation, mainly to address the three challenges mentioned above. The service providers are kept anonymous for privacy reasons.

APCS IN CALIFORNIA AND THE SERVICE PROVIDER SELECTION PROCESS

Caltrans APCS, Data to Be Collected and Key Innovations

As indicated in the Request for Proposal (RFP number 45a0002) for Caltrans’ annual APCS and the associated documents, the raw data items to be collected by the APCS service provider include:

- Low resolution longitudinal profiles (maximum 25 mm longitudinal sampling interval);
- High resolution longitudinal profiles (1 mm longitudinal sampling interval);
- Transverse profiles (maximum 25 mm transverse sampling interval and maximum 300 mm longitudinal sampling interval);
- Front-view right-of-way (ROW) images; and
- Downward-view pavement surface images.

These raw data will be analyzed in compliance with the interim Caltrans Automated Pavement Condition Survey Manual to yield pavement distress/condition measurements reported for pavement data segments. For the purposes of this paper, pavement data segments can be considered to be the reporting interval for analyzed data, and each data segment is 10 m or one jointed concrete pavement slab long depending on the pavement surface type and data type. A few features of the Caltrans APCS are noteworthy:
- The survey has a 100% network coverage, i.e. all lanes will be surveyed over a two year period with overlap of a significant portion of the network between the two years;

- Distress classification/quantification and non-distress condition measurements are conducted according to the newly developed Caltrans APCS Manual;

- All data items, including raw data and analyzed data carry their own linear reference (and georeference if applicable) information, and location-references of all data items will be synchronized;

- File formats are either existing industry standards, specified by the agency, or proposed by the selected service provider, but all the file formats will be non-proprietary.

APCS Demonstration and Evaluation

Because of the stringent qualification and productivity requirements posted by the agency (Caltrans with technical support from the UCPRC), only two service providers passed the minimum qualification evaluation. As required by the RFP, they participated in an APCS demonstration organized by the agency. Ten short test sections, each 180 m long, covering various structure types, different levels of structural integrity, different levels of traffic volume, and typical distresses seen in California were selected by the agency. Each service provider collected APCS data on these sections and submitted both raw data and analyzed results to be evaluated by a panel consisting of Caltrans engineers and managers. The items subjected to this technical evaluation included: 1) general observations from each service provider’s data collection effort in the field (safety, efficiency, etc.), 2) data and file management, 3) accuracy of location-referencing information, 4) quality of collected images, 5) pavement distress/condition measurements based on images and profiles, 6) a data viewer supplied by the service provider, and 7) the service provider’s technical competence as demonstrated in an oral interview with the evaluation panel.

Most of the data items delivered by each service provider were evaluated on an item-by-item and section-by-section basis against predefined acceptance criteria. The percentage of all data items that meet the acceptance criteria, pooled among all the ten sections was used to calculate the service provider’s technical score. Although service providers are always evaluated in one way or another in APCS procurements of most highway agencies, this kind of item-by-item evaluation is somewhat uncommon. In the case histories reported in the literature, data subjected to comparison were either a small number of specific measurements such as International Roughness Inces (IRI) (2), or composite indices as an overall measurement of pavement condition calculated based on detailed roughness and distress quantifications (5,10,11). The decision to perform a detailed item-by-item evaluation in California was made for the following two reasons.

1. A composite pavement condition index cannot sufficiently represent the accuracy of service provider’s data. An undesirable yet very likely situation is that a service provider happens to obtain an “accurate” composite index by systematically underestimating some distresses but overestimating some others.
2. The data to be delivered for this APCS contract will be used for performance modeling and for other purposes in addition to computing composite indices for network-level decisions. It is highly desirable to ensure that every data type is accurately measured and properly processed.

In addition to the technical assessment, service providers will also be evaluated for their technical and cost proposals. The service provider that obtains the highest combined score will be selected for potential awarding of the contract. In the combined score, 40% is contributed by the technical assessment, 30% by the proposal evaluation, and 30% by the proposed cost. The latter two items are not discussed in this paper.

