Technical Paper

ISSN 1997-1400 Int. J. Pavement Res. Technol. 6(1):22-30

Developing a Quality Measure for Evaluating the Contractor Performance in Asphalt Pavement Construction

Ahmed Elyamany[1], Magdy Abdelrahman[2+], Fu-Chih Cheng[3], and Curt Doetkott[4]

Abstract: Most construction agencies have a quality management system to control and manage the quality of their final product. Once the project is over, the tests results are usually kept in archives with no use. Project bidding is a time when the quality information could be very helpful to identify qualified contractor. The proposed model is an effort to use quality tests in evaluating the contractor performance. The composite percent defective is proposed as the quality measure which combines percent defectives of a number of selected quality characteristics. The trend analysis is used to provide a better look at the contractor performance. The significance of the proposed measure is in its ability to reflect the true quality of the contractor work being calculated based on actual tests results.

Key words*: Acceptance; Pay factor; Percent defect; Probability; Quality measure; Trend analysis.*

Introduction

The current practice of a number of State Highway Agencies (SHAs) is to evaluate and rate the contractor performance. Most of these cases do not have a well established measure or index to track the contractor performance. Rather, they use subjective questionnaires filled by project engineers and managers. This is mainly because in most states public projects must be bid by traditional low-bid method [1]. Parties involved in construction think the low-bid method decreases the value of the construction. In low-bid contracting, the contractor focuses on how to lower the cost rather than developing a quality product with a low cost. Some owners think quality has no price and only look for the lowest bid which, in fact, is responsible for the cost and time overruns in their projects. Saving the taxpayer money is an acceptable reason to believe in the low-bid; however this belief must change as most of the studies show losses in the long run [2, 3]. The problems associated with the low-bid system encourage contractors to implement cost-cutting measures instead of quality-enhancing measures. Therefore, it is less likely that the contracts will be awarded to the best-performing contractors who will deliver the highest quality projects with minimum cost [4]. The significance of having the performance measure for quality is to compensate the good contractors by allowing them to bid on more projects or at least let the other contractors account for their bad performance.

Several attempts to develop a rational system of rating the quality of construction were documented. Most of these ratings were used for qualification and bidding processes. The construction industry lacks the existence of a standard rating system for quality based on measurable, not subjective data. Actually there are a lot of data especially for quality control and quality assurance purposes. However, the major challenge is how to combine multiple distributions of different quality characteristics.

Study Objective

The main propose of this study is to develop an analytical model for combining the effects of multiple deficiencies of quality characteristics distributions, for contractors working in the pavement industry, using basic statistical concepts. The inputs of the model are the individual percentage defectives of multiple quality characteristics, while the output is the composite percent defective. The Composite Percent Defective (PD_C) is based upon assuming specific underlying distributions for individual quality characteristics, as well as the relationships among the individual quality tests. When the underlying assumptions are applicable, statistical principles from mathematical statistics, such as the form of the joint distributions of independent random variables, may be used to derive the expected results for the proposed quality measure. In such cases, simulations may be used to supplement our understanding and to verify the expected theoretical results. The basic statistical concepts are used to classify the contractor performance in two categories: first is the work accepted with full pay and second is the work accepted with partial pay. This classification is common in the pavement industry where defective work can be accepted with partial pay.

Literature Review

Quality Control/Quality Assurance testing results are available in most DOTs, which represent a valuable indication of past performance. As one part of the quality assurance process, there is a need for comprehensive methods to evaluate a contractor's eligibility to engage in work from a quality perspective; thus, there is a need for examining quality performance measurement techniques and approaches. The main purposes of rating the contractor performance are for qualification, bidding, or payment schedules. Weed proposed a method for developing pay schedules based on the need for a rational method to relate As-Build quality to

[1] Department of Construction Engineering, Zagazig University, Zagazig-Egypt.
[2] Department of Civil Engineering, North Dakota State University, Dept. 2470, PO Box 6050, Fargo, ND 58108, USA.
[3] Statistics Department, North Dakota State University, USA.
[4] Information Technology Services, North Dakota State University.
[+] Corresponding Author: E-mail m.abdelrahman@ndsu.edu
Note: Submitted April 3, 2012; Revised August 1, 2012; Accepted August 6, 2012.

expected performance and for use in the development of reliable and defensible pay schedules [5]. This method is believed to reflect more accurately the value of failure to meet the design level of quality because the actions upon which any pay reduction is based are not a function of the thickness of the pavement layer or the bid price. In a different reference, Weed offered a method for combining the effects of multiple deficiencies [6]. A rational and feasible method for quantitatively calculating pay factors was described by Monismith for asphalt construction [7]. Performance models were developed for fatigue and rutting based on results of tests on Caltrans Heavy Vehicle Simulator and WesTrack accelerated pavement performance test program. Whiteley developed a method for generating pay factors based on the Life cycle Cost (LCC), as well as between LCC and pay factors [8]. Minchin developed a construction quality index which is a rating of the quality of materials and workmanship on highway projects [1]. This index used actual numerical test results to rate quality of highway construction projects. The drawback of the quality index is using a questionnaire to assign the weights of quality characteristics.

Measuring Quality

The majority of highway agencies have their own specifications for QC and Acceptance plans. The National Cooperative Highway Research Program (NCHRP) sponsored a project to develop a generic QC methodology for pavements constructed with Superpave mixes [9]. The QC process requires that the contractor collect samples of the asphalt concrete. A random sampling process is used to ensure the material is representative of the total amount of material placed during the project [10].

