Preliminary Analysis of AASHO Road Test Flexible Pavement Data Using Linear Mix Effects Models

Hsiang-Wei Ker and Ying-Haur Lee

***Abstract*—Multilevel data are very common in many fields. Because of its hierarchical data structure, multilevel data are often analyzed using Linear Mixed-Effects (LME) models. The exploratory analysis, statistical modeling, and the examination of model-fit of LME models are more complicated than those of standard multiple regressions. A systematic modeling approach using visual-graphical techniques and LME models was proposed and demonstrated using the original AASHO road test flexible pavement data. The proposed approach including exploring the growth patterns at both group and individual levels, identifying the important predictors and unusual subjects, choosing suitable statistical models, selecting a preliminary mean structure, selecting a random structure, selecting a residual covariance structure, model reduction, and the examination of the model fit was further discussed.**

***Index Terms*—Multilevel data, linear mixed-effects models, flexible pavement, AASHO road test.**

I. Introduction

Longitudinal data are used in the research on growth, development, and change. Such data consist of measurements on the same subjects repeatedly over time. To describe the pattern of individual growth, make predictions, and gain more insight into the underlying causal relationships related to developmental pattern requires studying the structure of measurements taken on different occasions [1]. Multivariate analysis of variance (MANOVA), repeated measures ANOVA, and standard multiple regression methods have historically been the most widely used tools for analyzing longitudinal data. Polynomial functions are usually employed to model individual growth patterns.

Classical longitudinal data analysis relies on balanced designs where each individual is measured at the same time (i.e., no missing observations). MANOVA, which imposes no constraints on residual covariance matrix, is one common approach in analyzing longitudinal data. However, an unconstrained residual covariance structure is not efficient if the residual errors indeed possess a certain structure, especially when this structure is often of interest in longitudinal studies. Repeated measures ANOVA have the assumption of sphericity. It is too restrictive for longitudinal data because such data often exhibit larger correlations between nearby measurement than between measurements that are far apart. The variance and covariance of the within-subject errors also vary over time. The sphericity assumption is inappropriate in longitudinal studies if residual errors exhibit heterogeneity and dependence.

In longitudinal studies, the focus is on determining whether subjects respond differently under different treatment conditions or at different time points. The errors in longitudinal data often exhibit heterogeneity and dependence, which call for structured covariance models. Longitudinal data typically possess a hierarchical structure that the repeated measurements are nested within an individual. While the repeated measures are the first level, the individual is the second-level unit and groups of individuals are higher level units [2]. Traditional regression analysis and repeated measures ANOVA fail to deal with these two major characteristics of longitudinal data, i.e. heterogeneity and dependence.

Linear Mixed-Effects (LME) models are an alternative for analyzing longitudinal data. These models can be applied to data where the number and the spacing of occasions vary across individuals and the number of occasions is large. LME models can also be used for continuous time. LME models are more flexible than MANOVA in that they do not require an equal number of occasions for all individuals or even the same occasions. Moreover, varied covariance structures can be imposed on the residuals based on the nature of the data. Thus, LME models are well suited for longitudinal data that have variable occasion time, unbalanced data structure, and constrained covariance model for residual errors.

A systematic modeling approach using visual-graphical techniques and LME models was proposed and demonstrated using the original AASHO road test flexible pavement data [3]. The proposed approach including characterizing the growth patterns at both group and individual levels, identifying the important predictors and unusual subjects, choosing suitable statistical models, selecting random-effects structures, suggesting possible residuals covariance models, and examining the model-fits will be further discussed [4-6].

II. Methods

Hierarchical linear models allow researchers to analyze hierarchically nested data with two or more levels. A two-level hierarchical linear model consists of two submodels: individual-level (level-1) and group-level

Manuscript received January 25, 2010. This work was supported in part by the National Science Council, Taiwan, Republic of China under Grants NSC97-2221-E-032-034 and NSC96-2211-E-032-036.

Hsiang-Wei Ker is with the International Trade Department, Chihlee Institute of Technology, Taiwan 220 (phone/fax: 886-2-2623-2408; e-mail: hker@mail.chihlee.edu.tw).

Ying-Haur Lee is with the Civil Engineering Department, Tamkang University, Taiwan 251 (e-mail: yinghaur@mail.tku.edu.tw).

(level-2). The parameters in a group-level model specify the unknown distribution of individual-level parameters. The intercept and slopes at individual-level can be specified as random. Substituting the level-2 equations for the slopes into the level-1 model yields a linear mixed-effects (LME) model. LME models are mixed-effects models in which both fixed and random effects occur linearly in the model function [7].

