NEW MODELING TECHNIQUES
FOR PAVEMENT PREDICTIONS

Ying-Haur Lee[1] and Michael I. Darter[2]

1. Associate Professor of Civil Engineering, Tamkang University, Tamsui, Taipei, Taiwan 25137

2. Professor of Civil Engineering, University of Illinois, 205 N. Mathews Ave., Urbana, IL 61801

1 ABSTRACT

Statistical regression algorithms have been utilized extensively in pavement engineering for more than three decades. Multiple linear regression, stepwise regression, and nonlinear regression techniques are the most popular ones for pavement predictive modeling. The advantages, the deficiencies, and the limitations of these regression techniques are reviewed.

To minimize these problems, the projection pursuit regression (PPR) introduced by Friedman and Stuetzle (1981) was selected to assist in the proper selection of functional forms. Through the use of local smoothing techniques, the PPR attempts to model the response surface as a sum of nonparametric functions of projections of the explanatory variables. The projected terms are essentially two-dimensional curves which can be graphically represented, easily visualized, and properly formulated. As a result, a two-step predictive modeling approach is proposed and demonstrated in a case study for the prediction of edge stress of a pavement slab.

This paper also demonstrates that statistical regression techniques should not be used alone to obtain a more reliable and comprehensive predictive model. The importance of subject-related engineering knowledge, the principles of dimensional analysis, the proper selection of functional forms, and applicable engineering boundary conditions are considered essential and are also demonstrated in this paper. A comparison of the predictive models previously developed and the proposed approach to "prediction" and "extrapolation" clearly shows the preference of the proposed approach and the promising features of the PPR algorithm.

2 INTRODUCTION

Statistical regression techniques have been utilized extensively in predicting complicated pavement performance indicators for more than three decades. Multiple linear regression, stepwise regression, and nonlinear regression techniques are the most popular ones among transportation engineers. Undoubtedly, regression analysis is a very important statistical tool for many sciences. This paper starts with a review of the advantages and the limitations of these currently used regression techniques. Some specific problems in the selection of proper functional forms and the violation of some embedded statistical assumptions using these techniques are discussed.

Traditional parametric regression techniques such as linear and nonlinear regressions require imposing a parametric form on the functions and then obtaining the parameter estimates afterwards. In situations where little knowledge about the shape and the form of a function exists, several new nonparametric regression techniques developed over the past ten years have gradually gained popularity. Without imposing an unjustified parametric assumption, nonparametric regression techniques strive to estimate the actual functional form which best fits the data through the use of scatterplot smoothers [1].

Consequently, the projection pursuit regression (PPR) introduced by Friedman and Stuetzle (1981) [2, 3] is selected because of its ability to handle variable interactions. As a result, a two-step predictive modeling approach is proposed and demonstrated in a case study for the prediction of concrete pavement edge stress due to finite slab length effect.

The demonstration is based on a comparison between the previous modeling approaches from the Zero-Maintenance study [4] and a stress analysis procedure [5] and the proposed methodology. The importance of the use of the principles of dimensional analysis and the proper selection of functional forms are the main focus points. The demonstration also illustrates how to select a proper functional form through the use of PPR algorithm as well as the aid of graphical representation and visual inspection. Nonlinear regression is used to obtain the parameter estimates for the specified functional forms. The proper selection of initial parameter estimates are also discussed to guarantee convergence of the iterative nonlinear regression routines.

An illustration on the incorporation of subject-related engineering knowledge to the modeling process is also provided. The applications of the resulting predictive models to "prediction" and "extrapolation" are also discussed. Not only more comprehensive insights to the performance of the predictive model can be obtained by selecting more proper functional forms, but reasonable results may also be obtained even when extrapolation beyond the range of the data is required.

3 MULTIPLE REGRESSION

Multiple regression is one of the most time-honored and widely-used regression techniques for the study of linear relationships among a group of measurable variables. Suppose there exists a true model to describe the relationship between response variables (y's) and explanatory variables (or predictors, x's) [6, 7, 8]:

$$y_i = x_i^T \beta + \varepsilon_i \qquad i = 1,\ldots,n \qquad (1)$$

where x_i^T is the *i*th row of the (n x p) matrix X of the column of 1's if including an intercept and the explanatory variables. The superscript "T" denoting the transpose of the column vector x_i is required because of the usual convention that all vectors are represented by column vectors. β is a (p x 1) vector of unknown regression coefficients and p and n are the number of parameter estimates in the model and the total number of observations, respectively.

The basic assumptions are usually that the random errors (ε's) are mutually uncorrelated and normally distributed with zero mean and constant variance, and additive and independent of the expectation function. For any arbitrary β value of $\hat{\beta}$, the residuals $r_i(\beta)$ can be determined by the following expression:

$$r_i(\beta) = y_i - x_i^T \beta \qquad i = 1,\ldots,n \qquad (2)$$

Based on above assumptions, multiple regression tries to find a set of parameters $\hat{\beta}$ such that the sum of the squared residuals given in equation (3) is minimized, which is also best known as the least squares (LS) method.

$$RSS(\hat{\beta}) = \sum_{i=1}^{n} \left\{ r_i^2(\hat{\beta}) \right\} \qquad (3)$$

The LS method is best known and most frequently adopted due to its simple structure, elegant least squares theory, fast computation time, as well as its estimators which may be obtained explicitly from the data and are interpretable to its user.

4 ALL-SUBSET REGRESSION AND STEPWISE REGRESSION

As far as variable selection is concerned, all-subset regression and stepwise regression techniques are often utilized for automatic variable selection for preliminary and exploratory analysis of linear relationships among a group of important variables. The all-subset regression procedure finds all possible combinations of variables in the model when the set of candidate variables is not too large. In order to select the "best" subset of variables, different measurements including the coefficient of determination (R-squared, R^2), adjusted R-squared (adj-R^2), and the C_p statistics due to Mallows have been proposed [8, 9]. The Mallows C_p statistic may be is defined as follows:

$$C_p = \frac{RSS}{s^2} + 2p - n \qquad \textbf{(4)}$$

where RSS represents the residual sum of squares of the regression, s^2 is the estimate of the residual variance from the full regression model.

For an adequate model, the expected value of Mallows C_p statistic is approximately equal to the number of parameters (p) in the model. The smaller model has a smaller C_p value, but the C_p of the larger model is closer to p. According to Draper and Smith (1981), a regression model with "**a low C_p value about equal to p**" is preferred such that the model fits the actual data better with least number of parameters. [9 (pp. 299-300), 8 (pp. 216-217)]

All-subset procedure produces optimal subset models in terms of R^2 for the given data sample, but they may not be necessarily optimal for the whole population. Due to the high computational cost of all-subset regression, stepwise regression has been used as the most popular technique in performing variable selection regardless of obtaining a sub-optimal set of the combination of candidate variables.