The normal distribution assumption could be used both to describe the quality level desired and to assess the quality level actually achieved. This characteristic has led many agencies to define the specification limits in terms of percent-within-limits (PWL) [11]. PWL is defined as the percentage of the lot falling above the lower Specification Limit (LSL), beneath the Upper Specification Limit (USL), or between the LSL and the USL [12]. This quality measure uses the sample mean and the sample standard deviation to estimate the percentage of the population (lot) that is within the specification limits. In theory, the use of the PWL method assumes that the population being sampled is normally distributed. In practice, it has been found that statistical estimates of quality are reasonably accurate, provided the sampled population is at least approximately normal, i.e., reasonably bell-shaped and not bimodal or highly skewed. The PWL is the recommended statistical measure for material and construction quality in many applications with a quality level of approximately 90% being considered acceptable, which means 10% is defective or of lesser quality [11].

Pay Factor

Assigning a score for the contractor based on the quality of the pavement material is equivalent to adjusting the payment with what is known as the Pay Factor (PF). The PF can be defined as a multiplication factor that is often used to determine the contractor pay for the unit of work. After the project or a project stage is completed, the owner/agency evaluates the product, and based on this evaluation, the contractor gets paid. The contractor could be paid in full, penalized, or rewarded, depending on the performance and the quality of the final product [12]. Pay Factor is calculated using empirical equations suggested by the agency such as that shown in Eq. (1) [11].

$$PF = 55 + 0.5 * PWL \quad (1)$$

This equation assumes the maximum and the minimum PF are 105 and 55 at 100PWL and 0PWL, respectively. Many practitioners and researchers suggest the Accepted Quality Limit (AQL) to be satisfied at 90PWL with a PF equal to 100. They also suggested the Rejected Quality Limit (RQL) to be satisfied at 50PWL with a PF equal to 80 [11]. Eq. (2) is studied as a non-linear PF equation [11].

$$PF = 2.4 * PWL - 0.01 * PWL2 - 35 \quad (2)$$

This equation assumes the minimum and the maximum PF are 0 and 105 at 15.6PWL and 100PWL, respectively. Since the minimum PF of 0 is not rational, this equation should have a minimum PWL between 40 and 50 to keep the minimum PF between 45 and 60. The comparison between both equations is illustrated in Fig. 1.

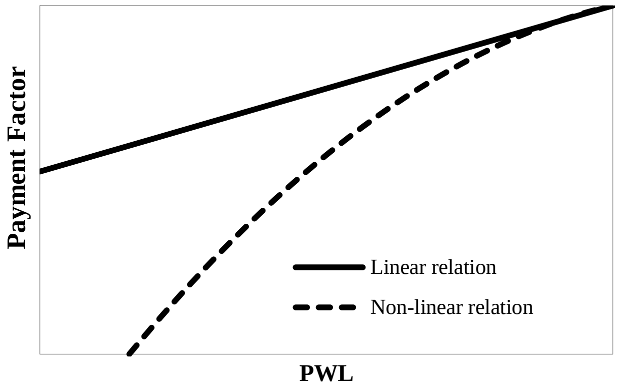

Fig. 1. Graphical Representation of Pay Factor Equations.

Table 1. Quality Characteristics Considered by State DOTs for PF.

Quality Characteristic	% Consideration
Density	77.3
Asphalt Content	77.3
Gradation	72.7
Air Voids	27.3
Void Mineral Aggregate	22.7
Smoothness	18.2
Thickness	9.1

Table 1 shows the quality characteristics considered by a sample of 22 State Highway agencies in the PF calculations [11].

Most project specifications will contain multiple quality characteristics, which results in using multiple pay factors. Combining multiple pay factors is essential to have a single pay factor for a lot. The ideal situation to combine multiple pay factors is to connect them to the long-term performance of the pavement. Unfortunately, such relations are not widely accepted at this time [11]. Various agencies have considered at least four different approaches for combining a number of pay factors for individual acceptance quality characteristics into a single combined pay factor

[11]. These approaches include [11]:

- Using the minimum individual pay factor.
- Summing the pay factors.
- Averaging (possibly with weighting factors) the individual pay factors.
- Multiplying the individual pay factors.

The approach using the minimum individual pay factor is based on the "weak link" theory, in which the lowest pay factor indicates the value of all the quality characteristics. For the other three approaches the concept is that all individual factors contribute to the total, but may return different results depending on the value of the individual pay factors. The multiplying approach for combining individual pay factors implicitly assumes that each individual pay factor is equally important. However, several agencies consider some quality characteristics to be more important than others. A questionnaire is used in the current study to estimate the relative importance of the quality characteristics as used in the averaging approach. The responses of the questionnaire are analyzed using the Analytical Hierarchy Process (AHP) to minimize the subjectivity and account for human thoughts and intuitions.

Analytical Hierarchy Process (AHP)

"The AHP provides a means of decomposing a problem into a hierarchy of sub-problems which can easily be comprehended and subjectively evaluated. The subjective evaluations are converted into numerical values and processed to rank each alternative on a numerical scale" [13]. The methodology of the AHP can be explained in the following steps:

The problem is decomposed into a hierarchy of goals, criteria, and alternatives.

Data are collected from experts corresponding to the hierarchic structure, in the form of comparison between pairwise alternatives on a qualitative scale and then converted into quantitative numbers from 1 to 9.

The pairwise comparisons of various criteria generated at step 2 are organized into a square matrix.

The principal eigenvalue and the corresponding normalized right eigenvector of the comparison matrix give the relative importance of the various criteria being compared.