**Data Acceptance Criteria**

For quantitative data items, such as the average IRI of a section or the average count of transverse cracks per concrete slab, data acceptance criteria are specified in the RFP package. The acceptance criterion for each data item is expressed as a tolerable range of absolute difference and a tolerable relative difference between the value reported by the service provider and the ground truth reported by the agency. These acceptance criteria were determined by the review panel on a consensus basis, supported by UCPRC research. The tolerable absolute differences generally reflect the inherent noise levels in the measurements, and the requirements for relative differences were posed by the anticipated needs of the PMS. If either the absolute difference or the relative difference is within the specified tolerance, this data item is considered to meet the quality requirement. As an example, the tolerable absolute difference for the average Mean Profile Depth (MPD) of a 180 m long section is 0.2 mm and the tolerable relative difference is 20%. Assume the ground truth average MPD determined by the agency is 0.5 mm, and the value reported by the service provider is 0.7 mm. The absolute difference is |0.7-0.5|=0.2 mm and the relative difference is 0.2/0.5×100%=40%. Although the relative difference criterion was not met (40%>20%), the service provider-reported MPD value is considered acceptable because it meets the absolute difference criterion (0.2 mm ≤ 0.2 mm).

**What Would Happen After the Service Provider Is Selected**

As specified by the RFP, the selected service provider will participate in an annual startup process before receiving approval for statewide data collection. In the annual startup, the service provider will collect and process data on a number of short sections selected by the agency. In addition, a collection–evaluation–feedback–correction and recollection loop will be followed until the data accuracy meets predefined acceptance criteria. The data collection and evaluation will follow similar procedures as those for the service provider evaluation, but much more stringent criteria (narrower tolerable range around the ground truth) will be employed. It is in the best interest of both the agency and the service provider that the problems identified and lessons learned in the technical evaluation stage can be used to help the service provider solve these problems and thereby improve data quality. During the statewide survey, experience gained in the service provider evaluation process will also benefit the agency by guiding the QA program.
DETERMINATION OF THE GROUND TRUTH

This section describes the determination of the ground truth condition for the ten test sections by the agency. The most obvious measure for ensuring the accuracy and credibility of the determined ground truth is to use certified, calibrated, and well-maintained equipment, and to follow standard procedures. Additionally, two means of “triangulation” were adopted in this study: First, two or more independent methods were used to collect and process most of the data items, and second, different data items were synchronized to “cross-examine” each other. The term “triangulation” refers to the research approach of using more than one method to study a specific problem, mostly used in social science disciplines. The evaluation team found it to be an effective approach for establishing the credibility of the ground truth data collected by the agency.

Transverse Profile

Transverse profiles were measured using a custom-built beam-type portable laser profilometer originally developed for accelerated pavement testing, as shown in Figure 1(a). The unit is powered by a rechargeable battery. A laptop computer controls the unit and collects data through a wireless connection. This beam profilometer has an effective measure span of 3.4 m, slightly shorter than typical lane widths (~3.6 m) in California but sufficient for covering both wheelpaths. Its vertical resolution is 0.1 mm and the sampling rate is 1 mm per data point. Therefore, the measured transverse profiles are essentially continuous. The productivity rate is approximately 2 minutes per profile and transverse profiles were acquired very 5 m along the 180 m long test sections as shown in Figure 1(b). A raw transverse profile is shown in Figure 1(d). A box placed on the pavement surface when the measurement was taken to mark the lane centerline appears in the middle of the profile, and two cracks across the transverse profile are shown as negative spikes at $x = -0.75$ m and $x = 0.5$ m.

A special algorithm was developed to calculate rut depths from the measured transverse profiles. Existing algorithms typically identify “highest” and “lowest” points in a profile trace and use the differences in height between these points as the rut depths. As shown in Figure 1(d), the newly developed algorithm first applies a special low-pass filter to the raw profile and the filtered profile is an upper envelope of the raw profile. Then the algorithm simulates two straightedges in the left and right wheelpaths, respectively, and the vertical distances between the straight edges and the lowest points in the filtered profile are taken as the rut depths. This algorithm is essentially an implementation of the ASTM E1703 Standard Test Method for Measuring Rut-Depth of Pavement Surfaces Using a Straightedge, and it can correctly consider the effects of macrotexture and cracks across the profile as shown in Figure 1(d). Note that if these effects are not handled properly, rut depths will be systematically overestimated, which is possibly one of the reasons why automated rut measurements are often larger than manual measurements as reported in the literature (4). More detailed description of this algorithm is beyond the scope of this paper.
Figure 1  Profile measurement and processing using a portable beam profilometer.
Longitudinal Profile