Stepwise regression uses a forward selection or a backward elimination procedure which iteratively adds or deletes one explanatory variable at a time to find the "best" subset of significant independent variables in the model. In forward selection, the procedure starts with a simple regression model containing the explanatory variable, x_i, which best correlates with the response variable, y. Subsequently, the procedure selects another variable which produces the largest F-value among the candidate variables to be included in the model. A pre-specified F-value is often used as the stopping rule if the F-value of the next iteration is less than this value. On the other hand, the backward elimination procedure first starts with a model containing all the explanatory variables. The variable which contributes the least to the current model is subsequently eliminated by the procedure. Similar stopping rules are also used in this procedure [10, pp. 107-108].

It should be noted that the stepwise procedure does not necessarily produce the "best" subset among a group of candidate variables. The first variable added or deleted may not be necessarily the best or the worst overall. It is quite possible that the first variable selected by forward selection becomes unnecessary in the presence of the other variables entered afterwards or vice versa. Therefore, a combination of forward selection and backward elimination may also be supplied as an option for variable selection in a commercial statistical package. However, this still doesn't guarantee an optimal subset of candidate variables. In summary, forward selection and backward elimination techniques may result in totally different subsets of variables with similar summary statistics. "It is unlikely that there is a single best subset, but rather several equally good ones," as observed by Gorman and Toman (1966) [11, 10, p.108].

Stepwise regression is one of the most popular techniques used in transportation research due to its fast computation, simplicity for use, and capability in progressively adding or dropping variables in a regression model. It is expected that some latent or lurking variables may be included in a regression model. Due to the ease of use and the efficiency of computation, stepwise regression may be easily abused in some respects as well if one fails to recognize possible physical interpretations on the parameter estimates and the embedded basic assumptions previously discussed. A very long regression model with dozens of parameter estimates may be the typical result of a stepwise regression which fits the data very well but with no comprehensive engineering insight to the data analyzed. Often times, some applicable engineering boundary conditions are violated as well.

5 NONLINEAR REGRESSION

Practical real world problems are often found to be nonlinear in nature. Due to its favorable feature of handling a complicated nonlinear model, nonlinear regression has been widely used as a modeling technique for pavement research. It is also because that some applicable boundary conditions may be incorporated into the specified nonlinear model form during the modeling process. Besides, the parameter estimates may be physically interpreted and may have their own physical meanings.

However, nonlinear models are more difficult to specify and develop than linear models. Suppose there exists a true model which best describes the relationship between response variables (y's) and explanatory variables (x's): [12, 7]

$$y_i = F(\beta, x_i) + \varepsilon_i \qquad i = 1,\ldots,n \qquad \textbf{(5)}$$

where $F(\beta, x_i)$ is a nonlinear function based on the predictors x_i's. β is a (p x 1) vector of unknown regression coefficients to be estimated and n is the total number of observations. Similar to linear regressions, the disturbance (or error) term is usually assumed to be additive, mutually uncorrelated and normally distributed with zero mean and constant variance. For any arbitrary β value of $\hat{\beta}$, the residuals $r_i(\beta)$ are:

Unlike linear regressions whose parameters can be explicitly estimated by a closed-form expression, nonlinear regressions must use an iterative routine to find the best parameter estimates (β) such that the sum of the squared residuals as given in equation (3) is minimized.

Before carrying out a nonlinear regression, engineers will first have to assume a feasible descriptive (or conceptual) model form with the most important parameters pre-specified based on previous experience and engineering knowledge. In a two-dimensional factor space, the model form can be easily determined as it can be visualized in a scatter plot. Subsequently, reliable parameter estimates may be obtained. The situation is quite different for the case of multi-dimensional (three or higher) factor space where the data trend and the possible interactions among candidate variables can not be easily identified. Unfortunately, real engineering problems most likely belong to the latter case where higher dimensions are concerned and the true functional forms are rarely known to engineers. Thus, the proper selection of a plausible functional form from a multi-dimensional factor space provides a real challenge for pavement engineers.

Guessing initial parameter estimates for a multi-dimensional model is also a very time-consuming process, even though a grid of values may be specified. It is not uncommon that the resulting model does not converge and/or contains insignificant parameter estimates with poor initial parameter estimates and an improper model form. In addition, parameter estimates are sometimes toward the wrong direction of their physical interpretations and they may heavily depend on the initial starting values as well. Specifying bounds for some parameters are not uncommon, though the final resulting model may eventually reach the assumed bounds but without achieving the specified convergence criterion. However, it should also be kept in mind that "*Some models are difficult to fit, and there is no guarantee that the procedure will be able to fit the model successfully* [7, p. 576]."

Furthermore, even when parameter estimates are successfully obtained, the assumption of constant error variance is frequently violated as indicated by residual analyses. This problem is often recognized, probably due to the improper selection of a model form though there is no quick remedy for this situation. It is very dangerous to develop such a nonlinear model without acquiring more knowledge of the complicated relationships between the predictors and the dependent variable. Thus, the prediction based on a inadequate model specification is very questionable.

Traditional parametric regression techniques such as linear and nonlinear regressions require imposing a parametric form on the functions and then obtaining the parameter estimates afterwards. In situations where little knowledge about the shape and the form of a function exists, several new nonparametric regression techniques developed over the past ten years have gradually gained popularity. Without imposing an unjustified parametric assumption, nonparametric regression techniques strive to estimate the actual functional form which best fits the data through the use of scatterplot smoothers [1].

The alternating conditional expectations (ACE) algorithm developed by Breiman and Friedman (1985) for optimal transformation of multiple regression and correlation does an excellent job in maximizing the squared multiple correlation, R^2 [1, 13, 14]. By using the same alternative backfitting procedures for variable transformations as used in the ACE algorithm, Tibshirani (1987) introduced the Additivity and Variance Stabilization (AVAS) algorithm which tries to achieve the constant error variance assumption of regression [1, 15]. Both techniques provide an invaluable tool in suggesting proper transformations of variables. However, neither the ACE nor the AVAS is capable of modeling interactions between the explanatory variables. The variable interactions must be explicitly specified in the assumed model form.