The Consistency Ratio (CR) can be determined using Eqs. (3) and (4) as follows [14, 15]:

$$CI = \frac{\lambda_{max} - \eta}{\eta - 1} \tag{3}$$

$$CR = CI / RI \tag{4}$$

where: CI = the matrix consistency index, η= matrix size, λ_{max} = the maximum eignvalue of the judgment matrix, and RI = random index. Saaty 1982 suggests the value of CR should be less than 10% [14].

The rating of each alternative is multiplied by the weights and aggregated to get global ratings.

Model Development

The equation of the Composite Percent Defective (PDC) is mainly

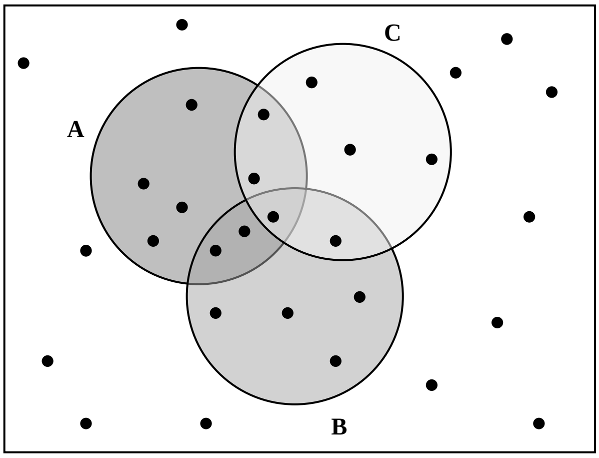

Fig. 2. Venn Diagram.

based on the concept of joint distributions of independent random variables. The Monte Carlo simulation technique is used to verify the proposed equation. The Composite Percent Defective model was developed in the following steps:

- Collecting and processing quality characteristic data.
- Estimating Percent Defective.
- Developing the Composite Percent Defective equation.
- Using the Simulation to verify the developed equation.

Quality Characteristics and Quality Data

The quality of Superpave® mixes is dependent on several materials and construction factors. As shown in Table 1, the specifications of a sample of 22 states, studied by Burati are analyzed in order to select the most important and applicable quality characteristics [11]. According to the study, the top 4 quality characteristics are Field Density, Asphalt Content (AC), Gradation (GR), and Air Voids (AV). Due to the availability of data, only AC, AV, and GR are used in the current model.

Estimating Percent Defective Statistically

The results of quality tests are transformed to Percent Defective (PD) which has been preferred in recent years because it simultaneously measures both the average and the variability level in a statistically efficient way. The PD can be calculated from PWL using the simple relationship, PD = 100 – PWL. The use of PD as a quality measure has some advantages, particularly with two–sided specifications, because the PD below the lower specification limit can simply be added to the PD above the upper specification limit to obtain the total PD value [16]. Conceptually, the PWL procedure is based on the normal distribution. The area under the normal curve can be calculated to determine the percentage of the population that is within certain limits. Similarly, the percentage of the lot that is within the specification limits can be estimated. The interested reader may refer to the reference by Burati for more details on how to estimate PWL and PD values [11].

Concept of the Composite Percent Defective

Consider the Venn diagram, Fig. 2, in which each of the dots represents an outcome and each of the three circles represents an event.

The probability of the union of three events, A, B and C is given by:

$$p(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) - P(B \cap C) + P(A \cap B \cap C) \quad (5)$$

where; P(A), P(B), and P(C) are the probability of occurring for event A, B, and C, respectively.

$P(A \cap B)$ is the set of outcomes that belong to both A and B. $P(A \cap C)$ is the set of outcomes that belong to both B and C. $P(B \cap C)$ is the set of outcomes that belong to both B and C. $P(A \cap B \cap C)$ is the set of outcomes that belong to A, B, and C.

The conditional probability of two events A and B is defined as the probability of one of the events occurring, knowing that the other event has already occurred. Eq. (6) denotes the probability of A occurring given that B has already occurred.

$$P(A|B) = P(A \cap B)/P(B) \quad (6)$$

If knowing B gives no information about A, then the events are said to be independent and the conditional probability expression reduces to:

$$P(A|B) = P(A) \quad (7)$$

From the definition of conditional probability, Eq. (6) can be written as:

$$P(A \cap B) = \mathrm{P}(A|B)\mathrm{P}(B) \quad (8)$$

Since events A and B are independent, Eq. (7) is substituted in Eq. (8). The intersection between A and B is represented in Eq. (9).

$$P(A \cap B) = P(A)P(B) \quad (9)$$

If a group of 3 events A, B, and C are independent, then:

$$P(A \cap B \cap C) = P(A)P(B)P(C) \quad (10)$$

In Eq. (5), the probability of the union of three events, A, B and C, can be rewritten as:

$$p(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A)(B) - P(A)P(C) - P(B)(C) + P(A)(B)(C) \quad (11)$$

Eq. (12) is a simpler form of Eq. (11).

$$P(A \cup B \cup C) = 1 - \left[\left(1 - P(A)\right) \times \left(1 - P(B)\right)\left(1 - P(C)\right)\right] \quad (12)$$

The term $\left(1 - P(A)\right)$ in Eq. (12) defines the probability of A not occurring and the same applies for B and C. To define the absence of each of the three events, $P(\bar{A}), P(\bar{B}), and\ P(\bar{C})$ are used as the complementary event of A, B, and C, respectively. The probability of the union of three events in Eq. (12) becomes:

The probability that A, B, and C do not occur is:

$$P(A \cup B \cup C) = 1 - [P(\bar{A})P(\bar{B})P(\bar{C})] = 1 - P(\bar{A} \cap \bar{B} \cap \bar{C}) \quad (13)$$

$$P(\bar{A} \cap \bar{B} \cap \bar{C}) = 1 - P(A \cap B \cap C) = (1 - P(A)) \times (1 - P(B))(1 - P(C)) \quad (14)$$

Composite Percent Defective Equation

The actual practice in the pavement industry does not reject the sample if the test result falls outside the specification limit. Shifting the focus from accepting/rejecting the sample to the PD allows the agencies to accept the contractor defected work. Two scenarios for accepting the contractor work based on the estimated PD: first to accept the work with full or bonus pay, second to accept the work with partial pay. The proposed model assumes each pavement sample is tested for three quality characteristics; AC, AV, and GR. Using the methodology discussed earlier, each sample will have three estimated PD values referred to as AC, AV, and GR. For each quality characteristic, PD is considered the probability of test result falling outside the specified limits. Based on the result of the three tests, samples could be classified into two groups; acceptance with full pay and acceptance with partial pay. The area outside the three circles, in Fig. 2, indicates the acceptance with full pay region, whilst the area of the three circles, including the intersection area, represents the acceptance with partial pay region. Both regions together form 100% of the space. The area of the intersection between the three circles represents the percentage of failed tests. Ideally, eliminating the contractor who fails in all the quality characteristics tests guarantees a better quality. However, the current practice penalizes the contractor for the poor performance without elimination.

Eqs. (12) and (14) are transformed to Eq. (15) in order to obtain the probability of the sample falling in the acceptance with partial pay region regions.

$$PD_C = 100 - [(100 - PD_{AG}) \times (100 - PD_{AV})(100 - PD_{GR})] / 10000 \quad (15)$$

where; PDC is the composite percent defective and PDAC, PDAV, and PDGR are Percent Defectives for AC, AV, and GR, respectively. Since the probabilities in Eqs. (5) through (14) have a scale of 0 to 1 while PD have a scale of 0 to 100, Eq. (15) have some coefficients in order provide the probability in a scale of 0 to 100. The proposed model will use the Composite Percent Defective (PDC), in Eq. (15), as the measure for the contractor performance.

Composite Percent Defective Using Simulation

Generally, Monte Carlo method is a computational algorithm that relies on repeated random sampling to compute the result. Monte Carlo methods tend to be used when it is unfeasible or impossible to compute the exact result with a deterministic algorithm [17]. The output of a Monte Carlo simulation is a probability distribution describing the probability associated with each possible outcome [10].

The simulation is used here as a tool to verify the use of Eq. (15) in calculating the Composite Percent Defective (PDC). The simulation procedure estimates the percentage of rejected samples as an equivalent to the Composite Percent Defective. The following, and Fig. 3, summarize the steps used to calculate the percentage rejection for each project.

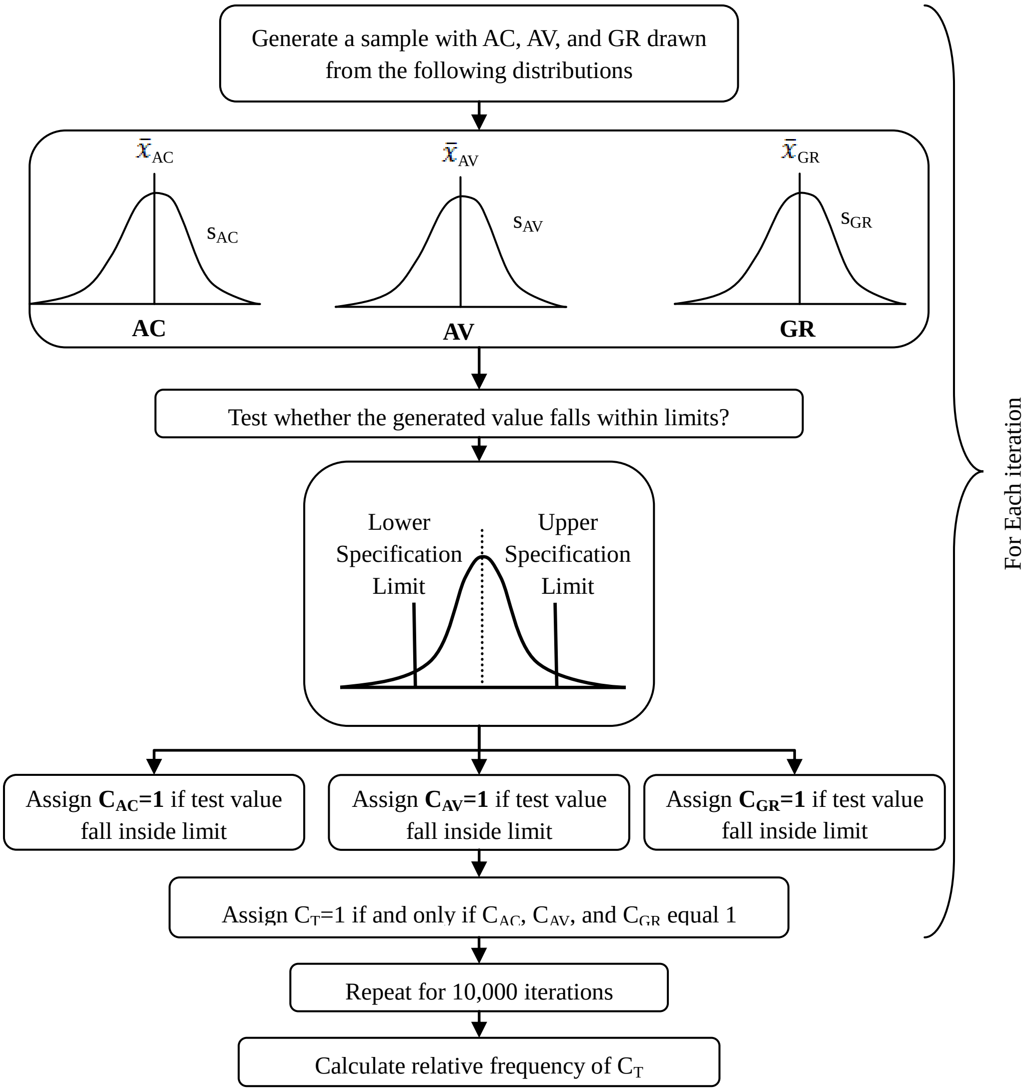

Fig. 3. Simulation Procedure.

- Estimate the sample mean ($\bar{x}$) and standard deviation (s) of the test value for each quality characteristic.
- Establish a normal distribution for the test values with the estimated mean ($\bar{x}$) and the estimate standard deviation (s).
- Obtain the specification limit for each quality characteristic.
- Randomly, select a sample value from the established distribution of the quality characteristic.
- Compare the test value to the specification limit.
- Check acceptance of the test value. Assign Ci=1 if the test value falls inside the specification limit or Ci=0 otherwise.
- Check the acceptance of all the quality characteristics, AC, AV, and GR.
- Accept the lot if, and only if, all the tests of the quality characteristics are accepted. Assign CT=1 if and only if CAC, CAV, and CGR equal 1.
- Run the simulation for 10,000 iterations.
- Estimate the percentage acceptance (PWL) as the percentage of iterations where the lot is accepted.
- Subtract the acceptance percentage from 100 to get the rejection percentage (PD).

Case Study

The proposed model is implemented on a case study of four contractors, who worked with the Nebraska Department of Roads (NDOR), to illustrate the merit and significance of using it in a realistic situation. The Quality Control data used in the analysis were obtained from the records of NDOR of some 500 projects completed between 2000 and 2006. The collected data belong to two types of SuperPave® mixes; SPS and SP4 which are used for Average Daily Truck Traffic (ADTT) <160 and >500, respectively, according to NDOR records [18], [19].

The methodology discussed above selects three quality characteristics to represent the contractor quality in the collected pavement projects. AC, AV, and GR may not be the only representatives of the pavement quality. However they have a well known effect on the pavement quality and are used to adjust the contractor pay. Since each of the collected projects has multiple lots, the sample mean ($\bar{x}$) and standard deviation (s) of each quality characteristic is calculated to form the distribution of the test value. Table 2 shows the calculated μ and σ for a sample of the case study

Table 2. CONA Quality Tests Data.

Project	Mix	AC (%)		AV (%)		GR (%) No. 8 (% Passing)		GR (%) No. 16 (% Passing)		GR (%) No. 30 (% Passing)		GR (%) No. 50 (% Passing)	
		$\bar{x}$	s	$\bar{x}$	s	$\bar{x}$	s	$\bar{x}$	s	$\bar{x}$	s	$\bar{x}$	s
1	SP4	5.47	0.22	4.06	0.65	46.17	1.87	28.01	1.33	17.88	0.80	11.82	1.30
2	SP4	5.34	0.28	4.04	0.72	54.40	3.34	32.27	2.58	20.46	1.88	15.09	1.35
3	SP4	5.81	0.17	4.03	0.58	50.55	2.41	28.75	1.90	18.53	1.10	12.03	0.90
4	SP4	5.51	0.22	4.09	0.51	46.20	1.94	26.83	1.16	17.59	1.63	11.14	1.12
5	SP4	5.17	0.01	4.15	0.01	43.70	0.01	28.20	0.01	18.90	0.01	11.40	0.01
6	SP4	5.46	0.32	4.46	1.12	46.63	3.37	28.62	2.31	17.74	1.51	10.36	0.90
7	SP4	5.21	0.38	4.17	1.49	45.75	1.99	28.36	1.95	17.14	0.73	11.11	0.20

Table 3. Specification Limits for Quality Characteristics.

Specification	LSL	USL
AC (%)	Project Dependent	Project Dependent
AV (%)	3	5
No.8 (% Passing)	39.1	39.1
No.16 (% Passing)	25.6	31.6
No.30 (% Passing)	19.1	23.1
No.50 (% Passing)	15.5	15.5

Table 4. Correlations Matrix of Quality Characteristics.

Pearson Correlation	PD_{AC}	PD_{AV}	PD_{GR}
PD_{AC}	1	0.103	-0.042
PD_{AV}	0.103	1	-0.180
PD_{GR}	-0.042	-0.180	1

Table 5. PD_C Values Adjusted for Project Size.

Mix	CONA	CONB	CONC	COND
SP4	80.52	53.97	77.81	79.98
SPS	59.23	70.87	74.33	51.37

projects. GR has four test points belong to different sieve sizes: No.8, No.16, No.30, and No.50. The Specification limits for the three quality tests are shown in Table 3.

Estimating PD

PD values are estimated for each quality characteristic on a project to project basis. Both μ and σ, calculated from the previous step, are utilized with the USL and LSL to estimate PD.

Composite Percent Defective

The proposed model uses the Composite Percent Defective (PDC) as the measure of the pavement Quality. PDC is calculated using Eq. (15) for individual projects using PDAC, PDAV, and PDGR. PDC values range from 0 for high quality pavement and 100 for poor quality pavement. This interpretation, which at first glance may seem counter-intuitive, makes sense when we recall that as the probability of acceptance with partial pay increases, the probability of acceptance with full pay decreases. The same data used for calculating PDC in Eq. (15) is used next in the simulation process to verify the statistical model results.

Simulation Results

Prior to running the simulation, it is necessary to check for correlation between the simulated variables; PDAC, PDAV, and PDGR. Table 4 indicates that the correlation coefficients of the three variables are very small, which means a weak correlations exist among them. The simulation is performed assuming that the variables are independent and the results are compared to those results from Eq. (15). The paired t-test is used to check the null hypothesis that the mean PDC values for the two methods are equal. The p-value of the t-test is 0.637 which is greater than 0.05. In this case we will not reject the H0 and conclude that the mean of the PDC calculated using statistics is equal to that calculated using simulation. This verifies using Eq. (15) in calculating the PDC.

Project Size Adjustment

The collected data include different project sizes which are characterized by the length of the road. Each contractor has multiple PDC values estimated for each project in the collected data. One option to combine multiple PDC for different projects is to calculate the average value. The other option, which considers the size of the project, is to calculate the weighted average based on the project length. Big projects require more effort and resources to construct than small ones. For example, the credit or penalty for completing a project 6 miles in length should not be equal to that for completing a 1 mile project. Definitely, the 6 mile long project deserves more weight than the 1 mile project. The estimated PDC values are averaged, with the project length as the relative weight, in order to provide a combined PDC value for the contractor.

The adjusted PDC values for two types of mixes are shown in Table 5. The estimated PDC values are averaged, with the project length as the relative weight. In SP4 mixes, this table indicates that CONB has the lowest probability to get partial pay, while CONA has the highest probability to get partial pay. This concludes that CONB performed, on average, better than the other contractors in projects with SP4 mixes. On the other hand, COND, on average, is the best among all other contractors in project with SPS mixes.

Analysis of the Contractor Quality over Time

Looking at one value of the quality measure may not tell everything about the evolving contractor quality. It would be helpful to see how the contractor quality changes over time. The main purpose of analyzing the contractor quality over time is to predict the future

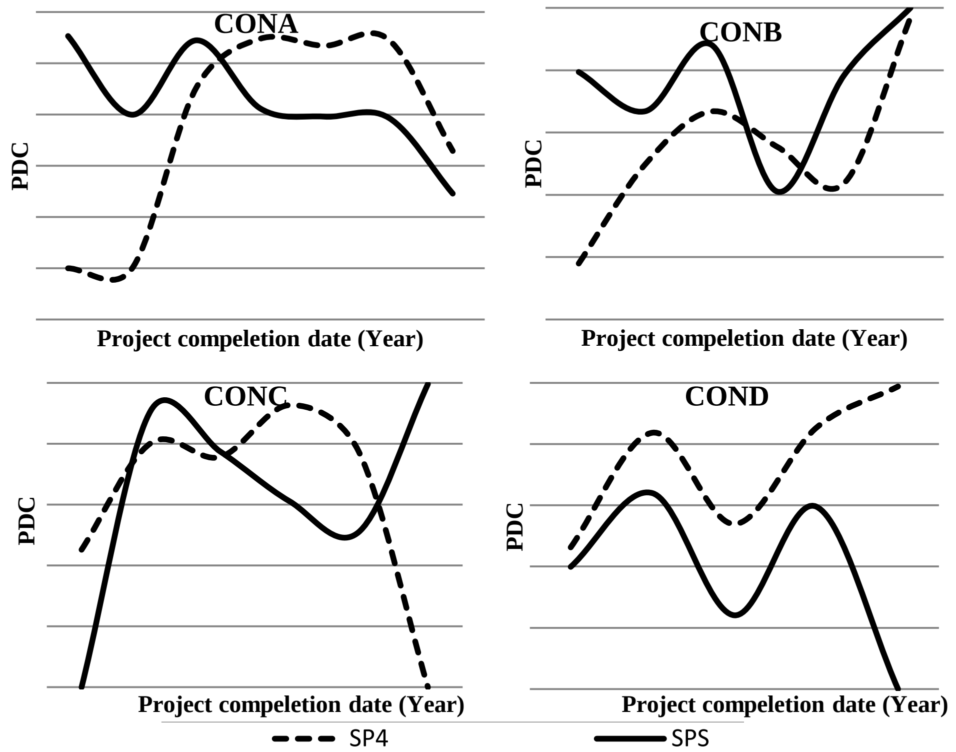

Fig. 4. Contractor Quality Versus Time.

Table 6. Analysis of PD_C Value Over Time.

Mix	Contractor	Fitted Trend Equation		Accuracy Measures
		Constant	Slope	MAPE
SP4	CONA	91.257	-4.97	21.75
	CONB	18.536	+10.43	32.39
	CONC	51.109	+7.71	15.27
	COND	41.596	+10.67	15.89
SPS	CONA	95.0143	- 7.73	17.26
	CONB	66.5773	+ 2.59	23.82
	CONC	78.754	- 0.91	22.45
	COND	42.09	+ 1.92	38.41

quality based on the prior knowledge. Time series analysis is used to study the existence of a trend in the contractor quality over time. A negative slope indicates the PDC value is decreasing over time. Generally, low PDC is better than higher values as it means a low Percent Defective and high quality product. One of the measures to assess the accuracy of forecasts is the Mean Absolute Percentage Error (MAPE). As shown in Eq. (14), MAPE is the mean absolute percentage difference between the true value and the forecast value using the model.

$$MAPE = \frac{1}{T}\sum_{t=1}^{T}\left|\frac{True_t - Forecast_t}{True_t}\right| \tag{14}$$

Fig. 4 shows the PDC values of the contractors plotted over time. A PDC value for each year is obtained from several projects constructed in the same year and averaged by length as mentioned earlier.

The trend analysis in Table 6 indicates that CONA has a negative slope which means the contractor performance is improving in SP4. Other contractors have positive slopes which indicate a deteriorating performance. Both CONA and COND have a negative slope in SPS, yet CONA slope is smaller than CONC and is the best. It is interesting to compare this information with the previous information we obtain by averaging PDC values. In SP4, the performance of CONB, which is the best using the average method, is deteriorating. On the other hand, the performance of CONA, the worst using the average method, is improving. In SPS, the performance of both COND and CONA, the best and second best using the average method, is improving. The conclusion of this analysis is to recommend using the trend analysis, when possible, to study the performance of the contractor over time rather than only looking at the average value. The average value could lead to a false conclusion about the contractor.

Combined Pay Factor

Agencies have developed their own equations to reward or penalize the contractor using the Pay Factor. Using the Pay Factor assumes that giving the contractor a fraction of the full pay would motivate improved performance. The previously mentioned approaches to combine multiple pay factors are compared to the pay factor calculated for PDC to illustrate the significance of using the proposed approach to combine the effect of multiple pay factors. A questionnaire was used to assign the weights of different pay factors when combining the effect of multiple quality characteristics using the averaging approach.

Quality Characteristics Weights Questionnaire

The quality characteristics weights questionnaire was designed to draw feedback from experts regarding the weights of the quality characteristics. The questionnaire was sent to the District Engineers of the DOTs in Iowa, Kansas, Minnesota, Nebraska, North Dakota and South Dakota, in order to gather their subjective opinions regarding the quality characteristics' weight. The questions are the following:

- Question 1; Select the importance of asphalt content of the bituminous mixture as a measure of the quality and long-term performance of the asphalt concrete pavement?
- Question 2; Select the importance of in-place air voids of the bituminous mixture as a measure of the quality and long-term performance of the asphalt concrete pavement?
- Question 3; Select the importance of aggregate gradation of the bituminous mixture as a measure of the quality and long-term performance of the asphalt concrete pavement?

The engineers were asked to select the most appropriate importance level from five levels ranging from "not important" to "absolutely important". The collected data from this questionnaire was used to develop the quality characteristics' weights. Eighty District Engineers from the six state DOTs were contacted by email and asked to answer the questionnaire questions. The response rate of the questionnaire was 38%.

The responses to the quality characteristics weights questionnaire show that the Engineers rated all three quality characteristics as "very strongly important" in 50% of the responses. Nearly 40% of the responses were distributed between "strongly important" and "absolutely important". Only 7% of the responses were distributed between "weakly important" and "not important". Since the main objective of the questionnaire is to determine the quality characteristics' weights, the AHP is used to analyze the questionnaire responses and estimate the quality characteristics' weights.

Analytical Hierarchy Process

The earlier mentioned steps in applying the AHP technique are carried out in order to generate the quality characteristics' weights. The pairwise comparison matrix has a dimension of 3×3. The CR for the matrix is 0.011 (less than 0.1), which means it is a consistent matrix and the weights of this matrix are justifiable and valid. Values of the quality characteristics weights that result from the AHP technique application are 0.33, 0.35, and 0.32 for AC, AV, and GR, respectively. It is noted that the weights are relatively equal, which is rational since the three quality characteristics are inter-correlated and believed to have the same importance and effect on the long-term performance of the asphalt pavement.

Comparison between Combining Approaches

The minimum, multiplying, average, and the proposed approach to combine multiple pay factors are studied to determine the significance of using the proposed approach. Each approach was used to combine three individual pay factors for AC, AV, and GR. The proposed approach provides a composite percent defective (PD_C), which is used in the pay factor equation to estimate the combined pay factor. All combinations of PD values for the three quality characteristics are tried using @risk software to draw the probability density function of the combined pay factor for each approach. Percent defective are changed on a scale of 0 to 100. The

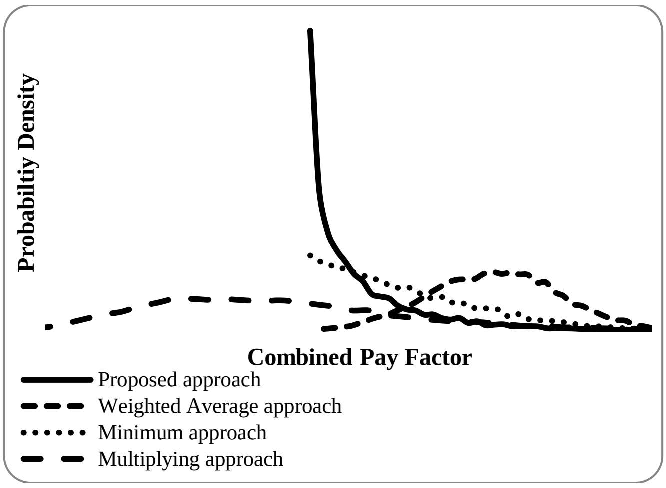

Fig. 5. Comparison between Combining Approaches.

probability density of the combined pay factor estimated using the four approaches are shown in Fig. 5. In this figure, it is shown that the distribution of the weighted average and multiplying approach are uniformly distributed over the range of PF. The average PF resulted from the weighted average and multiplying approaches are 80.3 and 65.1, respectively. The minimum and the proposed approaches are distributed heavily over the mid range of PF. The average PF resulted from the minimum and the proposed approaches are 78.7 and 78.9, respectively. This result illustrates that the proposed approach of estimating the combined PF provide the same result as the minimum approach, except the distribution is biased toward the mid range of PF. Fig. 5 shows that using the proposed approach result in combined PF greater than 55. Yet, only qualified contractors get high combined PF which is one of the major benefits of using the proposed approach.

Conclusions

Most states' public projects must be bid by the traditional low-bid method which results in many issues related to cost and time overruns. SHAs become more interested in adopting a rating system or alternative bidding process to overcome these problems. Utilizing current quality control and quality assurance data makes it easy to implement the proposed rating system, since the data exists and everyone is familiar with quality management systems. This study aims to develop a measure for pavement quality based on the results of quality tests. The Composite Percent Defective (PDC) is selected as the quality measure. The composite percent defective is proposed as the quality measure which combines percent defectives of a number of selected quality characteristics. The significance of the proposed measure is in its ability to reflect the true quality of the contractor work being calculated based on actual tests results. The proposed method has the advantage of using statistical concept to develop the PDC equation. The simulation verifies the results of the developed equation. The results of the case study show that CONB performed better than the other contractors in projects with SP4 mixes. On the other hand, COND is the best among all other contractors in projects with SPS mixes. The trend analysis of SP4 shows that the performance of CONA is improving while the others are deteriorating. It also shows the performance of both CONA and

COND is improving in SPS. The study recommends use of trend analysis, when possible, to study the performance of contractors over time, rather than only looking at the average value of the probability of acceptance.

References

1 Minchin, R.E., Hammons, M.I., Ahn, J., and Choubane, B. (2008). Highway construction quality index for Florida, *Transportation Research Record*, No. 2081, pp. 29-37.

2 Flyvbjerg, B., Holm, M., and Buhl, S. (2002). Underestimating Costs in Public Works Projects: Error or Lie?, *Journal of the American Planning Association*, 68(3), pp. 279-295.

3 Azhar, N., Farooqui, R., and Ahmed, S. (2008). Cost Overrun Factors In Construction Industry of Pakistan, *International Conference on Construction in Developing Countries*, Karachi, Pakistan.

4 NAVFAC (1996). Instructions for developing request or proposals for source selection design/build projects, *Specification Instruction 00011*, Naval Facilities Engineering Command: Southern Division, Rep. code 076 (RAM/SHP), Pensacola, Florida, USA.

5 Weed, R.M. (1998). A rational method for relating as-build quality to pavement performance and value, *Transportation Research Record*, No. 1632, pp. 32-39.

6 Weed, R.M. (2000). Method to model performance relationships and pay schedules, *Transportation Research Record*, No. 2081, pp. 117-124.

7 Monismith, C.L., Popescu, L., and Harvey, J. (2004). Performance-based pay factors for asphalt concrete construction: Comparison with a currently used experienced-based approach, *Journal of Association of Asphalt Paving Technologists*, No. 73, pp. 147-194.

8 Whiteley, L.S., Tighe, S.L., and Zhang, Z. (2005). Incorporating variability into pavement performance, life cycle cost analysis and performance-based specification pay factor, *Transportation Research Record,* No. 1940, pp. 13-20.

9 Cominsky, R.J., Killingsworth, B.M., Anderson, R.M., Anderson, D.A., and Crockford, W.W. (1998). Quality Control and Acceptance of Superpave-Designed Hot Mix Asphalt, *Report 409*, National Cooperative Highway Research Program, Washington, DC, USA.

10 Zaniewski, J.P. and Adams, T. (2005). *Evaluation of quality control parameters for Superpave hot mix asphalt*, Asphalt Technology Program, West Virginia University, Morgantown, West Virginia, USA.

11 Burati, J., Weed, R., Hughes, C., and Hill, H. (2003). Optimal procedures for quality assurance specifications, *Publication No. FHWA-RD-02-095*, Federal Highway Administration, Washington, DC, USA.

12 TRB (2005). Glossary of highway quality assurance terms, *Transportation Research Circular: No. EC074*, Transportation Research Board, Washington, DC, USA.

13 Bhushan, N. and Rai, K. (2004) *Strategic decision making: Applying the analytic hierarchy process*, Springer Publications, New York City, New York, USA.

14 Saaty, T. (1982). *Decision making for leaders: the analytic hierarchy process for decision in a complex world*, Lifetime Learning Publications, Belmont, California, USA.

15 Saaty, T. (1991). *Decision making with dependence and feedback: The analytic network process*, RWS Publications, Pittsburgh, Pennsylvania, USA.

16 Breakah, T.M., Williams, R.C., Kandil, A., and Shane, J.S. (2007). Implementing percent within limits for hot mix asphalt, *Mid-Continent Transportation Research Symposium*, Iowa State University, Ames, Iowa, USA.

17 Hubbard, D. (2007). How to measure anything: Finding the value of intangibles in business, John Wiley & Sons, Hoboken, New Jersey, USA.

18 Schram, S. and Abdelrahman, S. (2008). *Evaluation of Superpave construction quality on low volume roads*, *Journal of the Association of Asphalt Paving Technologists*, No. 77, pp. 299-326.

19 Nebraska Department of Roads (2005). *Nebraska Superpave Manual*, Nebraska Department of Roads HMA Design Manual, Nebraska Department of Roads, Lincoln, Nebraska, USA.