Three methods were used to collect high and low resolution longitudinal profiles. The agency owns a Road Surface Profilometer® 5051 Mark III manufactured by Dynatest® mounted on a passenger car. Two months prior to the testing, this unit was certified by the Texas Transportation Institute (TTI) based on AASHTO PP49 Certification of Inertial Profiling Systems. The first method used this unit’s inherent capability of acquiring low resolution (25 mm sample interval) longitudinal profiles, and data collection was performed at 96 km/h (60 mph). Ten replicate measurements were conducted on each section for both wheelpaths. The second method used the same unit but different acquisition and processing software to acquire high resolution longitudinal profiles (1 mm sampling interval) in the right wheelpath. Since this equipment was designed to measure Mean Profile Depth (MPD) at highway speeds and discard the high-resolution profile data, and is not designed to measure raw profiles with sampling intervals smaller than 25 mm, some major modifications to the software were necessary and the vehicle had to be operated at a speed of approximately 6 km/h in order to save the raw readings of the laser sensors and the accelerometers. The raw data were then processed with a custom-developed software package in an office environment to obtain the high resolution profiles. Although these two methods used the same inertial profilometer, since the data processing was performed using different software packages, the measurements were considered largely independent.

To provide an independent check of the vehicle mounted profilometer, the portable beam profilometer shown in Figure 1 was used to measure high resolution longitudinal profiles in the right wheelpath, which was the third method. The profilometer was manually moved at an interval of 3.0 m as shown in Figure 1(c), and the profiles were “stitched” together to obtain a continuous measurement. Since the measurement span is 3.4 m, there was a 0.4 m overlap between any two adjacent short profiles to establish continuity. Apart from strict equipment certification and calibration, longitudinal profiles obtained with these three methods were examined against each other to establish the credibility of the measurement as shown later in this paper.

Image-Based Distress Quantification

In the Caltrans APCS, imaged-based pavement distresses are classified and quantified from pavement images with resolution better than 2 mm per pixel according to the Caltrans APCS Manual. The raw pavement images submitted by the service providers participating in the demonstration were checked against visual observations made from the shoulder, and also subjectively assessed by the evaluation panel. After passing this image quality evaluation, they were considered to represent the ground truth condition in terms of surface features of the sections and the ground truth for image-based distress measurements were derived from these images. Three factors were considered to justify the approach. First, studies in the literature reported that image quality is generally not an issue for modern equipment thanks to recent significant advances in digital imaging technologies (12). Second, in an APCS rodeo organized by Caltrans and the UCPRC in 2008 (13), it was found that most service providers can deliver images clear...
enough for characterizing cracks and other distresses that Caltrans is interested in. Third, use of images delivered by service providers to check distress quantification data reported by the same service providers is a common practice in APCS QA (8).

To classify and quantify distresses in a convenient and efficient manner, a comprehensive pavement information reviewer, CaPCIV (California Pavement Condition Information Viewer) was developed at the UCPRC. The basic functions of CaPCIV include: 1) synchronization and display of all the raw data collected in the APCS, 2) flexible yet easy image manipulation and profile filtering, and 3) generation of reports of distress classification and quantification. In CaPCIV, crack/patch/pothole maps are produced manually by the operator but quantification and report generation is automatic. The productivity of the software is not high enough for network level APCS work, but it is sufficient for data evaluation and QA of short sections sampled from the network. Developing such a tool was necessary because 1) viewers developed by APCS service providers are proprietary software packages and their source codes are not available, 2) these viewers are compatible with neither the data collected by the agency nor data collected by other service providers while data format openness is essential to APCS in California, and 3) image and profile manipulation and synchronization in these packages do not meet the high expectations of the demonstration evaluation team and the future QA team in California.