Unlike the ACE and the AVAS algorithms, the projection pursuit regression (PPR) introduced by Friedman and Stuetzle [2, 3] is capable of modeling interactions between predictor variables when suggesting nonparametric transformations to improve the fit. This algorithm is an exploratory nonlinear regression technique which tries to model the response surface (y's) as a sum of nonparametric functions of projections of the predictor variables (x's) through the use of super smoothers [2, 3]. Assuming there exists a true model given as follows [16, Vol. 2]:

$$y = \bar{y} + \sum_{m=1}^{M_0} \beta_m \phi_m (a_m^T x) + \varepsilon$$

where $x = (x_1, x_2, ..., x_p)^T$ denotes the vector of predictor variables and ÿ is the expected value (or mean) of response variable. The "projection" part of the "projection pursuit regression" indicates that the vector of predictors, x, is projected onto the direction vectors, a_m, to get the lengths of the projection $a_m^T x$, where m = 1, ..., M_0. The "pursuit" part indicates that an optimization technique is used to find the best direction vectors, a_m. $\phi_m(a_m^T x)$ stands for the unknown nonparametric transformation functions of the projected lengths $a_m^T x$ to be estimated. ε is the random error.

Notice that ϕ_m have been standardized to have zero mean and unity variance given in the following expression:

$$E[\phi_m(a_m^T x)] = 0, \quad E[\phi_m^2(a_m^T x)] = 1, \quad m = 1,\ldots,M_0 \qquad \textbf{(8)}$$

The PPR algorithm strives to minimize the mean squared residuals, $E[r^2]$, over all possible combinations of β_m, ϕ_m, and a_m values. This algorithm first performs a forward stepwise fitting procedure starting with a single term and ending up with M_{max} terms. Supersmoother is used to estimate the conditional expectation functions. After having fitted the M_{max}-term model in a forward stepwise manner, a backward stepwise procedure is followed and stopped with an M_{min}-term model. Notice that M_{max} and M_{min} are the user-specified maximum and minimum number of terms in the model, respectively. The relative importance of each term can be measured by the absolute value of β_m as well. The expectation function of the squared residuals is given as follows:

$$E[r^2] = E\left[y - \bar{y} - \sum_{m=1}^{M_0} \beta_m \phi_m(a_m^T x)\right]^2 \qquad \textbf{(9)}$$

Conceptually, the explanatory variables x's are projected onto the direction vectors a_1, a_2, ..., a_m, to find good nonlinear transformations ϕ_1, ϕ_2, ..., ϕ_m for the multi-dimensional response surface. In other words, the response surface is broken down into a series of smooth projected terms adding all together. Thus, the largest trend in the factor space is captured as the first projected term ϕ_1 and the residuals become the sources for additional projections afterwards. More technical details about the development process, the application, and the demonstration on modeling interactions of the PPR algorithm can be found in References [2, 3, 16].

7 PROPOSED TWO-STEP MODELING APPROACH

The proper selection of regression techniques is one of the most important factors to the success of prediction modeling. Since most of the regression algorithms currently available do not directly consider interaction effects during the modeling process, the interaction terms must be subjectively determined prior to performing a regression analysis. With the multi-dimensional pavement engineering problems in mind, several untangled deficiencies are frequently identified in the use of stepwise regression and nonlinear regression. These include problems in the selection of correct functional form, violations of the embedded statistical assumptions, and failure to satisfy some engineering boundary conditions as previously discussed.

The projection pursuit regression (PPR), however, appears to have the most favorable features in handling these problems. As a result, a two-step regression analysis procedure is proposed herein to better find the correct functional form and to better fit the response surface as well:

1. With the help of the PPR, a multi-dimensional response surface is broken down into the sum of several smooth projected curves which are graphically representable in two dimensions. Plausible functional forms and applicable boundary conditions may then be easily identified and specified through visual inspection and/or engineering knowledge of physical relationships to model these individual projected curves separately.
2. Traditional parametric regression techniques such as linear and nonlinear regressions are then utilized for these purposes with higher confidence in the parameter estimates.

The overall regression statistics and the goodness of fit often clearly show the preference of the proposed two-step modeling approach against the traditional counterparts.

A new statistical package named S-PLUS which has been widely used by statisticians for data analysis [16, 17, 18, 19] was selected due to the availability of these new nonparametric regression techniques. S-PLUS is very strong in its graphics, data exploration tools, and flexibility but weak in database management as compared to the most well-known and widely used statistical package, SAS. The strengths and weaknesses of S-PLUS versus SAS have been briefly discussed by Therneau [20].

8 EDGE STRESS DUE TO LOADING AND FINITE SLAB LENGTH EFFECTS: A CASE STUDY

Since transverse cracking is one of the major structural distresses in jointed concrete pavements, the determination of the maximum bending stress at the longitudinal edge (midway between the transverse joints) is very crucial in any mechanistic or mechanistic-empirical design procedures.

However, the theoretical Westergaard edge stress solution [21] based on infinite slab assumption may not be well-suited for the actual finite slab size conditions. To access the effect of a finite slab length, a series of finite element (F.E.) runs are often performed. Unfortunately, F.E. model can not be easily implemented as a part of a design procedure due to the required running time.

Several attempts in edge stress prediction which proved to be very successful in the past were reviewed and discussed below. The basic concepts behind those analysis procedures were to design and perform a full factorial of F.E. runs under various combinations of different conditions. Statistical regression techniques were utilized to develop predictive models for different situations. The predictive equations were then used as alternatives for design purposes with sufficient accuracy in estimating stresses.

To demonstrate the proposed two-step predictive modeling procedures, the following case study was performed. The purpose of this case study was to demonstrate the advantages of incorporating dimensional analysis and selecting proper functional forms for the predictive models analyzed. Three different approaches are discussed in the following sections:

1. Use arbitrary but "best" linear combinations of individual variables with interaction terms for regression;
2. Introduce as many mechanistic variables as possible based on engineering knowledge, and also find the "best" linear combinations of these variables with interaction terms for regression; and
3. Introduce as many mechanistic variables (or clusters of variables) as before, and also try to find the best functional forms based on the proposed two-step predictive modeling procedures.

__8.1 Principles of Dimensional Analysis__

According to Westergaard [21, 22], the closed form stress solution due to a circular edge loading on a semi-infinite slab over a dense liquid foundation based on medium-thick plate theory is given as follows:

$$\sigma = \frac{3(1+\mu)P}{\pi(3+\mu)h^2}\left[\log_e\frac{Eh^3}{100ka^4} + 1.84 - \frac{4\mu}{3} + \frac{1-\mu}{2} + 1.18(1+2\mu)\frac{a}{l}\right] \tag{10}$$

Where:

- σ = theoretical Westergaard edge stress, [FL^{-2}];
- P = total applied wheel load, [F];
- a = radius of the applied circular load, [L];
- E = elastic modulus of the concrete slab, [FL^{-2}];
- h = slab thickness, [L];
- μ = slab Poisson's ratio;
- k = modulus of subgrade reaction, [FL^{-3}]; and
- l = radius of relative stiffness of the slab subgrade system, [L], given by

$$l = \sqrt[4]{\frac{Eh^3}{12(1-\mu^2)k}} \tag{11}$$

Note that the basic dimensions are abbreviated as length [L] and force [F].