In a typical hierarchical linear model, the individual is the level-1 unit in the hierarchy. An individual has a series of measurements at different time points in longitudinal studies [8]. When modeling longitudinal data, the repeated measurements are the level-1 units (i.e., a separate level below individuals). The individual is the second-level unit, and more levels can be added for possible group structures [2]. The basic model at the lowest level, also regarded as repeated-measures level, for the application of hierarchical linear model in longitudinal data can be formulated as:

$$\text{Level-1:} \quad \mathrm{Y}_{tj} = \beta_{0\mathrm{j}} + \beta_{1j} c_{tj} + \beta_{2\mathrm{j}} x_{\mathrm{tj}} + r_{\mathrm{tj}} \tag{1}$$

Where Y_{tj} is the measure for an individual j at time t, c is the time variable indicating the time of measurement for this individual, x_{tj} is the time-varying covariate, and r_{tj} is the residual error term.

$$\begin{aligned} \text{Level-2:} \quad & \beta_{0j} = \gamma_{00} + \gamma_{01} W_{j1} + u_{0j} \\ & \beta_{1j} = \gamma_{10} + \gamma_{11} W_{j1} + u_{1j} \\ & \beta_{2j} = \gamma_{20} + \gamma_{21} W_{j1} + u_{1j} \end{aligned} \tag{2}$$

In this level-2 equation, W is the time-invariant covariate for this individual. After substituting level-2 equations into level-1, the combined or the linear mixed-effects model is:

$$\begin{aligned} Y_{tj} &= \gamma_{00} + \gamma_{10} c_{tj} + \gamma_{20} x_{tj} + \gamma_{01} W_{j1} + \gamma_{11} W_{j1} c_{tj} \\ &\quad + \gamma_{21} x_{tj} W_{j1} + u_{0j} + u_{1j} c_{tj} + u_{2j} x_{tj} + r_{tj} \end{aligned} . \tag{3}$$

The level-1 model is a within-individuals model and the level-2 model is a between-individuals model [9]. Note that there is no time-invariant covariate in level-2 before introducing the variable W. The variance and covariance of the u's are the variances and covariances of the random intercept and slopes. After introducing the variable W, the variance and the covariance of u's are the variance and covariance of residual intercept and slopes after partitioning out the variable W. More time-invariant variables can be added sequentially into level-2 to get different models. The reduction in variance of u's could provide an estimate of variance in intercepts and slopes accounted for by those W's [10]. This linear mixed-effects model does not require that every individual must have the same number of observations because of possible withdrawal from study or data transmission errors.

Let $\boldsymbol{Y}_{tj}$ denotes the t^{th} measurement on the j^{th} individual, in which t = 1, 2, … n_i measurements for subject j, and j = 1, 2, … N individuals. The vector $\boldsymbol{Y_j}$ is the collection of the observations for the j^{th} individual. A general linear mixed-effects model for individual j in longitudinal analysis can be formulated as:

$$\mathbf{Y}_j = \mathbf{X}_j \boldsymbol{\beta} + \mathbf{Z}_j \mathbf{U}_j + \mathbf{R}_j , \tag{4}$$

Where $\mathbf{X_j}$ is a $(n_j \times p)$ design matrix for the fixed effects; and $\boldsymbol{\beta}$ is a $(p \times 1)$ vector of fixed-effect parameters. $\mathbf{Z_j}$ is a $(n_j \times r)$design matrix for the random effects; and $\boldsymbol{U}_j$ is an $(r \times 1)$ vector of random-effect parameters assumed to be independently distributed across individuals with a normal distribution, $\boldsymbol{U}_j$ ~NID($\mathbf{0}$, $\mathbf{T}$). The $\boldsymbol{U}_j$ vector captures the subject-specific mean effects as well as reflects the extra variability in the data. $\boldsymbol{R}_j$ is an $(n_j \times 1)$ vector for the residuals. The within errors, $\boldsymbol{R}_j$, are assumed normally distributed with mean zero and variance $\sigma^2\mathbf{W}_j$, where $\mathbf{W_j}$ (stands for "within") is a covariance matrix with a scale factor σ^2. The matrix $\mathbf{W_j}$ can be parameterized by using a few parameters and assumed to have various forms, e.g., an identity matrix or the first-order of autoregression or moving-average process [11-12]. They are independent from individual to individual and are independent of random effects, $\boldsymbol{U}_j$.

Other choices for variance-covariance structures that involve correlated within-subject errors have been proposed. Using appropriate covariance structures can increase efficiency and produced valid standard errors. The choice among covariance depends upon data structures, subject-related theories and available computer packages (Louis, 1988). In some cases, heterogeneous error variances can be employed in the model because the variances in this model are allowed to increase or decrease with time. The assumption of common variance shared by all individuals is removed [11, 13]. There are some possible forms for modeling heterogeneity for within-subject errors. For example, a variance model with different variances for each level of a stratification variable x_{ij} is appropriate if the data shows that different variances appear for the levels of the stratification variable. A variance model with power function can be used if the within-subject variability is to increase with some power of the absolute value of a covariate x_{ij}.

