Technical Report Documentation Page

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
FHWA/TX-02/4040-2		

4. Title and Subtitle	5. Report Date
FINAL RESEARCH PLAN OF STATISTICAL ANALYSIS FOR SMERP DATA	August 2001 Resubmitted: September 2001; April 2002
	6. Performing Organization Code
7. Author(s) Haobo Ren, Thomas J. Freeman, and Clifford H. Spiegelman	8. Performing Organization Report No. Report 4040-2
9. Performing Organization Name and Address Texas Transportation Institute The Texas A&M University System College Station, Texas 77843-3135	10. Work Unit No. (TRAIS) 11. Contract or Grant No. Project No. 0-4040
12. Sponsoring Agency Name and Address Texas Department of Transportation Research and Technology Implementation Office P.O. Box 5080 Austin, Texas 78763-5080	13. Type of Report and Period Covered Final Research Plan: September 2000 – April 2001 14. Sponsoring Agency Code

15. Supplementary Notes

Research performed in cooperation with the Texas Department of Transportation and the U.S. Department of Transportation, Federal Highway Administration.

Research Project Title: Analysis of the Supplemental Maintenance Effectiveness Research Program (SMERP) Experiment

16. Abstract

The primary objective of this research is to provide statistical guidance on the performance of the various maintenance treatments placed as part of the Supplemental Maintenance Effectiveness Research Program (SMERP). A key part of this will be the choice of the statistical model.

This report documents the choice of the form of the final model and describes the criterion used to select the model.

17. Key Words	18. Distribution Statement
SMERP, Maintenance, Effectiveness, Analysis, Cost-effective	No restrictions. This document is available to the public through NTIS: National Technical Information Service 5285 Port Royal Road Springfield, Virginia 22161

19. Security Classif.(of this report)	20. Security Classif.(of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	28	

Form DOT F 1700.7 **(8-72)** **Reproduction of completed page authorized**

FINAL RESEARCH PLAN OF STATISTICAL ANALYSIS FOR SMERP DATA

by

Haobo Ren
Graduate Assistant
Texas A&M University Statistics Department

Thomas J. Freeman
Engineering Research Associate
Texas Transportation Institute

and

Clifford H. Spiegelman
Professor and Research Scientist
Texas Transportation Institute

Report 4040-2
Project Number 0-4040
Research Project Title: Analysis of the Supplemental Maintenance Effectiveness Research Program (SMERP) Experiment

Sponsored by the
Texas Department of Transportation
In Cooperation with the
U.S. Department of Transportation
Federal Highway Administration

August 2001
Resubmitted: September 2001

TEXAS TRANSPORTATION INSTITUTE
The Texas A&M University System
College Station, Texas 77843-3135

DISCLAIMER

The contents of this report reflect the views of the authors, who are responsible for the opinions, findings, and conclusions presented herein. The contents do not necessarily reflect the official views or policies of the Texas Department of Transportation (TxDOT) or the Federal Highway Administration. This report does not constitute a standard, specification, or regulation. Additionally, this report is not intended for construction, bidding, or permit purposes.

ACKNOWLEDGMENTS

Special thanks are given to Joe Leidy, Elias Rmeili, Larry Buttler, James Brown, and James Sassin of TxDOT for their assistance in the development and construction of the SMERP experiment, to Donald A. Pinchott for his help in collecting the data, and to the Texas Department of Transportation and the Federal Highway Administration for their continued funding of this research.

TABLE OF CONTENTS

PAGE

LIST OF FIGURES

FIGURE **PAGE**

LIST OF TABLES

TABLE **PAGE**

FINAL RESEARCH PLAN OF STATISTICAL ANALYSIS FOR SUPPLEMENTAL MAINTENANCE EFFECTIVENESS RESEARCH PROGRAM (SMERP) DATA

One of the tasks in this project is to develop a set of tools to analyze the Supplemental Maintenance Effectiveness Research Program (SMERP) data. In pursuit of this, we have reviewed several different models and approaches.