Through the use of the principles of dimensional analysis, the above stress equation may be reduced to the following equation form in which both sides of the equation are dimensionless [23, 24].

$$\frac{\sigma h^2}{P} = f\left(\frac{a}{l}\right) \tag{12}$$

Therefore, the above normalized stress equation ($\sigma h^2/P$) is simply a function of the dominating variable, a/l, representing the combination effect of the input parameters, namely E, h, k, and a, for constant μ (usually $\mu \approx 0.15$).

In order to account for finite slab length effect, Ioannides *et al.* [22] introduced a normalized length term, L/l, when comparing F.E. results to the theoretical Westergaard solution. Similarly, a normalized width term, W/l, was introduced to account for the effect of finite slab width. Note that L and W represent slab length and slab width, respectively. However, it is also recognized that the finite slab width effect (W/l) is not as significant as the other two dimensionless parameters (i.e., a/l and L/l), as suggested by earlier investigators [24, 5].

An adjustment (or multiplication) factor may be introduced to account for the theoretical difference between the Westergaard solution (σ) and the results from the F.E. model using only these two dimensionless factors (a/l and L/l). The adjustment factor (R) is defined by the following expression:

$$R = \frac{\sigma_i}{\sigma} = f\left(\frac{a}{l}, \frac{L}{l}\right) \tag{13}$$

where σ_i is the edge stress determined by the F.E. model.

The advantages of using the principles of dimensional analysis to derive the above formulation are evident. They can be very helpful in identifying the governing dependent and independent variables (or clustered variables) of a given relationship. In addition, they may also help to reduce the number of factorial F.E. runs needed for the analysis. The above formulation is not only easy to comprehend but also dimensionally correct. Thus, the resulting edge stress from F.E. studies (σ_i) can be determined by multiplying the theoretical Westergaard solution and the adjustment factor (R) together, if it is properly formulated.

__8.2 Using Arbitrary Linear Combination of Variables__

The analysis of edge stress due to different axle loads, slab lengths, slab thicknesses, foundation supports, thermal gradients, and edge supports was initially studied by Darter [4]. A large full factorial of finite element (F.E.) runs was performed to determine the edge stresses for different conditions and subsequently separate predictive models were developed. For the case of loading only, an 18-kip single axle loading was applied on the following factorial pavement sections:

1. Slab length, L = 15, 20, 25, 30 ft;
2. Slab thickness, h = 8, 10, 14 in.;
3. Foundation support, k = 50, 200, 500 pci; and
4. Edge support, ES = 0, 12, 36, 60 in.

where other pertinent input parameters are: $E = 5 * 10^6$ psi, W = 12 ft, $\mu = 0.15$, and the loaded area per wheel = 12 x 15 in.^{2}. The results of these F.E. runs were directly used to derive a regression model for edge stress (σ_L) prediction.

Stepwise regression procedure was utilized to pick the "best" linear combinations of individual variables including interaction terms during the modeling process. The slab length (L) was not proved to affect the edge stress prediction significantly, and thus was excluded from the model. The proposed predictive model for single axle load without the loss of edge support (ES = 0) is given as follows:

$$\sigma_L = \frac{load}{18\ h^2}\left(17.358 - 0.054\frac{h^3}{k} + 7.417\ \log_{10}\frac{h^3}{k}\right) \quad \textbf{(14)}$$

Using Equation (14) to predict the actual calculated edge stresses from the above F.E. runs, a fairly good agreement is observed.

However, it is also noticed that this model does not contain the elastic modulus of the concrete slab (E) and the finite slab length (L) due to the limitations of the factorial runs conducted. The variables included in this model are numerically significant, but they are not the actual dominating factors with correct dimensions. Thus, the applicability of using this equation beyond the ranges of the data where the model was originally developed such as different E values and other shorter slabs is very questionable.

8.3 Introducing Mechanistic Variables (Clustered Variables)

The analysis of edge stress was further studied by Salsilli-Murua [5]. The concept of equivalent single axle radius (ESAR) was utilized for analyzing different load configurations including duel wheels, tandem gears, and tridem gears. The effects of slab sizes, widened outer lanes or tied shoulders, subbase and subgrade layers, and thermal curling were also analyzed separately. Through the use of the principles of dimensional analysis, a series of small factorial F.E. runs covering most of practical engineering conditions were performed and separate predictive models were developed.

As previously discussed, the effects of wheel loading and finite slab length can be best represented by two dimensionless factors: a/l and L/l, respectively. Thus, only a minimum number of factorial F.E. runs based these two factors were performed. The adjustment (or multiplication) factor (R) as defined in equation (13) was also introduced for edge stress prediction. Note that the Westergaard solution (σ) for infinite slab condition is approximated by the resulting ILLI-SLAB edge stress (σ_∞) for L/l = 7.0 herein when calculating the adjustment factor ($R = \sigma_i/\sigma_\infty$).

Multiple stepwise regression procedure was utilized again because of its favorable feature of progressively adding and dropping variables in a regression model. The proposed predictive model is given as follows:

$$R = 0.582 - 0.533\left(\frac{a}{l}\right) + 0.182\left(\frac{L}{l}\right) - 0.0198\left(\frac{L}{l}\right)^2 + 0.109\left(\frac{a}{l}\right)\left(\frac{L}{l}\right) \quad \textbf{(15)}$$

Statistics: N = 12, R^2 = 0.996, SEE = 0.0028, CV = 0.29%
Limits: $3 \le L/l \le 5$, $0.05 \le a/l \le 0.3$

(Note: N = number of data points, R^2 = coefficient of determination, SEE = standard error of estimates, CV = coefficient of variation)

His work has repeatedly demonstrated the advantages of using dimensional analysis to discover the underlying dominating factors for the analysis. In addition, the proposed predictive model not only has fairly good agreements with the data analyzed, but it may also be applied to a wide range of input parameters (i.e., E, h, k, and *a*). This model is also dimensionally correct. The prediction of edge stress is just simply a matter of multiplying the theoretical Westergaard solution by the predicted R value based on this model.

Nevertheless, the conclusions should be restricted to the ranges of the data, where the model was originally developed. This is also true for any regression technique chosen to develop such a predictive model. Besides, a reasonable doubt is raised as to the adequacy of the functional form, especially when an asymptotic trend of the relationship between R and L/l is observed. This asymptotic trend could not be captured simply because of the multiple stepwise regression technique chosen. Thus, the usefulness of the model beyond the specified limits is highly questionable.