LME models generally assume that level-1 residual errors are uncorrelated over time. This assumption is questionable for longitudinal data that have observations closely spaced in time. There typically exists dependence between adjacent observations. This is called serial correlation and it tends to diminish as the time between observations increases. Serial correlation is incorporated into models for time-series data. Serial correlation is part of the error structure and if it is present, it must be part of the model for producing proper analysis [11]. If the dependent within-subject errors are permitted, the choice of the model to represent the dependence needs careful consideration. It would be preferable to incorporate as much individual-specific structure as possible before introducing a serial correlation structure into within-subject errors [14]. For modeling the dependency of within-subject errors, autoregressive models, moving average models, and a mixture of autoregressive-moving average models may be used.

III. DATA DESCRIPTION

The AASHO road test was a large-scale highway research project conducted near Ottawa, Illinois from 1958 to 1960, and has had by far the largest impact on the history of pavement performance analysis. The test consisted of 6 loops, numbered 1 through 6. Each loop was a segment of a four-lane divided highway and centerlines divided the pavements into inner and outer lanes, called lane 1 and lane 2. Pavement designs varied from section to section. The axle loads on each loop and lane are listed in Table 1. All sections had been subjected to almost the same number of axle load applications on any given date.

Pavement performance data was collected based on the trend of the pavement serviceability index at 2-week interval. The last day of each 2-week period was called an "index day." Index days were numbered sequentially from 1 (November 3, 1958) to 55 (November 30, 1960) [3, 15].

Table 1 Magnitude of Axle Loads on Each Loop and Lane

Lane	Loop					
	1	2	3	4	5	6
1	0	2-S	12-S	18-S	22.4-S	30-S
2	0	6-S	24-T	32-T	40-T	48-T

Note: The axle loads are in thousand pounds (kips); S stands for single axle; T stands for tandem axle.

Empirical relationships between pavement thickness, load magnitude, axle type, accumulated axle load applications, and performance trends for both flexible and rigid pavements were developed after the completion of the road test. Several combinations of certain rules, mathematical transformations, analyses of variance, graphs, and linear regression techniques were utilized in the modeling process to develop such empirical relationships. A load equivalence factor was then established to convert different configurations of load applications to standard 18-kip equivalent single-axle loads (ESAL). This ESAL concept has been adopted internationally since then. As pavement design evolves from traditional empirically based methods toward mechanistic-empirical, the ESAL concept used for traffic loads estimation is no longer adopted in the recommended Mechanistic-Empirical Pavement Design Guide (MEPDG) [16], although many researchers have argued that it is urgently in need of reconsideration [3, 17-18].

During the road test, it was found that the damage rate was relatively low in winter but was relatively high in spring for flexible pavements. Therefore, load applications were adjusted by "seasonal weighting function" such that a better "weighted" flexible pavement equation was developed. Lee [1993] has pointed out that the error variance increases when the predicted number of weighted load repetitions (W) increases. To serve the needs of predicting pavement serviceability index (PSI) after certain load applications on a given section, it is not uncommon that engineers would rearrange the original flexible pavement equation into the following form:

$$PSI = 4.2 - 2.7 * 10^{\left[0.4 + \frac{1094}{(SN+1)^{5.19}}\right] * \left[\log(ESAL) - 9.36 * \log(SN+!) + 0.2\right]} \quad (5)$$
$$SN = 0.44 D_1 + 0.14 D_2 + 0.11 D_3$$

In which the regression statistics are: R^2=0.212, SEE=0.622, N=1083. Fig. 1 shows the predicted versus actual PSI values. Note that PSI ranges from 0 to 5 (0-1 for very poor; 1-2 for poor; 2-3 for fair; 3-4 for good; and 4-5 for very good conditions). D_1 is the surface thickness (in.); D_2 is the base thickness (in.); D_3 is the subbase thickness (in.).

IV. EXPLORATORY ANALYSIS

Exploratory analysis is a technique to visualize the patterns of data. It is detective work of exposing data patterns relative to research interests. Exploratory analysis of longitudinal data can serve to: (a) discover as much of the information regarding raw data as possible rather than simply summarize the data; (b) highlight mean and individual growth patterns which are of potential research interest; as well as (c) identify longitudinal patterns and unusual subjects. Hence plotting individual curves to carefully examine the data should be performed first before any formal curve fitting is carried out. For the nature of this flexible pavement data, the exploratory analysis includes exploring "growth" patterns and the patterns regarding experimental conditions.

A. Exploring "Growth" Patterns

The first step, which is perhaps the best way to get a sense of a new data, is to visualize or plot the data. Most longitudinal data analyses address individual growth patterns over time. Thus, the first useful exploratory analysis is to plot the response variable against time including individual and overall mean profiles. Individual mean profiles, which summarize the aspects of response variable for each individual over time, can be used to examine the possibility of variations among individuals and to identify potential outliers. The overall mean profile summarizes some aspects of the response variable over time for all subjects and is helpful in identifying unusual time when significant differences arise.

Fig. 2 shows the lines connecting the dependent variable (mean PSI) over time for each subject (loop/lane). Most subjects have higher mean PSIs at the beginning of the observation period, and they tend to decrease over time. The spread among the subjects is substantially smaller at the beginning than that at the end. In addition, there exist noticeable variations among subjects. The overall mean growth curve over time is presented in Fig. 3. The overall mean PSIs are larger at the beginning and decrease over time; and the rate of deterioration is higher at the beginning than that at the end.