**Location-Referencing**

In each of the ten test sections, geographic coordinates (longitudes and latitudes) of two “anchor” locations were collected using a survey-grade handheld GPS unit with differential correction. The error of such measurements is smaller 0.5 m. These locations were marked with paint before the demonstration so that they are visible in the images collected by the service providers. In CaPCIV, after the images on a section are “stitched” together, these reference locations are identified and the coordinates are input. As long as the section is straight, CaPCIV can accurately calculate georeference information of any point in the stitched images based on the coordinates at these two anchor points. This function was used to evaluate the accuracy of reference information delivered by service providers as presented in one of the subsequent sections of this paper.

**Synchronization and Cross-Examination**

In this study, synchronization of different measurements so that they can cross-examine each other was an important means to ensure the quality of the data. An example is shown in Figure 2. This is a jointed portland cement concrete pavement section with skewed joints and varying joint spacing. The upper portion of Figure 2 shows stitched pavement images of a 20 m long segment, and the lower portion shows four synchronized profiles measured using different equipment. Note that Profiles a, b, and c were all measurements by the agency and profile d was submitted by a service provider. A high-pass filter (subtracting the 3 m moving average from the raw data) was applied to all the profiles to facilitate comparison of the short-wavelength features. It can be observed that the discontinuities in the profiles (at distances of 103, 107, 112, and 118 m) correspond to pavement joints accurately, and the slab
warping/curling is also reflected in all the profiles. In Profiles a, b, and c, there are unusual spikes at the
distance of 109 m that are not related to joints. A close observation of the high resolution images found
that there was a refilled core hole approximately 100 mm (4 inches) in diameter at this location and the
refilled material was slightly higher (by approximately 5 mm) than the slab surface. This feature is not
visible in Profile d because its sampling rate was not high enough. The high level of agreement between
these independent profile measurements as well as between the profiles and the images provided great
confidence in their accuracy. Although not shown in this paper, the synchronization and
cross-examination between pavement images and transverse profiles was also successful, with cracks
across the profiles traces clearly showing in the profiles.

![Figure 2](image)

**FIGURE 2** Synchronization and cross-examination of profiles and images.
Note: Profile a is a high resolution (sampling rate of 1 mm) profile measured with the agency’s inertial
profilometer; Profile b is the stitched high resolution (1 mm) profile measured using the portable beam
profilometer by the agency; Profile c is the low resolution (25 mm) profile measured with the agency’s
inertial profilometer; Profile d is the low resolution (~100 mm) profile delivered by one of the service
providers.
EVALUATION RESULTS

The Accuracy of Georeference Information

In the Caltrans APCS, all data items are required to carry independent location-reference information, in the form of either geographical coordinates (latitudes and longitudes), linear reference, or both. One of the objectives of the APCS demonstration was to evaluate whether the proposing service providers can modify their system to accommodate this requirement and also to evaluate the accuracy of the georeference information. Figure 3 shows an example of evaluating a service provider’s georeference accuracy on two sections. Section A is a jointed concrete pavement with skewed joints and varying joint spacing. Section B is also a jointed concrete pavement, but the joints are perpendicular to the traffic direction and the joints have a relatively uniform spacing of 4.5 m. Section A is not cracked while extensive cracking is visible in Section B. All joints and cracks are sealed. Traffic is from left to right in the images and a compass is also shown for each section.

![Figure 3 Evaluation of the accuracy of georeference information delivered by a service provider.](image)

Service providers are required to report geographical coordinates for each pavement slab at the slab center, as denoted by the triangular symbols in Figure 3. The geographical coordinates reported by one of the service providers are denoted by circular symbols. The arrows shown connecting the triangular symbols and the circular symbols denote the errors in the service provider’s geographical coordinates. The reported locations were consistently ahead of the true locations by 4 to 5 m and left of the true locations by 0.5 to 1.5 m. This trend is true for both Section A and Section B, as well as for other test sections not shown in Figure 3 regardless of the directions of traffic. Although this error was out of the acceptable range for Caltrans APCS, the clear and consistent pattern indicates that it is a systematic error that would be easy to fix. This identified pattern will help this service provider (if eventually selected) take corrective measures during the annual startup process and finally meet the accuracy requirements.