8.4 Selecting Proper Functional Forms

The importance of incorporating subject-related engineering knowledge into the modeling process, as has been done in the previous example in identifying dominating clustered variables, cannot be overemphasized. In addition, the selection of proper functional forms to satisfy some applicable engineering boundary conditions is also a crucial component to a successful predictive model. The proposed two-step modeling approach was adopted herein to illustrate this point. The following factorial ILLI-SLAB runs were performed:

a/l : 0.05, 0.1, 0.2, 0.3

L/*l* : 2.0, 2.5, 3.0, 3.5, 4.0, 4.5, 5.0, 6.0, 7.0

Note that W/*l* was set to 7.0 for all runs. The finite element grids were generated according to the recommendations of Ioannides [25, pp. 187-188] to provide sufficient accuracy of the results. The pertinent input parameters and the results of these factorial runs are presented in Table 1. A three-dimensional perspective plot providing a very clear picture of the relationship among R, *a/l*, and L/*l* is shown in Figure 1.

Knowing that the relationship is a nonlinear three-dimensional surface, a nonlinear regression model may be better suited for this prediction purpose than an arbitrarily selected linear or polynomial model form. But, what kinds of functional forms to be used in the model and how to choose a proper one are still remained unanswered.

The proposed two-step modeling approach is aimed at searching for a series of two-dimensional projected curves first to ease these important modeling decisions. First, the PPR technique was utilized to determine the direction vectors for *a/l* and L/*l* and the suggested single-term nonlinear transformation $\phi_1(a_1^T x)$ using only those data points within the specified limits of equation (15). The results of this PPR trial are given in the following equations:

$$R = 0.9667 + 0.0330\,\Phi_1(ATX1) \quad \textbf{(16)}$$

$$ATX1 = -0.8945\frac{a}{l} + 0.4471\frac{L}{l} \quad \textbf{(17)}$$

A graphical representation of this PPR trial is shown in Figure 2. The first two plots are the scatter plots of the one-to-one relationships of R versus *a/l* and L/*l*. The next plot shows the resulting first PPR projected term versus the length of the projection, i.e., $\phi_1(a_1^T x)$ versus $a_1^T x$. Notice that $a_1^T x$ is represented by the axis label "ATX1" whereas ϕ_1(ATX1) is represented by the axis label "1st Projected Term" in this plot. The goodness-of-fit of this PPR model including a plot of the predicted versus actual values and a plot of the residual versus predicted values is displayed in the last two plots. Under the condition that this curve is properly modeled together with the use of the above two equations, fairly good predictions with a relatively high R^2 of 0.995 may be obtained.

Thus, the task was reduced to simply find a feasible model for this projected curve. An asymptotic trend was clearly identified through visual inspection. Therefore, an asymptotic nonlinear model with two to three (or at most four) parameters may describe this curve adequately. The following four-parameter equation form was selected to offer the maximum flexibility in the modeling process:

$$\Phi_1(ATX1) = a_1 + \frac{1}{a_2 + a_3 ATX1^{a_4}} \quad \textbf{(18)}$$

The nonlinear regression technique was then used to obtain the parameter estimates. The first derivative of this equation form with respect to each parameter was also provided to help the iterative regression routine converge faster. The starting value of each parameter was carefully chosen in a way that a_1 and a_4 were large negative values, e.g., -4.0, and a_1 was the minimum possible value of ϕ_1. Since a_2 and a_3 were still undecided, two small positive values, e.g., 0.01, were chosen as their starting points. The following equation satisfies the convergence criteria even with many other different starting values:

$$\Phi_1(ATX1) = -5.587 + \frac{1}{0.147 + 0.263\,ATX1^{-5.166}} \quad \textbf{(19)}$$

The overall statistics for the prediction of R (or σ_i/σ_∞) due to slab length effect using equations (16), (17) and (19) are summarized as follows:

Statistics: N = 20, R^2 = 0.994, SEE = 0.0027, CV = 0.28%

Limits: $3 \le L/l \le 5$, $0.05 \le a/l \le 0.3$

To illustrate how the three-dimensional response surface (shown in Figure 1) might be broken down into the sum of a series of two-dimensional curves, the PPR algorithm was utilized again to obtain a two-term model and a three-model as shown in Figure 3 and Figure 4, respectively. By analogy, the axis labels "ATX2" and "ATX3" represent $a_2^T x$ and $a_3^T x$ whereas the axis labels "2nd Projected Term" and "3rd Projected Term" represent ϕ_2(ATX2) and ϕ_3(ATX3) in equation (7), respectively. The other plots are similar to those plots defined in Figure 2.

As an example, the results of the two-term PPR model are given in the following equations:

$$R = 0.9667 + 0.033\,\Phi_1(ATX1) + 0.0021\,\Phi_2(ATX2) \quad \textbf{(20)}$$

$$ATX1 = -0.8945\frac{a}{l} + 0.4471\frac{L}{l} \quad \textbf{(21)}$$

$$ATX2 = 0.9970\frac{a}{l} + 0.07787\frac{L}{l} \quad \textbf{(22)}$$

The second projected term contributes little to the prediction of R, since its coefficient is very small when compared to the coefficient of the first projected term (i.e. $\beta_2 = 0.0021$ versus $\beta_1 = 0.0330$). The third projected term also contributes very little to the accuracy of the prediction when compared to the first one and the second one, and thus the associated equations for the three-term model are not presented herein. In fact, using only one or two projected curves can adequately describe this three-dimensional response surface as suggested by these figures.

The plausible model forms suggested by these PPR trials are: an asymptotic nonlinear model for the first projected curve; a second-degree polynomial for the second projected curve; and a third-degree polynomial for the third projected curve if necessary. Of course, some piece-wise linear and nonlinear curves may also be formulated for these purposes without much loss of generality.

8.5 Prediction and Extrapolation

From the above demonstration, it is evident that using different regression procedures may result in totally different predictive models but with excellent agreements with the same set of data. Under the umbrella of statistical regression analysis, they are all feasible solutions to the targeted problem as long as the conclusions are restricted to the range of the data. In other words, "extrapolation beyond the range of the data" is never considered feasible with statistical regression analysis procedures.

However, it is also recognized that "extrapolation" is often necessary in actual engineering practice due to limited resources available and technical limitations. To judge which equation is more preferable, both equations (15) and (16) were checked by plotting the predicted versus the actual R values within the specified data ranges as displayed by the symbol "o" in Figure 5 and Figure 6, respectively. It is clearly shown that both equations result in very good predictions within the range of the data from which they were originally developed. In addition, Figure 5 and Figure 6 also present the results when extrapolating beyond the specified limits using these two models and the data from all of the 36 F.E. runs, respectively. These data points as displayed by the symbol "∗" in the graphs have clearly shown the difference.