B. Exploring the Patterns of Experimental Conditions

In addition to time (in terms of index day), different major experimental conditions may be considered. This exploratory analysis is intended to discover the overall and individual patterns of each experimental condition and their interactions on mean PSIs. Subsequently, the patterns of mean PSIs for each subject and the patterns of overall mean PSIs on each experimental condition and their interactions over time are investigated. Fig. 4 is an example plot of mean PSIs for each subject on different surface thickness over time. Generally speaking, the mean PSIs for pavements with higher surface thickness are higher than those with thinner surface layer.

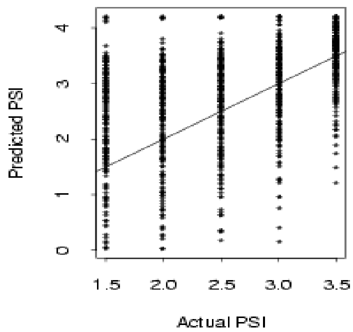

Fig. 1 Predicted versus Actual PSI

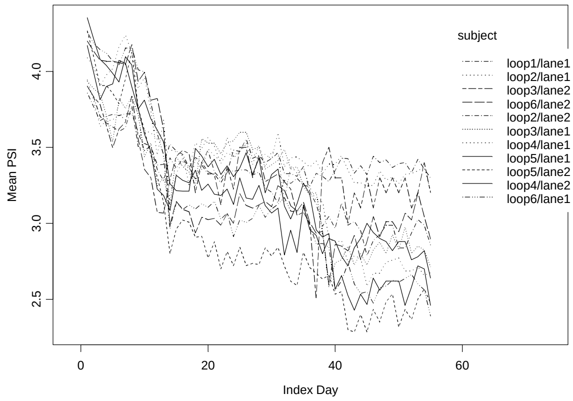

Fig. 2 Mean PSI for each subject (loop/lane) versus index day

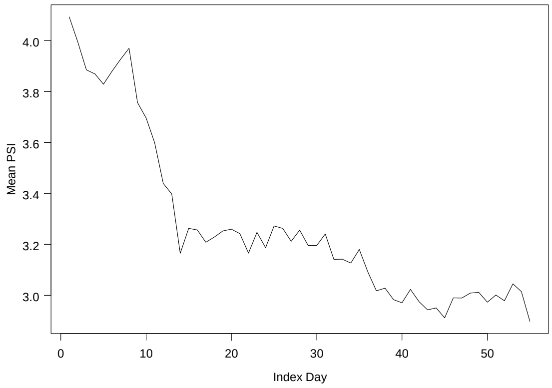

Fig. 3 Overall Mean PSI versus index day

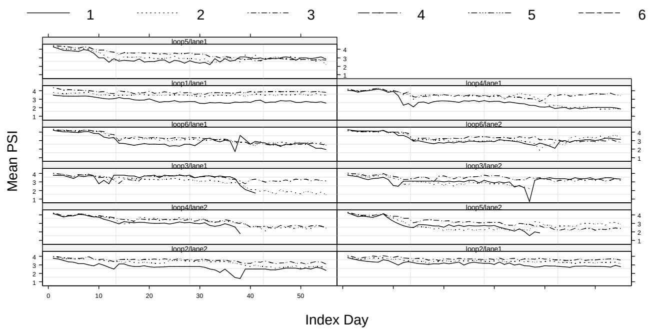

Fig. 4 Mean PSI for each subject (loop/lane) on different surface thickness versus index day

V. LINEAR-MIXED EFFECTS MODELING APPROACH

The following proposed modeling approach is generally applicable to modeling multilevel longitudinal data with a large number of time points. Model building procedures including the selection of a preliminary mean structure, the selection of a random structure, the selection of a residual covariance structure, model reduction, and the examination of the model fit are subsequently illustrated.

A. Selecting a Preliminary Mean Structure

Covariance structures are used to model variation that cannot be explained by fixed effects and depend highly on the mean structures. The first step to model building is to remove the systematic part and remove this so that the variation can be examined. The dataset includes the following explanatory variables: thick, basethk, subasthk, uwtappl, FT. In which, thick is the surface thickness (in.); basethk is the base thickness (in.); subasthk is the subbase thickness (in.); uwtappl is the unweighted applications (millions), and FT is monthly the freeze-thaw cycles.