The repeated measures, linear covariate, and non-linear covariate models have been fitted to the SMERP data. We treated the distresses (fatigue or alligator cracking, all other cracking, and bleeding) as three univariate response variables in the models. Since the results for each distress are similar, cracking data was selected to be representative. In addition, in the following context, treatment types 1 - 7 refer to seven pavement treatments: 1 - asphalt rubber, 2 - microsurfacing, 3 - polymer modified emulsion seal coat, 4 - latex modified seal coat, 5 - conventional asphalt cement seal coat, 6 - fog seal, and 7 – control section with no treatment applied.

A plot of all other cracking is shown graphically in Figure 1.

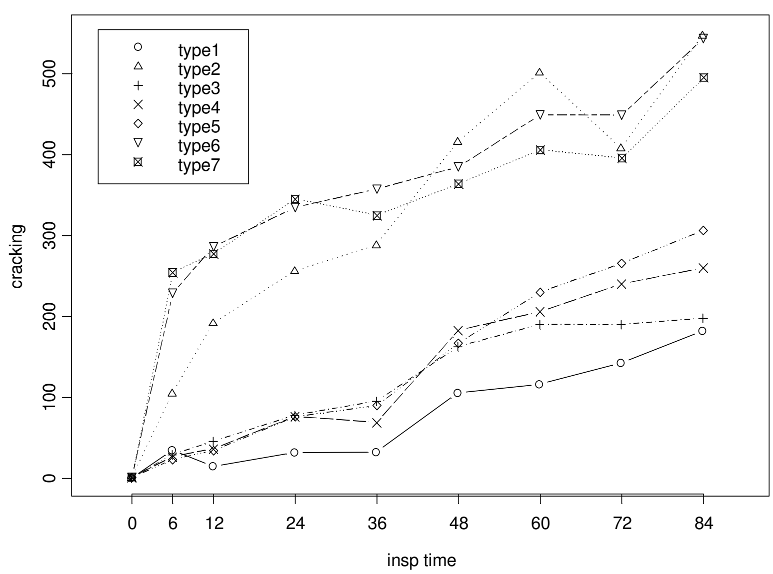

Figure 1. Cracking.

REPEATED MEASURES MODEL

The repeated measures model was the first candidate model tested because for each experimental section, the pavement condition was measured several times and the inspection intervals were nearly equal.

There are two main ways to analyze a repeated measures design: split-plot and multivariate approaches. Multivariate analysis is a necessary complement to verify the validity of the split plot analysis.

For the cracking, the split plot model is

$$Y_{ijk} = \mu + \alpha_i + \rho_k + d_{ik} + \beta_j + (\alpha\beta)_{ij} + e_{ijk},$$

where

i = 1, 2,..,7,

k = 1, 2,…, 20,

j = 1,2, …, 9,

α_i = fixed pavement treatment,

ρ_k = fixed site effect (block),

β_j = fixed time effect,

$(\alpha\beta)_{ij}$ = fixed treatment × time interaction effect, and

d_{ik} = whole plot random error.

The subplot errors are assumed to satisfy the Huynh-Feldt condition of equal variances for differences between all inspection times.

The following is an example analysis of variance (ANOVA) table output for this model:

Table 1. Example ANOVA Table for Repeated Measures Model ($R^2 = 77\%$).

Source	DF	Sum of Squares	Mean Squares	F Value	P-value
Model	195	76038405.6	389940.5	14.89	0.0001
TYPE	6	8895403.0	1482567.2	11.16	0.0001
INSP	8	12330871.5	1541358.9	58.86	0.0001
INSP*TYPE	48	2699722.8	56244.2	2.15	0.0001
PLACE	19	30689140.5	1615217.9	9.96	0.0001
TYPE*PLACE	114	19096833.0	167516.1	6.40	0.0001
Error	866	22676605.2	26185.0		
Corrected Total	1061	98715010.8			

The first p-value listed in the last column is less than 0.0001. A value this low indicates that the model is significant with a satisfied squared multiple correlation, R^2 value of 77 percent. This correlation means that 77 percent of variation in cracking is captured by this split-plot model. The other p-values are all much less than 0.05, hence the effects of treatment type, inspection date, site location, and interaction are all significant. We are especially concerned with the effects of treatment type, so we implemented a post-ANOVA analysis to provide more detail. The Tukey Honestly Significant Difference (HSD) test (Kuehl, 1994), which was developed by Tukey for pairwise comparison of all treatment types, gives the result shown in Table 2.

Table 2. Results of the Tukey's HSD Test.

Grouping	Mean	N	TYPE
A	313.54	139	6
A	298.80	143	7
A	277.95	156	2
B	113.82	156	5
B	105.94	156	4
B	98.62	156	3
B	63.59	156	1

This table shows that types 2, 6, and 7 are not significantly different; types 1, 3, 4, and 5 do not differ significantly. The data can be arranged into groups A (2, 6, and 7) and B (1, 3, 4, and 5), and the types in the different groups are significantly different.

The trend analysis was conducted to determine the quantitative trend relationship between cracking condition and inspection time, which can be simplified by examining orthogonal contrasts among the inspection time levels that measure the linear, quadratic, and higher-level polynomial effects. These contrasts, known as orthogonal polynomials, enable us to evaluate the importance of each polynomial component with a specific contrast. Table 3 shows the output of the orthogonal polynomial analysis.

Table 3. Analysis of Variance for the Orthogonal Polynomial Model.

Contrast	DF	Contrast SS	Mean Square	F Value	P-value
Linear	1	11350662.1	11350662.1	433.47	0.0001
Quadratic	1	37030.4	37030.4	1.41	0.2347
Cubic	1	41671.3	41671.3	1.59	0.2075
Quartic	1	41048.2	41048.2	1.57	0.2109
Fifth	1	110755.0	110755.0	4.23	0.0400
Sixth	1	17843.2	17843.2	0.68	0.4093

This table indicates that the only significant trend is the linear model, but the small p-value for the fifth order component stands, though it is not significant.

The Hynh – Feldt (H-F) condition is required for the usual analysis of variance for the above model, which means that variance of the difference between any pair of observations receiving the same treatment but at the different time must be equal. This structure, also termed sphericity, is a necessary and sufficient condition for the F tests to be appropriate. Multivariate analysis of variance can test the H-F condition. The basic model for cracking values is:

$$Y_{ijk} = \mu_{ij} + e_{ijk},$$

where

i = treatment type (1, 2,..,7);

k = site number (1, 2,…, 20); and

j = inspection number (1,2, …, 9).

The result of the Mauchly's test, which can be applied to test the null hypothesis of sphericity (Littell, Freund, and Spector, 1991) extracted from the multivariate analysis of variance, is shown in Table 4.

Table 4. Test for Sphericity.

Mauchly's Criterion	Chisquare Approximation	DF	P-value
0.0000298	772.4463	35	0.0000

Data shown in Table 4 confirm that the split-plot analysis may not be valid. The Wilks' Lambda likelihood ratio test on the general linear hypothesis of no-time effect has a p-value < 0.0001 and the no-time-treatment interaction effect has a p-value = 0.5653, which shows the significance of inspection time but a non-significant interaction effect between time and treatment.

A final analysis forms contrast (orthogonal polynomials) in the time variable and tests the differences in this contrast across the levels of the type variable. Since we have nine time periods, we can form eight orthogonal polynomials, which summarize the data across the repeated factor (time period). The polynomial trend can thus be considered the response variable in a complete random design with seven treatments and eight responses. Table 5 shows the p-values from the polynomial analysis.

Table 5. P-values of Tests for Polynomial Trends.

Polynomial Order	Mean	Type
Linear	0.0001	0.4888
Quadratic	0.1279	0.4809
Cubic	0.0501	0.0035
Quartic	0.0368	0.6827
5^{th}	0.0002	0.8954
6^{th}	0.6379	0.6043
7^{th}	0.1878	0.0098
8^{th}	0.0279	0.2457

The column labeled "Mean" tests the hypothesis that, averaged over all the observations, the mean of the specified contrast variable is 0. The column labeled "Type" tests the hypothesis that the mean of the contrast variable is the same for each level of type tested. Since the eight tests are not independent, we would use $\alpha_{PC} = 0.05/8 \approx 0.0063$ as the significance level in these multiple comparisons. The SAS output shows that the linear trend (p-value < 0.0001) and fifth order trend (p-value = 0.0002) are significant, but the non-significant type effects for linear and fifth order shows that the curvatures are almost the same for the seven treatment types.

Although the data tested do not satisfy the H-F condition, the repeated measures model is still a good approach for this project. We grouped the treatment types and found the significant non-linear time trend by ANOVA and post-ANOVA. In the following models, we treat inspection times as exact days rather than as several levels.

ANALYSIS OF COVARIANCE (ANCOVA)

Basically, ANCOVA (Huitema, 1980) is useful when the researcher wishes to examine the relationship among at least two quantitative variables and at least one additional categorical variable. Especially, the researcher may be interested in examining the relationship between two quantitative variables but find that a categorical variable is confounding that relationship. ANCOVA allows one to examine the relationship in question "controlling for" the confounding categorical variable.

The model of ANCOVA can be presented as:

dependent variables = constant + (effect of treatment type) + (effect of covariate) + error, which has two forms:

(1) Traditional model:

$$y_{ij} = \mu_i + \beta\left(x_{ij} - \bar{x}_{i.}\right) + e_{ij},$$

where

μ_i = treatment mean,

β = coefficient for the linear regression of y_{ij} on x_{ij}.

Two additional key assumptions for this model are that the regression coefficient β is the same for all treatment groups and the treatments do not influence the covariate x.

(2) Heterogeneous ANCOVA model:

$$y_{ij} = \mu_i + \beta_i\left(x_{ij} - \bar{x}_{i.}\right) + e_{ij},$$

which allows different slopes for different treatments.

We applied this model to study the relationship among cracking, time, and sites for each pavement treatment. Time is considered as the continuous covariate (here, we calculated the exact days between the inspection date and the construction date, while in the repeated measures model, we only used the approximate months) with type and site as the categorical variable.

At first, we fit a more general covariate model:

$$y_{ijk} = \mu + \alpha_i + \rho_j + \beta_{ij} x_{ijk} + e_{ijk},$$

where

i = 1, 2, .., 7,

j = 1, 2, …, 20,

k = 1, 2, …, 9, α_i = fixed pavement treatment,

ρ_j = fixed site effect (block),

β_{ij} = coefficient for the linear regression of y_{ijk} on x_{ijk}, and

y_{ijk} = cracking index - ratio of cracking area to pavement area.

The ANCOVA table output for this model is listed in Table 6.

Table 6. Example ANCOVA Table for Linear Covariate Model (R^2 =91%).

Source	DF	Sum of Squares	Mean Squares	F Value	P-value
Model	159	2.03087849	0.01277282	55.77	0.0001
TIME	1	0.28783045	0.28783045	1256.67	0.0001
TYPE	6	0.24615822	0.04102637	179.12	0.0001
SITE	19	0.67984932	0.03578154	156.22	0.0001
TIME*TYPE*SITE	133	0.81704049	0.00614316	26.82	0.0001
Error	889	0.20361924	0.00022904		
Corrected Total	1048	2.23449772			

The first p-value listed in the last column is less than 0.0001. This small value indicates that the model is significant with a very high squared multiple correlation, R^2 value of 91 percent. This correlation means that 91 percent of the variation in cracking is captured by this linear covariate model. The other p-values are all much less than the 0.05 level. The significant interaction among time, treatment type, and site indicates a heterogeneous ANCOVA model.

We further tried the ANCOVA for each type of treatment. Tables 7 through 13 contain extracted ANCOVA data showing that for each type of pavement, the model is significant and the heterogeneous ANCOVA is necessary.

Table 7. ANCOVA Table for Treatment Type 1 ($R^2 = 85\%$).

Source	DF	Sum of Squares	Mean Squares	F Value	P-value
TIME	1	0.00819032	0.00819032	111.60	0.0001
SITE	19	0.02891665	0.00152193	20.74	0.0001
TIME*SITE	19	0.01113912	0.00058627	7.99	0.0001

Table 8. ANCOVA Table for Treatment Type 2 ($R^2 = 92\%$).

Source	DF	Sum of Squares	Mean Squares	F Value	P-value
TIME	1	0.09764018	0.09764018	301.76	0.0001
SITE	19	0.22332694	0.00152193	36.33	0.0001
TIME*SITE	19	0.08718655	0.00458877	14.18	0.0001

Table 9. ANCOVA Table for Treatment Type 3 ($R^2 = 94\%$).

Source	DF	Sum of Squares	Mean Squares	F Value	P-value
TIME	1	0.01651027	0.01651027	283.81	0.0001
SITE	19	0.05657327	0.00297754	51.18	0.0001
TIME*SITE	19	0.02648082	0.00139373	23.96	0.0001

Table 10. ANCOVA Table for Treatment Type 4 ($R^2 = 92\%$).

Source	DF	Sum of Squares	Mean Squares	F Value	P-value
TIME	1	0.02687644	0.02687644	245.41	0.0001
SITE	19	0.06861267	0.00361119	32.97	0.0001
TIME*SITE	19	0.05267105	0.00277216	25.31	0.0001

Table 11. ANCOVA Table for Treatment Type 5 ($R^2 = 94\%$).

Source	DF	Sum of Squares	Mean Squares	F Value	P-value
TIME	1	0.03112358	0.03112358	420.09	0.0001
SITE	19	0.07002234	0.00368539	49.74	0.0001
TIME*SITE	19	0.04408005	0.00232000	31.31	0.0001

Table 12. ANCOVA Table for Treatment Type 6 ($R^2 = 92\%$).

Source	DF	Sum of Squares	Mean Squares	F Value	P-value

TIME	1	0.06774384	0.06774384	154.94	0.0001
SITE	17	0.32281973	0.01898940	43.43	0.0001
TIME*SITE	16	0.08458547	0.00528659	12.09	0.0001

Table 13. ANCOVA Table for Treatment Type 7 ($R^2 = 91\%$).

Source	DF	Sum of Squares	Mean Squares	F Value	P-value
TIME	1	0.05575945	0.05575946	118.42	0.0001
SITE	17	0.3519001	0.02070000	43.96	0.0001
TIME*SITE	16	0.07847784	0.00490486	10.42	0.0001

The linear covariate model captures the data more precisely than the other models, since we treat the inspection time as a continuous covariate. From the view of the goodness-of-fit, this model is better than the repeated measures model. However, as with the split-plot model, the data cannot satisfy an important assumption of the linear covariate model which requires that the errors be independent. Therefore, this approach is still an approximation.

NON-LINEAR COVARIATE MODEL

The above linear covariate model has already shown to be a good approximation. However, some high order polynomial trends presented in the above polynomial trend analyses indicate strong non-linearity. Previous pavement studies (Freeman, 2000) recommend an S-shaped curve:

$$g = e^{-\left(\frac{\rho}{W}\right)^{\beta}}$$

where

g = damage index,

W = pavement age depending upon the distress type under consideration, and

ρ, β = scale and shape parameters, respectively.

An assumed exponential error structure:

$$g = e^{-\left(\frac{\rho}{W}\right)^{\beta} e^{\varepsilon}}$$

can give us a linear model after transformation:

$$\log[-\log(g)] = \beta\log(\rho) - \beta\log(W) + \varepsilon \, .$$

It was interesting to exert more covariate analysis with this general linear relation, called non-linear covariate analysis. We applied this model to study the relationship among log[-log(cracking)] (the new response variable), log(time) (the new covariate), and site (still the categorical variable) for each pavement treatment. The results are very close to the linear covariate analysis. In order to compare these two approaches, we also did the linear covariate analysis based on the formula:

$$g = a + bW + \varepsilon \, .$$

As with the linear covariate modeling, we fit a more general covariate model after the above transformation. The ANCOVA table output for this model is included in Table 14.

Table 14. Example ANCOVA Table for Non-linear Covariate Model ($R^2 = 90\%$).

Source	DF	Sum of Squares	Mean Squares	F Value	P-value
Model	121	82.5054439	0.6818632	30.61	0.0001
W	1	6.4265315	6.4265315	288.47	0.0001
TYPE	6	12.5931327	2.0988555	94.21	0.0001
SITE	18	40.9097879	2.2727660	102.02	0.0001
W*TYPE*SITE	96	22.5759918	0.2351666	10.56	0.0001
Error	400	8.9113327	0.0222783		
Corrected Total	521	91.4167766			

The first p-value listed in the last column is less than 0.0001. This very small value indicates that the model is significant with a very high squared multiple correlation, R^2 value of 90 percent. This correlation means that 90 percent of the variation in cracking is captured by this non-linear covariate model. The other p-values are all much less than the 0.05 level. The significant interaction among time, treatment type, and site indicates a heterogeneous ANCOVA model.

Further, we tried this ANCOVA model for each type of treatment. All of the ANCOVA results show that for each type of pavement, the model is significant and the heterogeneous ANCOVA is necessary. We list the R^2 for both models in Table 15.

Table 15. Results of ANCOVA for Each Treatment.

Type	Non-linear Model	Linear Model
1	81.9%	85.0%
2	96.3%	91.6%
3	92.7%	93.7%
4	92.2%	92.1%
5	89.2%	94.4%
6	95.4%	91.7%
7	95.5%	91.2%

The R^2 values are all very similar. Figures 2 and 3 illustrate this point. In Figure 2, the linear model is better than the non-linear, but the non-linear model fits more closely in Figure 3. Overall, the S-shaped covariate analysis is slightly better than linear covariate analysis.

Obviously, we still cannot overcome the problem of the existence of the covariance when modeling by linear ANCOVA. Also, because there are many zeros in the original dataset, the logarithm transformation leads to more undefined values. However, the S-shaped modeling is an appropriate way to handle non-linearity.

CONCLUSION

We have tried several statistical models. The linear models work well, but non-linearity also exists, which has been proven by several analyses. We will continue to study the univariate analysis by the more complicated non-linear models with more parameters (Haas, Hudson and Zaniewski; 1994, Han and Lukanen, 1994; and Visser, Queiroz and Caroca, 1994) then move to multivariate analysis, Finally, we will cope with the data sets where the sites were taken out of service (right censored, with competing risks).

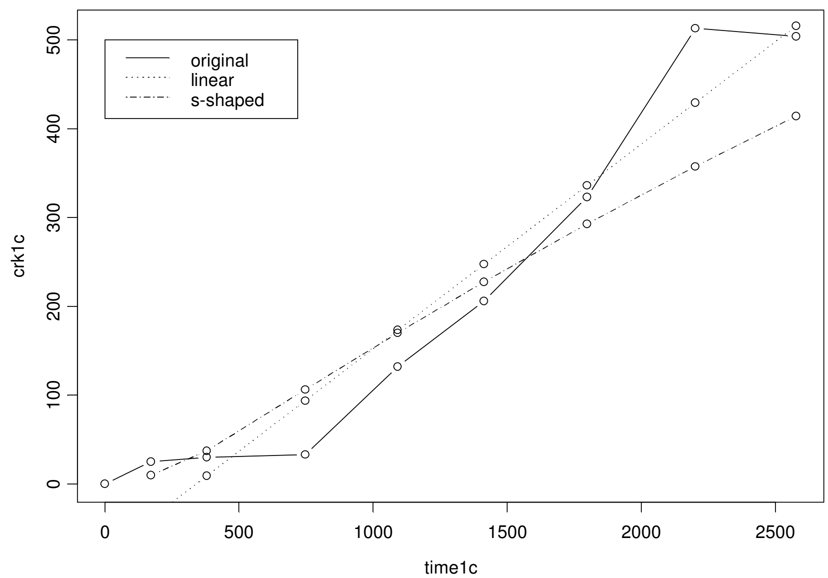

Figure 2. Cracking Type = 1 Site = C.

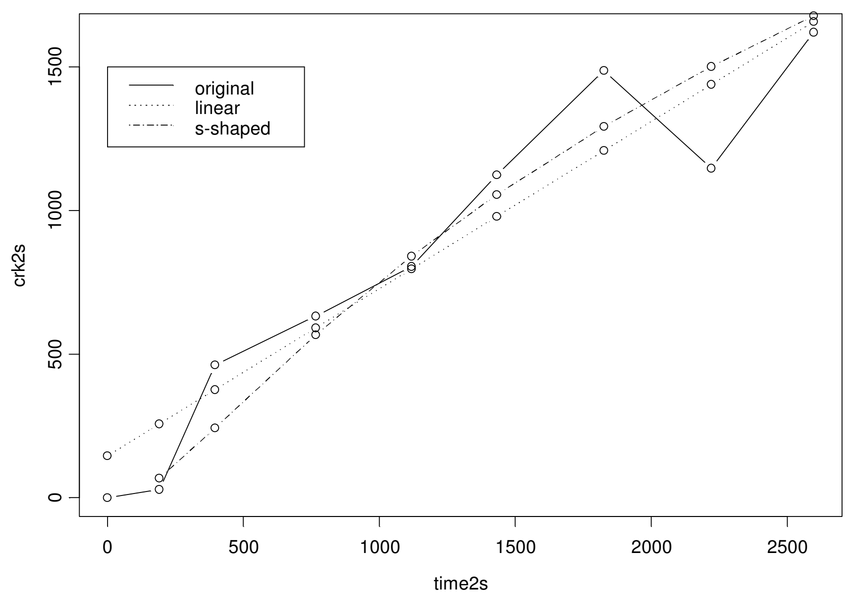

Figure 3. Cracking Type = 2 Site = S.

RESULTS OF LITERATURE SEARCH

The following is a list of much of the reference material used to develop the models. Not all models studied were used, but valuable information was discovered that helped guide the research. Other sources were consulted, but these represent the majority of the reference.

Allison, P. D. (2001) Missing Data, Sage University Series Paper on Quantitative Applications in the Social Sciences, Thousand Oaks, CA: Sage.

This report provided a basic idea about mechanisms for handling missing data which was to help us to model the missing data in SMERP.

Chinchill, V. M. and Vonesh, E. F. (1997), Linear and Nonlinear Models for the Analysis of Repeated measurements, New York: M. Dekker.

A review of the general repeated measures models, which were used in the data analysis, was provided in this reference.

Diggle, R. J., Liang, K. – Y. and Zeger, S. L. (1994), Analysis of Longitudinal Data, Oxford University Press Inc.

This document reviews the general strategy for analyzing longitudinal data. This approach was followed in the exploratory data analysis stage.

Fox, J. (1999), Nonparametric Regression Analysis: Smoothing Scatterplots, University Series Paper on Quantitative Applications in the Social Sciences, Thousand Oaks, CA: Sage.

This report provides an excellent review of non-parametric smoothing techniques to help us understand the “LOWESS” procedure used in the exploratory data analysis stage.

Freeman, T. (2000), Project 0-4040 Proposal.

An introductory background of the SMERP study and some terminology used throughout the whole statistical analysis was included in this report. An attempt was made during the first research stage to fit the nonlinear model included in this proposal. We are currently revisiting that analysis.

Gallant, A. R. and Fuller, W. A. (1973), Fitting Segmented Polynomial Regression Model Whose Joint Points Have to Be Estimated, Journal of the American Statistical Association, 68, 144-147.

The information in this journal helped us to understand the numerical techniques for segmented regression. Since the method outlined in this paper requires stronger conditions on the data than we have, we cannot use the method directly.

Haas, R., Hudson, W. and Zaniewski, J., (1994), Modern Pavement Management, Krieger Publishing Company.

This book was used a background book to help the statisticians know more about pavement knowledge, terminology, and especially, the many kinds of distress.

Han, C. and Lukanen, E. O., (1994), “Performance History and Prediction Modeling for Minnesota Pavements”, Third International Conference on Managing Pavements, Volume I.

The methods proposed for a modeling procedure based on simple, two variable models which relate distress density to age, and additional variables such as surfaces type, traffic and structure were described. We did not use this procedure.

Hand, D. J. and Crowder, M. J. (1996), Practical Longitudinal Data Analysis, Chapman & Hall.

Information in this reference compares several models for longitudinal data. We used the random-coefficient model and some ideas of handling non-normal and non-linear in current study.

Hazelrig, J. B., Turner, M. E. and Blackstone, E. H. (1982), Parametric Survival Analysis Combining Longitudinal and Cross-sectional-censored and Interval censored Data with Concomitant Information, Biometrics, 38, 1-15.

The analysis techniques included in this report were studied very carefully during the second research stage. In the final analysis, the methods were not used it treated the missing data as dropouts, but not as censored.

Huitema, B. E. (1980), The Analysis of Covariance and Alternatives, Wiley-Interscience Publication.

This book helped us to understand the covariance model, and was used in the first research stage.

Kuehl, R. O. (1994), Statistical Principles of Research Design and Analysis, Duxbury Press.

This report reviews classical experimental designs. We used the split-plot and covariance models introduced in the book.

Laird, N. C. and Ware, J. H. (1982) Random-effects Models for Longitudinal Data, Biometrics, 38, 963-974.

This report is the pioneering work on longitudinal data. We used the random-coefficient model found in this report.

Littell, R. C., Freund, R. J. and Spector, P. C. (1991), SAS System for Linear Models, 3rd Edition, SAS Institute Inc.

This programming reference helped us to program in the SAS statistical language using PROC GLM to implement the repeated measures and covariance models in the first research stage.

Littell, R. C., Milliken, G. A., Stroup, W. W. and Wolfinger, R. D. (1996), SAS System for Mixed Models, SAS Institute Inc.

This programming reference helped us to program in the SAS statistical language using PROC MIXED to implement the linear mixed model in the current study.

Palmer, M. J. and Phillips, B. F. and Smith, G. T. (1991), Applications of Nonlinear Models with Random Coefficients to Growth Data, Biometrics, 47, 623-635.

In this reference, we learned to combine the nonlinearity and random coefficients to model the growth curve which aids us in our current model development.

Rutherford, A. (2001), Introducing ANOVA and ANCOVA: a GLM Approach, SAGE publications.

This is a general introduction to ANOVA and ANCOVA which helped us in beginning modeling stage.

Smith, P. L. (1979), Spline: as a Useful and Convenient Statistic Tool, The American Statistician, 30, 57-62.

The information in this reference introduced the topic of splines, which we used to explore piecewise regression. However, the model described assumes fixed knots, which we cannot use directly.

Verbek, G. and Molenberghs, G. (2000), Linear Mixed Models for Longitudinal Data, Springer-Verlag New York, Inc.

This book is a good review for this field and introduces pattern-mixture models for dropout in longitudinal data. We used this efficient model.

Visser, A. T., Queiroz, C. and Caroca, A. (1994), "Total Cost Rehabilitation Design Method for use in Pavement Management", Third International Conference on Managing Pavements, Volume I.

We studied the multiple regression techniques in pavement management modeling including relating cracking index to many variables provided in this reference. We were enlightened by some of the ideas, but did not apply the model.

REFERENCES USED IN THIS REPORT

Freeman, T. (2000), Project 0-4040 Proposal.

Haas, R., Hudson, W. and Zaniewski, J., (1994), Modern Pavement Management, Krieger Publishing Company.

Han, C. and Lukanen, E. O., (1994), "Performance History and Prediction Modeling for Minnesota Pavements", Third International Conference on Managing Pavements, Volume I.

Huitema, B. E. (1980), The Analysis of Covariance and Alternatives, Wiley-Interscience Publication.

Kuehl, R. O. (1994), Statistical Principles of Research Design and Analysis, Duxbury Press.

Littell, R. C., Freund, R. J. and Spector, P. C. (1991), SAS System for Linear Models, 3rd Edition, SAS Institute Inc.

Visser, A. T., Queiroz, C. and Caroca, A. (1994), "Total Cost Rehabilitation Design Method for use in Pavement Management", Third International Conference on Managing Pavements, Volume I.