Extrapolation using equation (15) results in totally unacceptable predictions outside the specified limits as shown in Figure 5. However, this problem is less pronounced as shown in Figure 6 when using equation (16). In fact, the extrapolation of L/l greater than 5.0 provides excellent agreements with the actual data, since the functional form was properly selected to be asymptotic such that the predicted R values are very close to unity as they should be. Nevertheless, some discrepancies still exist when extrapolating equation (16) to smaller L/l values, i.e., $L/l < 3.0$, which generally result in smaller R values.

This problem may be explained by the fact that the first parameter (a_1, or the minimum possible value of ϕ_1(ATX1)) of equation (19) was estimated to be -5.5869 such that equation (16) will never result in a value below 0.7823. This approximation is acceptable for the ranges of data which were used to develop equations (16), (17) and (19). However, it may not be as accurate when extrapolating too far away from this specified range. Thus, the bad part of the extrapolation as shown in Figure 6 is not surprising. If the minimum possible value of R is known based on engineering knowledge, the boundary condition may also be imposed onto the predictive model. Thus, better agreements with the data can be obtained even in the case of extrapolation.

In summary, this case study not only demonstrates the benefits of using mechanistic variables through the use of the principles of dimensional analysis, but it also emphasizes the importance of selecting proper functional forms. Correct functional forms provide more comprehensive insights to the performance of the predictive model and may also lead to reasonable results even when extrapolation beyond the range of the data is required. In other words, it can be concluded that proper functional form is also a very crucial component to the success of a modeling process. The proper selection of functional forms to satisfy some applicable physical boundary conditions as well as subject-related engineering knowledge are the best supplements to statistical regression algorithms.

9 CONCLUSIONS

Several traditional linear and nonlinear regression techniques which have been widely adopted in pavement research were reviewed. Their advantages and limitations were discussed. Applying these regression techniques to practical engineering problems often shows the deficiencies in the selection of proper functional forms, the violation of some embedded statistical assumptions, and failures to satisfy some applicable engineering boundary conditions.

Through the use of projection pursuit regression (PPR) technique, a two-step predictive modeling approach was proposed in an attempt to minimize these problems. The proposed modeling approach was demonstrated in a case study for edge stress prediction due to finite slab length effect. Without imposing an unjustified parametric assumption, the PPR provides a unique routine for breaking higher dimensional engineering problems down into a series of sensible projected curves. By doing so, the individual projected curves may be formulated separately with better choices of feasible functional forms as well as possible boundary conditions. Not only did the demonstration re-emphasize the advantages of using the principles of dimensional analysis, but it also illustrated the importance of the proper selection of functional forms. Correct functional form provides more comprehensive insights to the performance of the predictive equation and may also leads to reasonable results even when extrapolation beyond the range of the data should be required.

The demonstration was based on the data from the theoretical finite element solutions which no erroneous data is expected. As for applying the PPR to field or laboratory data with some possible errors, the promising features of the PPR algorithm is also helpful to identify any extremely strange behavior due to data errors. Engineering knowledge on the physical mechanism of the problem may be applied in identifying the possible bad data points as well.

It is also believed that only through the joint efforts of both engineering and statistical knowledge and techniques, will a more theoretically correct predictive equation be developed. If properly developed, it may also be applied beyond the range of the data. Nevertheless, the conclusions should be restricted to the ranges of the data.

10 ACKNOWLEDGMENTS

This research work was partially sponsored by the Illinois Department of Transportation under the project titled "Implementation of Illinois Pavement Feedback System" for which different modern regression techniques were utilized for predictive model development. The authors would like to express much appreciation to Mr. James Hall, Illinois Department of Transportation for his assistance in this project.

11 REFERENCES

1. Hastie, T. J. and R. J. Tibshirani, Generalized Additive Models, Monographs on Statistics and Applied Probability 43, Chapman and Hall, 1990.
2. Friedman, J. H. and W. Stuetzle, "Projection Pursuit Regression," Journal of the American Statistical Association, Vol. 76, pp. 817-823, 1981.
3. Friedman, J. H., "SMART User's Guide," Technical Report No. 1, LCS, Department of Statistics, Stanford University, Stanford, CA, 1984.
4. Darter, M. I., "Design of Zero-Maintenance Plain Jointed Concrete Pavement: Vol. I - Development of Design Procedures," Federal Highway Administration Report No. FHWA-RD-77-111, Washington, D.C., June 1977.
5. Salsilli-Murua, R. A., "Calibrated Mechanistic Design Procedure For Jointed Plain Concrete Pavements," Ph.D. Thesis, University of Illinois, Urbana, 1991.
6. SAS Institute, Inc., SAS User's Guide: Basics, Version 5 Edition, 1985.
7. SAS Institute, Inc., SAS User's Guide: Statistics, Version 5 Edition, 1985.
8. Weisberg, S., Applied Linear Regression, Second Edition, Weiley Series in Probability and Mathematical Statistics, John Wiley & Sons, 1985.
9. Draper, N. R. and H. Smith, Applied Regression Analysis, Second Edition, Weiley Series in Probability and Mathematical Statistics, John Wiley & Sons, Inc., 1981.
10. Rousseeuw, P. J. and A. M. Leroy, Robust Regression and Outlier Detection, Wiley Series in Probability and Mathematical Statistics, John Wiley & Sons, Inc., New York, 1987.
11. Gorman, J. W. and R. J. Toman, "Selection of Variables for Fitting Equations to Data," Technometrics, 8, pp. 27-51, 1966.
12. Bates, D. M. and D. G. Watts, Nonlinear Regression Analysis and Its Applications, Wiley Series in Probability and Mathematical Statistics, John Wiley & Sons, Inc., New York, 1988.
13. Breiman, L. and J. H. Friedman, "Estimating Optimal Transformations For Multiple Regression And Correlation, (With Discussion)", Journal of the American Statistical Association, Vol. 80, pp. 580-619, 1985.
14. Buja, A., "Remarks on Functional Canonical Variates, Alternating Least Squares Methods and ACE," Annals of Statistics, Vol. 18, No. 3, pp. 1032-1069, 1990.
15. Tibshirani, R., "Estimating Transformations for Regression via Additivity and Variance Stabilization," Journal of the American Statistical Association, Vol. 83, pp. 394-405, 1987.
16. Statistical Sciences, Inc., "S-PLUS User's Manual", Vol. 1 and Vol. 2, Seattle, Washington, November 1991.
17. Statistical Sciences, Inc., "S-PLUS Reference Manual", Seattle, Washington, November 1991.
18. Chambers, J. M. and T. J. Hastie, Statistical Models in S, Wadsworth & Brooks/Cole Computer Science Series, AT&T Bell Laboratories, 1992.
19. Becker, R. A., J. M. Chambers, and A. R. Wilks, The New S Language - A Programming Environment for Data Analysis and Graphics, The Wadsworth & Brooks/Cold Computer Science Series, AT&T Bell Laboratories, 1988.
20. Therneau, T. M., "S-PLUS", The American Statistician, Vol. 44, No. 3, pp. 239-241, 1990.
21. Westergaard, H. M., "New Formulas for Stresses in Concrete Pavements of Airfields," American Society of Civil Engineers (ASCE), Transactions, Vol. 113, 1948, pp. 425-444. Also in ASCE Proceedings, Vol. 73, No. 5, May, 1947.
22. Ioannides, A. M., M. R. Thompson, E. J. Barenberg, "The Westergaard Solutions Reconsidered," Transportation Research Record 1043, Transportation Research Board, 1985.
23. Tayabji, S. D. and D. J. Halpenny, "Thickness Design of Roller-Compacted Concrete Pavements," with Discussion by Ioannides, A. M., Transportation Research Record 1136, Transportation Research Board, National