A model containing all main effects, and all the two-way, three-way interaction terms was first investigated. This model (called model-1) has the form:

$$\begin{aligned}\overline{PSI}_{ij} &= \beta_{0j} + \beta_{1j}(thick)_{ij} + \beta_{2j}(basethk)_{ij} + \beta_{3j}(subasthk)_{ij} + \beta_{4j}(uwtappl)_{ij} \\ &\quad + \beta_{5j}(uwtappl)^2_{ij} + \beta_{6j}(FT) + \text{two-way internaction terms of thick,} \\ &\quad \text{basethk, subasthk, and uwtappl} + \text{three-way internaction terms of} \\ &\quad \text{thick, basethk, subasthk, and uwtappl} + R_{ij}\end{aligned} \tag{6}$$

B. Selecting a Preliminary Random Structure

The second step is to select a set of random effects in the covariance structure. An appropriately specified covariance structure is helpful in interpreting the random variation in the data, achieving the efficiency of estimation, as well as obtaining valid inferences of the parameters in the mean structure of the model. In longitudinal studies, the same subject is repeatedly measured over time. The data collected from longitudinal study is a collection of correlated data. The within-subject errors are often heteroscedastic (i.e., having unequal variance), correlated, or both.

1) Exploring preliminary random-effects structure

A useful tool to explore the random-effects structure is to remove the mean structure from the data and use ordinary least square (OLS) residuals to check the need for a linear mixed-effects model and decide which time-varying covariate should be included in the random structure.

The boxplot of residuals by subject corresponding to the fit of a single linear regression by using the same form of the preliminary level-1 model was conducted. This is the case when grouping structure is ignored from the hierarchy of data. Since the residuals are not centered around zero, there are considerable differences in the magnitudes of residuals among subjects. This indicates the need for subject effects, which is precisely the motivation for using linear mixed-effects model. Separate linear regression models were employed to fit each subject to explore the potential linear relationship.

To assist in selecting a set of random effects to be included in the covariance model, the plots of mean raw residuals against time and the variance of residuals against time are useful. If only random-intercepts models hold, the residual has the form, $e_{ij} = U_{0j} + R_{ij}$, in which $\boldsymbol{U_{0j}}$ is the random effect for intercepts and $\boldsymbol{R_{ij}}$ is the level-1 error. If this plot shows constant variability over time or the curves are flat, then only random intercept model is needed. If random-intercepts-and-slopes models hold, the residual has the form, $e_{ij} = U_{0j} + U_{1j}x_{1ij} + \ldots + U_{qj}x_{qij} + R_{ij}$, where $\boldsymbol{U_{0j}}$ is the random effect for the q-th slope. In the case of random-intercepts-and-slopes model, the plot would show the variability varies over time or there are some unexplained systematic structures in the model. One or more random effects, additional to random intercept, have to be added.

2) Selecting a variance-covariance structure for random effects

Three possible variance-covariance structures including general positive definite (unstructured), diagonal, and block-diagonal based on different assumptions were investigated. General positive-definite is a general covariance matrix parameterized directly in terms of variances and covariances. Diagonal covariance structure is used when random-effects are assumed independent. Block-diagonal matrix is employed when it is assumed that different sets of random effects have different variances. The forms of these three variance-covariance structures are given in equation (7).

$$Unstructured = \begin{bmatrix} \sigma_1^2 & \sigma_{21} & \sigma_{31} \\ \sigma_{21} & \sigma_2^2 & \sigma_{32} \\ \sigma_{31} & \sigma_{32} & \sigma_3^2 \end{bmatrix} \qquad Diagonal = \begin{bmatrix} \sigma_1^2 & 0 & 0 \\ 0 & \sigma_2^2 & 0 \\ 0 & 0 & \sigma_3^2 \end{bmatrix}$$
$$Block-diagonal = \begin{bmatrix} \sigma_1^2 & 0 & 0 \\ 0 & \sigma_2^2 & 0 \\ 0 & 0 & \sigma_2^2 \end{bmatrix} \tag{7}$$

Table 2 displays the model comparison of these three models. The unstructured model has the smallest absolute value of log-likelihood among them. The likelihood ratio test for unstructured model versus diagonal model is 168.57 with p-value less than 0.0001. Thus, unstructured variance-covariance model will be used hereafter.

Table 2 Model Comparison Using Three Variance-Covariance Structures

Model	df	AIC	BIC	logLik	Test	L.Ratio	p-value
(1) Unstr	29	12910.29	13117.74	-6426.14			
(2) Diag	22	13056.52	13213.90	-6506.26	1 vs 2	160.234	< 0.0001
(3) Bk-diag	21	13060.14	13210.37	-6509.07	2 vs 3	5.621	0.0177

The random effects of the preliminary level-2 model include intercept, uwtappl, quadratic term of uwtappl, and FT. The variance-covariance structure is a general positive-definite matrix. Putting the preliminary level-1 and level-2 models together, the preliminary linear-mixed-effects model is then:

$$\begin{aligned} \overline{PSI}_{ij} &= \gamma_{00} + \gamma_{10}(thick)_{ij} + \gamma_{20}(basethk)_{ij} + \gamma_{30}(subasthk)_{ij} + \gamma_{40}(uwtappl)_{ij} \\ &+ \gamma_{50}(uwtappl)_{ij}^2 + \gamma_{60}(FT)_{ij} + \gamma_{70}(thick*basethk)_{ij} + \gamma_{80}(thick*subasthk)_{ij} \\ &+ \gamma_{90}(basethk*uwtappl)_{ij} + \gamma_{100}(subasthk*uwtappl)_{ij} \\ &+ \gamma_{110}(basethk*subasthk*uwtappl)_{ij} \\ &+ \gamma_{120}(thick*basethk*subasthk*uwtappl)_{ij} \\ &+ U_{0j} + U_{4j}(uwtappl)_{ij} + U_{5j}(uwtappl)_{ij}^2 + U_{6j}(FT)_{ij} + R_{ij} \end{aligned} \tag{8}$$

C. Selecting a Residual Covariance Structure

The absolute value of log-likelihood for this heteroscedastic model is 6273.29. The need of heteroscedastic model can be formally checked by using the likelihood ratio test, which is summarized in Table 3. The small p-value indicates that the heteroscedastic model explains the data significantly better than homoscedastic model.

Table 3 Comparison of Heteroscedastic and Homoscedastic Models

Model	Df	AIC	BIC	logLik	L.Ratio	p-value
Homoscedastic	29	12910.29	13117.74	-6426.14		
Heteroscedastic	34	12614.57	12857.79	-6273.29	305.71	< 0.0001

1) Modeling dependence

Correlation structures are used to model dependence among the within-subject errors. Autoregressive model with order of 1, called AR(1), is the simplest and one of the most useful models [7]. The autocorrelation function (ACF), which begins autocorrelation at lag 1 and then declines geometrically, for AR(1) is particularly simple. Autocorrelation functions for autoregressive model of order greater than one are typically oscillating or sinusoidal functions and tend to damp out with increasing lag [19].

Thus, AR(1) may be adequate to model the dependency of the within-subject errors. The absolute value of log-likelihood for this heteroscedastic AR(1) model is 6207.24. The estimated single correlation parameter ϕ is 0.125. The heteroscedastic model (corresponding to $\phi = 0$) is nested within the heteroscedastic AR(1) model.

Likewise, the need of heteroscedastic AR(1) model can be checked using likelihood ratio test as given in Table 4. The small p-value indicates that the heteroscedastic AR(1) model explains the data significantly better than heteroscedastic model, suggesting that within-group serial correlation is present in the data.

Table 4 Comparison of Heteroscedastic and Heteroscedastic AR(1) Models

Model	Df	AIC	BIC	LogLik	L.Ratio	p-value
Heteroscedastic	34	12614.57	12857.79	-6273.29		
Heteroscedastic AR(1)	35	12484.47	12734.85	-6207.24	132.10	< 0.0001

D. Model Reduction

After specifying the within-subject error structure, the next step is to check whether the random-effects can be simplified. It is also desirable to reduce the number of parameters in fixed effects in order to achieve a parsimonious model that can well represent the data. A likelihood ratio test statistic, whose sampling distribution is a mixture of two chi-squared distributions, is used to test the need for random-effects. The p-value is determined by equal weight of the p-values of a mixture of two chi-squared distributions. To assess the significance of the terms in the fixed effects, conditional *t*-tests are used.

1) Reduction of random effects

As suggested by Morrell, Pearson, and Brant [20], the matrix of known covariates should not have polynomial effect if not all hierarchically inferior terms are included. The same rule applies to interaction terms. Hence, significance tests for higher-order random effects should be performed first. The random effects included in the preliminary random-effects structure are: intercept, uwtappl, uwtappl^2, and FT. Table 5 shows the models and the associated maximum log-likelihood values. The small p-value indicates that the preliminary random-effects structure (Model 1) explains the data significantly better than the others. Thus, no reduction of random effects is needed.

Table 5 Random-Effects Models with the Associated Maximum Log-Likelihood Values

Model	df	AIC	BIC	logLik	Test	L.Ratio	p-value
(1) Intercept, uwtappl, $uwtappl^2$, FT	35	12484.5	12734.9	-6207.2			
(2) Intercept, uwtappl, FT	31	12729.6	12951.4	-6333.8	1 vs 2	132.10	< 0.0001
(3) Intercept, uwtappl, $uwtappl^2$	31	12573.5	12795.2	-6255.7	1 vs 3	96.99	< 0.0001

2) Reduction of fixed effects

An adequate and appropriately specified random-effects structure implies efficient model-based inferences for the fixed effects. When considering the reduction of fixed effects, one model is nested within the other model and the random-effects structures are the same for the full and the reduced models. Likelihood ratio tests are appropriate for the model comparison. The parameter estimates, estimated standard errors, *t*-statistics and p-values for the fixed effects of the heteroscedastic AR(1) model are revisited. The heteroscedastic AR(1) model can be reduced to a more parsimonious model due to the existence of some insignificant parameter estimates. The reduction of fixed effects starts with removing the parameters with largest p-values, insignificant terms, and combining the parameters not changing significantly. These processes are repeated until no important terms have been left out of the model.

E. Proposed Preliminary LME Model

The final proposed preliminary linear mixed-effects model is listed in Table 6. The fixed-effects structures of the proposed model contain significant treatment effects for thick, basethk, subasthk, uwtappl, uwtappl^2, FT, and several other two-, three-, and four-way interaction terms. The positive parameter estimates for thick, basethk, and subasthk indicates that higher mean PSI values tend to occur on thicker pavements. The parameter estimate of uwtappl is negative indicating that lower PSI values for higher load applications.

Furthermore, the preliminary LME model also indicates that: The standard error for the pavements with surface thickness of 1 in. or 4 in. is about 48% or 20% higher than those with surface thickness of 2 in., respectively. There exists dependency in within-subject errors. The estimated single correlation parameter for the AR(1) model is 0.126.

F. Examination of the Model Fit

Fig. 5 depicts a plot of the population predictions (fixed), within-group predictions (Subject), and observed values versus time for the proposed preliminary LME model by subjects. Population predictions are obtained by setting random-effects to zero whereas within-group predictions use estimated random effects. The prediction line of the within-group predictions follows the observed values more closely indicating the proposed LME model provides better explanation to the data.

VI. CONCLUSIONS

A systematic modeling approach using visual-graphical techniques and LME models which is generally applicable to modeling multilevel longitudinal data with a large number of time points was proposed in this paper. The original AASHO road test flexible pavement data was used to illustrate the proposed modeling approach.

Exploratory analysis of the data indicated that most subjects (loop/lane) have higher mean PSIs at the beginning of the observation period, and they tend to decrease over time. The spread among the subjects is substantially smaller at the beginning than that at the end. In addition, there exist noticeable variations among subjects.

A preliminary LME model for PSI prediction was developed. The positive parameter estimates for thick, basethk, and subasthk indicates that higher mean PSI values tend to occur on thicker pavements. The parameter estimate of uwtappl is negative indicating that lower PSI values for higher load applications. The prediction line of the within-group predictions (Subject) follows the observed values more closely than that of the population predictions (fixed) indicating the proposed LME model provides better explanation to the data.

Table 6 Proposed Preliminary LME Model

Random Effects	Intercept	uwtappl	$uwtappl^2$	FT	Residual
Standard Deviation	0.170	1.679	0.765	0.00722	0.448

Fixed Effects: Parameter	Value	Std.Error	DF	t-value	p-value
(Intercept)	2.4969	0.0703	9423	35.51	< 0.0001
thick	0.2629	0.0122	9423	21.48	< 0.0001
basethk	0.0590	0.0066	9423	8.91	< 0.0001
subasthk	0.0386	0.0041	9423	9.37	< 0.0001
uwtappl	-3.6191	0.5254	9423	-6.89	< 0.0001
uwtappl^2	1.1524	0.2481	9423	4.65	< 0.0001
FT	0.0148	0.0023	9423	6.39	< 0.0001
thick*basethk	-0.0062	0.0016	9423	-3.81	< 0.0001
thick*subasthk	-0.0082	0.0010	9423	-8.07	< 0.0001
basethk*uwtappl	0.1275	0.0172	9423	7.40	< 0.0001
subasthk*uwtappl	0.1355	0.0181	9423	7.50	< 0.0001
thick*basethk*uwtappl	-0.0155	0.0045	9423	-3.43	0.0006
thick*subasthk*uwtappl	-0.0077	0.0036	9423	-2.16	0.0307
basethk*subasthk*uwtappl	-0.0291	0.0029	9423	-9.87	< 0.0001
thick*basethk*subasthk*uwtappl	0.0073	0.0006	9423	11.53	< 0.0001

Note. (a) Model fit: AIC=12481.77, BIC=12710.69, logLik=-6208.89. (b) Correlation structure: AR(1); parameter estimate(s): Phi= 0.126. (c) Variance function structure: for different standard deviations per stratum (thick= 2, 1, 3, 4, 5, 6 in.), the parameter estimates are: 1, 1.479, 0.935, 1.199, 0.982, 0.959.

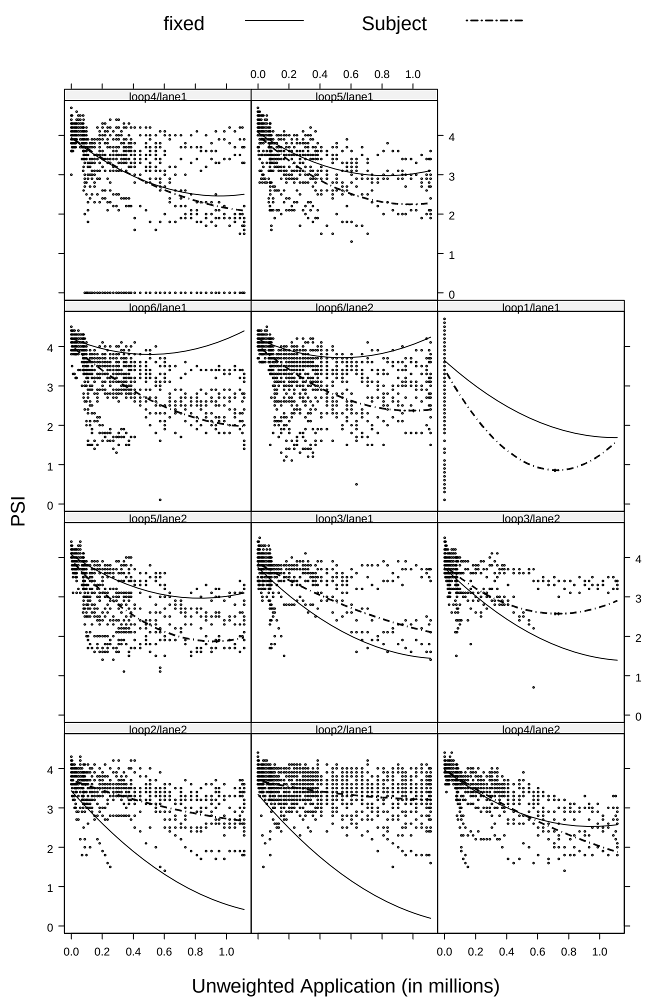

Fig. 5 Population Prediction (fixed), within-group Predictions (Subject), and observed values for the proposed LME model

References

[1] H. Goldstein, *The Design and Analysis of Longitudinal Studies,* New York: Academic Press Inc., 1979.

[2] J. J. Hox, "Multilevel analysis of grouped and longitudinal data," in *Modeling Longitudinal and Multilevel Data: Practical Issues, Applied Approaches and Specific Examples,* T. D. Little, K. U. Schnabel, & J. Baumert , Ed., New Jersey: Lawrence Erlbaum Associates, 2000, pp. 15-32.

[3] Highway Research Board, *The AASHO Road Test, Report 5, Pavement Research, Special Report 61E*. Publication No. 954, National Research Council, Washington, D.C., 1962.

[4] H. W. Ker, Application of Regression Spline in Multilevel Longitudinal Modeling, Doctoral Dissertation, University of Illinois, Urbana, Illinois, 2002.

[5] Y. H. Lee, and H. W. Ker, *Reevaluation and Application of the AASHTO Mechanistic-Empirical Pavement Design Guide, Phase I,* Summary Report, NSC96-2211-E-032-036, National Science Council, Taiwan, 2008. (In Chinese)

[6] Y. H. Lee, and H. W. Ker, *Reevaluation and Application of the AASHTO Mechanistic-Empirical Pavement Design Guide, Phase II,* NSC97-2221-E-032-034, Summary Report, National Science Council, Taiwan, 2009. (In Chinese)

[7] J. C. Pinherio, and D. M. Bates. *Mixed-Effects Models in S and S-Plus*, New York: Springer-Verlag, 2000.

[8] N. M. Laird, and J. H. Ware, "Random effects models for longitudinal data," *Biometrics, 38*, 1982, pp. 963-974.

[9] C. J. Anderson (2001, October 29). *Model building*. Available: http://www.ed.uiuc.edu/courses/edpsy490ck.

[10] R. C. MacCallum, and C. Kim, "Modeling multivariate change," in *Modeling Longitudinal and Multilevel Data: Practical Issues, Applied Approaches and Specific Examples,* T. D. Little, K. U. Schnabel, and J. Baumert, Ed., NJ: Lawrence Erlbaum Associates, 2000, pp. 51-68.

[11] R. H. Jones, *Longitudinal Data with Serial Correlation: A State-Space Approach*. London: Chapman & Hall, 1993.

[12] E. F. Vonesh, and V. M. Chinchilli, *Linear and Nonlinear Models for the Analysis of Repeated Measurements*. New York: Marcel Dekker, Inc., 1997.

[13] B. P. Carlin, and T. A. Louis, *Bayes and Empirical Bayes Methods for Data Analysis*. London: Chapman & Hall, 1996.

[14] H. Goldstein, M. J. R. Healy, and J. Rasbash, "Multilevel time series models with application to repeated measures data," *Statistics in Medicine, 13*, pp. 1643-1655.

[15] Y. H. Huang, Pavement Analysis and Design, 2nd ed., Pearson New Jersey: Education, Inc., 2004.

[16] ARA, Inc. ERES Consultants Division, *Guide for Mechanistic-Empirical Design of New and Rehabilitated Pavement Structure*. NCHRP 1-37A Report, Transportation Research Board, National Research Council, Washington, D. C., 2004.

[17] Y. H. Lee, *Development of Pavement Prediction Models*, Doctoral Dissertation, University of Illinois, Urbana, Illinois, 1993.

[18] H. W. Ker, Y. H. Lee, and P. H. Wu "Development of fatigue cracking performance prediction models for flexible pavements using LTPP database." Journal of Transportation Engineering, ASCE, 134(11), 2008, pp. 477-482.

[19] R.S. Pindyck, and D. L. Rubinfeld, *Econometric Models and Economic Forecasts,* 4th ed. New York: McGraw-Hill, Inc., 1998.

[20] C. H. Morrell, J. D. Pearson, and L. J. Brant , "Linear transformations of linear mixed-effects models," *The American Statistician, 51*, 1997, pp. 338-343.