Evaluation of Rut Depth Measurement

For each test section, service providers reported mean rut depths for the left and right wheelpaths, at a reporting interval of 10 m. The APCS specification does not have a specific requirement for the type of
rut sensors (scanning laser or a group of point lasers), however, a large number of point lasers would be needed for the rut depth definition and to meet the data acceptance criteria. Both service providers in the technical evaluation used scanning lasers. The mean rut depths of each test section was calculated and evaluated against the ground truth values determined by the agency. The correlations between the agency measurements and service provider measurements are plotted in Figure 4(a) and (b) for Service provider A and Service provider B, respectively. The agency only collected transverse profiles on six test sections due to limited traffic closure availability, and therefore, in each part of the figure, twelve (6 sections x 2 wheelpaths) data points are shown. The data acceptance criterion in the RFP allows a maximum allowable difference of 3 mm or 20% (whichever is larger) between the service provider data and the ground truth, evaluated on a section-by-section basis. In Figure 4, as long as a data point is within the shaded area, this service provider’s data on the corresponding section is considered acceptable. The six sections happen to have relatively small (all smaller than 6 mm) rut depths, in part because asphalt rutting is generally difficult to find on the Caltrans network, and therefore the absolute difference criteria were always more critical than the relative difference criteria in this evaluation.

Apart from the section-by-section evaluation, the overall accuracy of each service provider’s data was evaluated by the regression method (4,5). The linear trend lines are forced to go through the coordinate origins, and the slopes indicate the overall bias of the service providers’ data compared to the ground truth measurement. The coefficients of determination ($R^2$) quantify how strong the correlations are. As shown in Figure 4(a) and (b), Service provider A’s rut depth data generally agree with the ground truth measurements, while Service provider B reported rut depths are consistently larger than the ground truth, on average by 46 percent. Service provider B’s data also appear to be more scattered ($R^2=0.70$) than Service provider A’s data ($R^2=0.82$). Although Service provider B only failed two out of twelve data points according to the data acceptance criteria in this evaluation, if it were to be selected to perform the statewide APCS service, it would have to considerably improve its rut depth measurement to meet the more stringent criteria (error less than 2 mm or 10 percent) in the annual startup process. The possible reason for Service provider B’s systematic overestimation of rut depths could be poorly calibrated laser sensors, or improper handling of macrotexture or cracks when calculating rut depths from the transverse profiles. All the data acceptance criteria including those in the demonstration, in the annual startup and during statewide data production are included in the RFP and therefore cannot be changed after awarding the contract. However, the evaluation results indicated that the agency should consider including a few test sections with more severe rutting in the annual startup to better evaluate the selected service provider’s ability to quantify this distress.
(a) Rut depth, Service provider A  
(b) Rut depth, Service provider B  
(c) MPD, Service provider A  
(d) MPD, Service provider B  
(e) IRI, Service provider A  
(f) IRI, Service provider B

FIGURE 4 Evaluation of profile-based distress/condition data reported by service providers.
Evaluation of MPD Measurements

Mean Profile Depth (MPD) measurements reported by the service providers were evaluated following the same approach as that for rut depths. MPD was only required for the right wheelpath of flexible surfaced pavements, so only six data points were available for each service provider. Figure 4(c) and (d) show the correlations between a service provider’s data and the ground truth, and the acceptance criterion (0.2 mm or 10 percent) is shown as the shaded areas. Service provider A failed on one section and Service provider B failed three out of six sections. Moreover, Service provider A appeared to slightly but consistently underestimate MPD, whereas Service provider B tended to overestimate MPD for surfaces with smooth macrotexture (MPD<1.0 mm) and underestimate MPD for surfaces with rougher macrotexture (MPD>1.5 mm). The evaluation team attempted to identify the reasons for these problems in MPD by checking the raw high resolution profiles (1 mm sampling interval) submitted by the service providers. However, the raw profiles turned in by one of the service providers appeared to be “over-filtered”. Features such as concrete slab joints and transverse cracks could not be identified in the profiles and synchronization between these profiles and images as well as the agency’s raw profiles was not successful. The other service provider failed to submit high resolution profiles to the evaluation team. Because communication between the evaluation panel and the service providers was prohibited during the evaluation process for legal reasons, these issues were not resolved. The annual startup process after the APCS contract is awarded will provide a mechanism in which the agency and the selected service provider can effectively communicate and work together to identify and resolve technical problems.