Research Council, Washington, D.C., 1987.

24. Ioannides, A. M., and R. A. Salsilli-Murua, "Temperature Curling in Rigid Pavements: An Application of Dimensional Analysis," Transportation Research Record 1227, 1989.

25. Ioannides, A. M., "Analysis of Slabs-on-Grade for a variety of Loading and support Conditions," Ph.D. Thesis, University of Illinois, Urbana, 1984.

Table 1 - ILLI-SLAB Runs on Slab Size Effects

a/l	L/l	c	a	P	l	L	W	E	k	h	σ
0.05	2.0	2.5	1.995	1250	39.89	79.79	279.26	5	200	10.59	40.868
0.05	2.5	2.5	1.995	1250	39.89	99.74	279.26	5	200	10.59	43.559
0.05	3.0	2.5	1.995	1250	39.89	119.68	279.26	5	200	10.59	45.458
0.05	3.5	2.5	1.995	1250	39.89	139.63	279.26	5	200	10.59	46.744
0.05	4.0	2.5	1.995	1250	39.89	159.58	279.26	5	200	10.59	47.536
0.05	4.5	2.5	1.995	1250	39.89	179.52	279.26	5	200	10.59	47.973

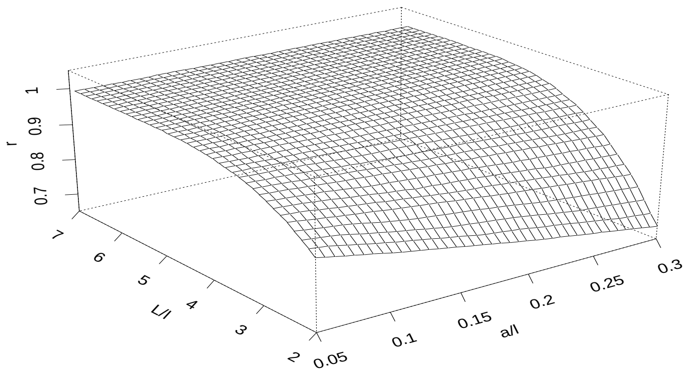

Figure 1 - Slab Size Effect, 3-D plot

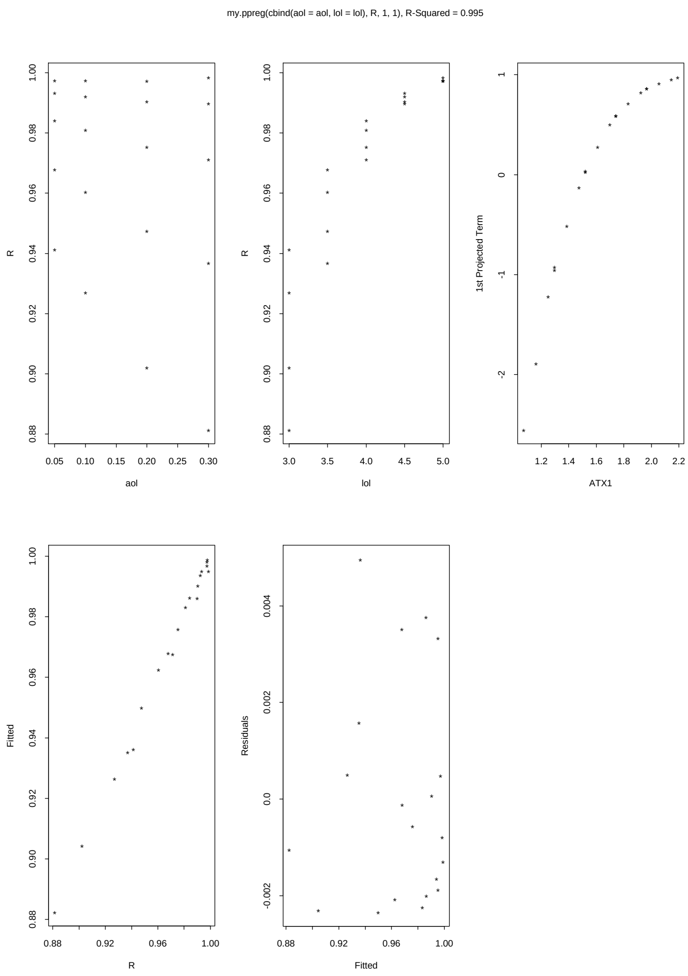

Figure 2 - One-Term PPR Model

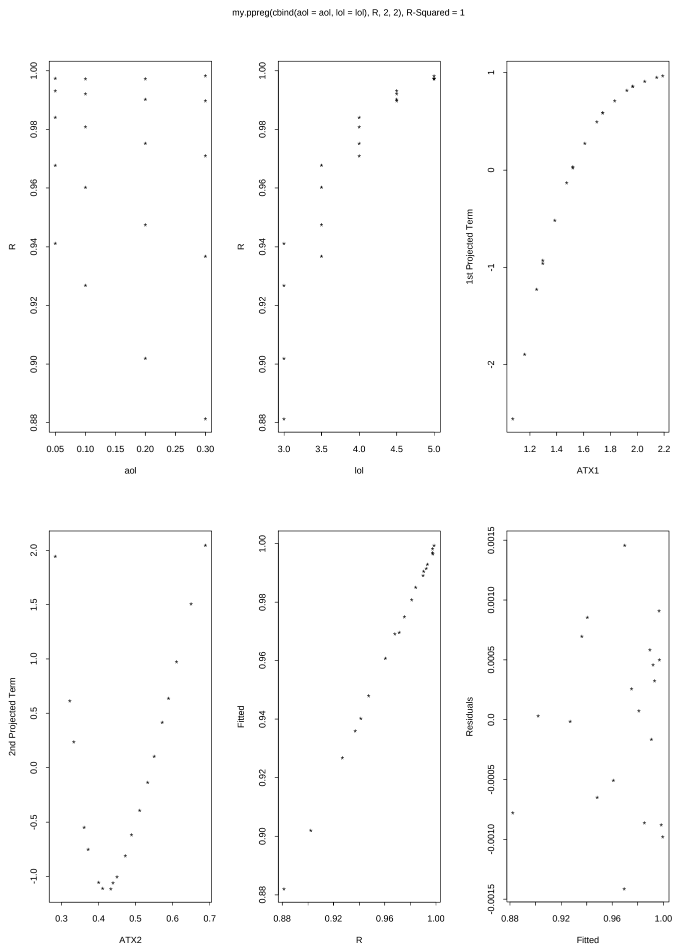

Figure 3 - Two-Term PPR Model

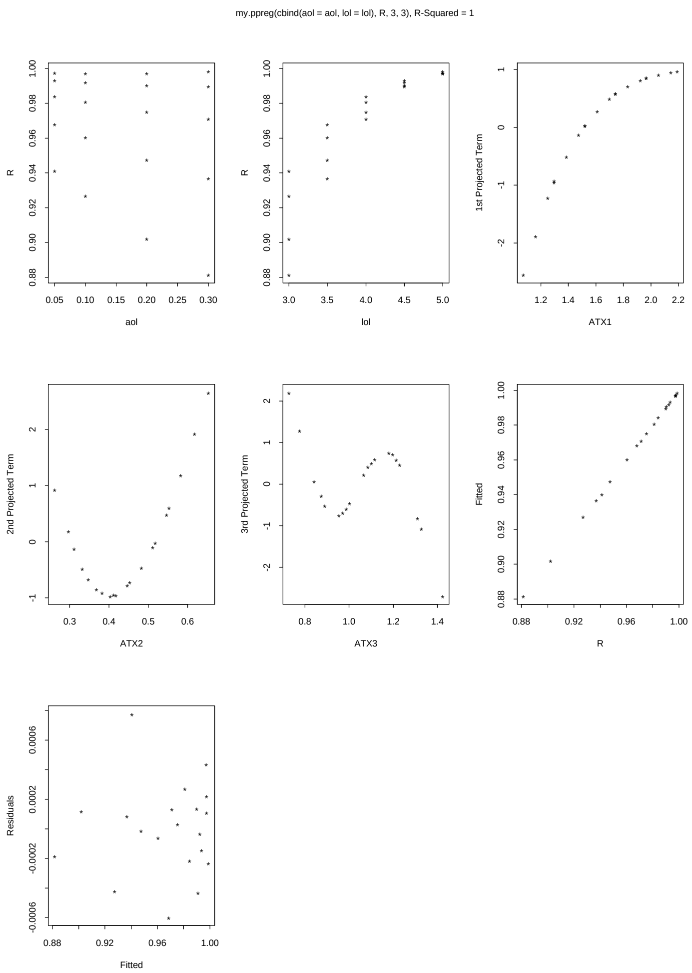

Figure 4 - Three-Term PPR Model

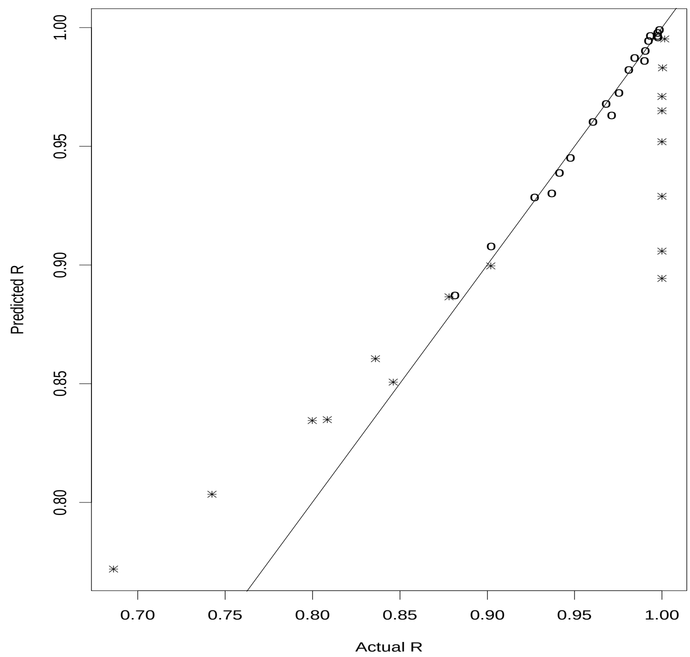

Figure 5 - Extrapolated Prediction Plot Using Equation (15)

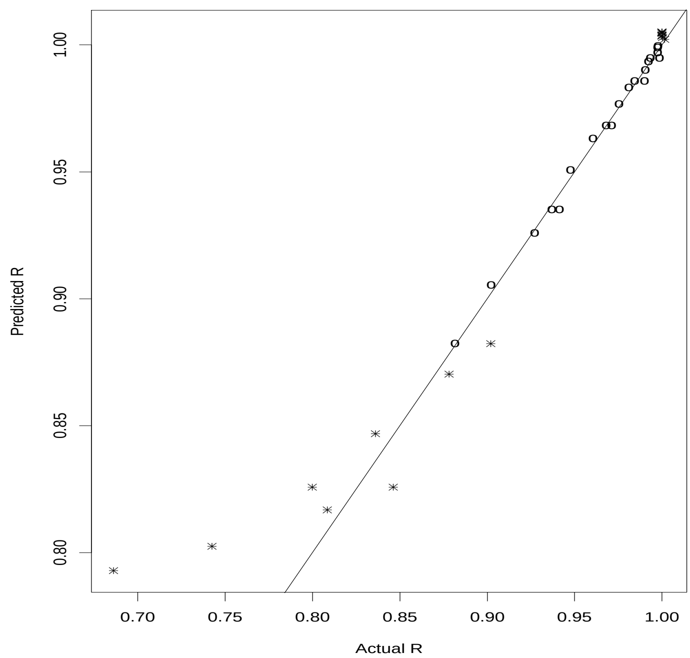

Figure 6 - Extrapolated Prediction Plot Using Equations (16), (17) and (19)