Evaluation of IRI Measurements

As shown in Figure 4(e) and (f), both service providers did very well with their IRI measurements. IRI values were reported for both wheelpaths on all ten test sections, and therefore Figure 4(e) and (f) shows twenty data points. Ground truth values were the average of ten replicate agency measurements. A comprehensive data review found that one service provider switched lanes in the middle of a section and the other identified the wrong start point for a section. The data affected by these mistakes are denoted by cross symbols in Figure 4(e) and (f), and they are excluded from the linear regression. Apart from this issue, Service provider A had almost perfect IRI measurements, while Service provider B’s data are slightly more scattered. The accuracy of IRI measurement by APCS service providers was mostly reported to be satisfactory in the literature (4), possibly due to the long history of equipment improvement and procedure standardization (14,15).

Image-Based Distress Measurement

Service providers’ image-based distress classification and quantification was assessed on an item-by-item and section-by-section basis. These distresses are mostly related to cracking and patching. The main objective of the assessment was to evaluate whether each service provider could correctly understand and implement the new distress classification and quantification system in Caltrans’ new APCS Manual. The results showed that misunderstanding of certain distress types was apparent, but presenting detailed results
of this evaluation is beyond the scope of this paper. It is expected that the annual startup process will provide the agency with good opportunities to educate the selected service provider on the new distress classification and quantification system, especially on the misunderstandings revealed in the evaluation. Significant improvement on related data quality is anticipated. The efficiency of service providers’ automated image analysis is another important concern in APCS, but this was indirectly but sufficiently reflected in their cost proposal and thus not included in the technical evaluation.

CONCLUDING REMARKS

The California Department of Transportation (Caltrans) with support from the University of California Pavement Research Center (UCPRC) (collectively referred to as the agency in this paper) recently evaluated a few APCS service providers as part of the RFP process for awarding a statewide APCS service contract. This paper presented the technical innovations made, experience gained, and lessons learned in this evaluation.

Technologies involved in APCS are constantly and rapidly evolving, and are often outside the skill set of practicing pavement engineers and managers. Evaluation of APCS service providers’ technical competence requires the highway agency to actively make technical innovations in evaluation methodologies and especially in ground truth pavement condition determination. As two examples, precise synchronization of different data items on the same section was an effective way to establish the credibility of the ground truth data, and the development of a comprehensive data viewer with open standards greatly enhanced the efficiency in handling complex data in this work. Caltrans found that collaborating with researchers in pavement engineering and pavement evaluation was essential to the success of this project.

The choice made by the agency to evaluate service providers’ data on an item-by-item and section-by-section basis (instead of assessing a composite index) was proven to be rewarding. The focus in the RFP and APCS manual on performance requirements as opposed to detailed specification of particular technologies makes this process easier to implement. Many detailed problems that would have been hidden in a composite index were identified, which will help improve the selected service provider’s data quality before the execution of the statewide APCS contract.

Communication between the agency (which will be the eventual data user) and service providers is crucial to improving data quality. Although the evaluation team was not allowed to communicate with the service providers during the evaluation for legal reasons required by the contracting officers, open communication is anticipated once the service provider is selected and the contract signed. The agency and the service provider will work together to improve the data quality in the annual startup process, which will save resources and time for both the agency and the service provider.

Acknowledgements:

The work presented in this paper was sponsored by the California Department of Transportation, Division of
Research and Innovation, and the Pavement Management Program of the Division of Maintenance. Caltrans sponsorship is gratefully acknowledged. The contents of this paper reflect the views of the authors who are responsible for the facts and accuracy of the data presented herein. The contents do not necessarily reflect the official views or policies of the State of California or the Federal Highway Administration. This paper does not constitute a standard, specification, or regulation. Creative inputs and feedback from the Caltrans APCS Demonstration Evaluation Panel are gratefully appreciated. Panel members included Mr. Tom Pyle, Dr. T.J. Holland, Dr. James N. Lee, Mr. Peter Vacura, and Mr. Robert Hogan. The authors also want to acknowledge the efforts of Mr. Mark Hannum, Mr. Mark Troxler and other participants of this study at the UCPRC in equipment improvement and data collection.

References:


