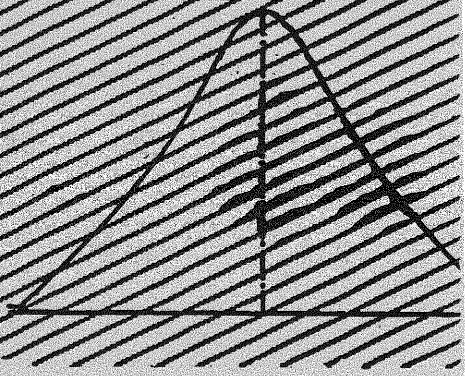
OKLAHOMA DEPARTMENT OF TRANSPORTATION
5 6208 10012 3001

TG 325
.S81
1995
Vol. 1

STATISTICAL MODELS FOR IDENTIFICATION OF PROBLEMATIC BRIDGE SITES AND ESTIMATION OF APPROACH SETTLEMENTS

STUDY 2188, ORA 125-6074



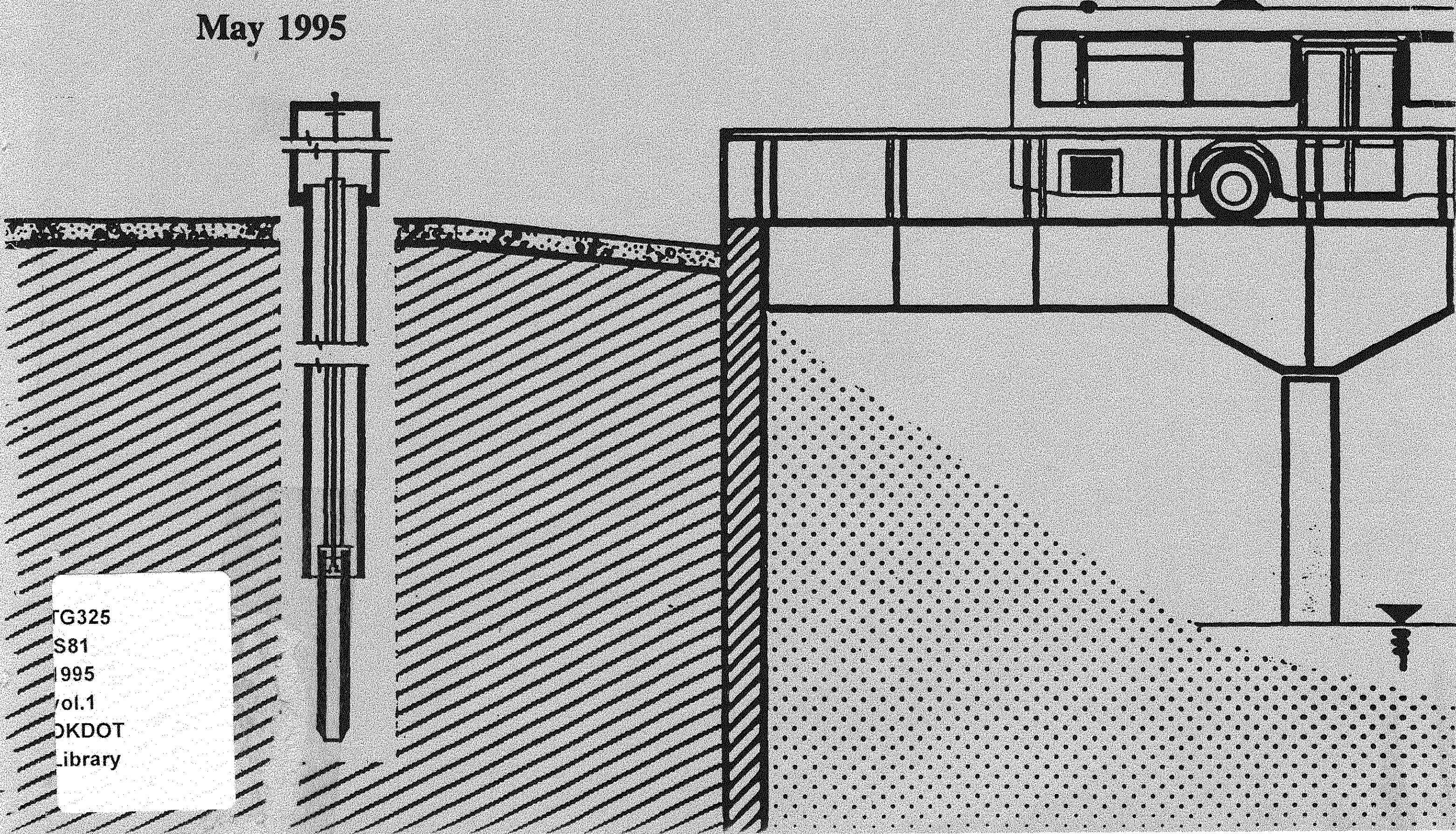
Submitted to
Oklahoma Department of Transportation

Prepared by
Musharraf Zaman
Joakim G. Laguros
Ratan K. Jha

University of Oklahoma

From
University of Oklahoma
Office of Research Administration

May 1995



TG325
S81
1995
vol.1
OKDOT
Library

TECHNICAL REPORT STANDARD TITLE PAGE

1. REPORT NO.
FHWA/OK 94(02)

2. GOVERNMENT ACCESSION NO.

3. RECIPIENT'S CATALOG NO.

4. TITLE AND SUBTITLE
Statistical Models for Identification of
Problematic Bridge Sites and Estimation of
Approach Settlements

5. REPORT DATE
May 1995

6. PERFORMING ORGANIZATION CODE
ORA 125-6074

7. AUTHOR(S)
M. M. Zaman, J. G. Laguros, and R. K. Jha

8. PERFORMING ORGANIZATION REPORT

9. PERFORMING ORGANIZATION AND ADDRESS
The University of Oklahoma
Norman, OK 73019

10. WORK UNIT NO.

11. CONTRACT OR GRANT NO.
Item No. 2188

12. SPONSORING AGENCY NAME AND ADDRESS
Oklahoma Department of Transportation
Research and Development Division
200 NE 21st Street
Oklahoma City, OK 73015

13. TYPE OF REPORT AND PERIOD COVERED

14. SPONSORING AGENCY CODE

15. SUPPLEMENTARY NOTES
Done in cooperation with FHWA

16. ABSTRACT

Bridge approach settlement is caused by the differential movements of the bridge abutment and the approach pavement. The resulting unsafe riding surface not only requires frequent and costly maintenance but also leads to numerous traffic accidents. A number of factors contributing to such approach settlement have been identified, of which age of approach, embankment height, foundation soil thickness, skewness of approach, traffic volume, and the site-specific embankment and foundation soil characteristics are considered in the present study. The site-specific SPT and CPT values (viz. tip resistance and friction ratio) were used to represent the embankment and the foundation soil characteristics at a site. A statistically-based model, using SAS package program, was developed based on the data from 26 bridge sites in Oklahoma. Also, a special "Field Test Model" was developed with the aim of identifying the problematic bridge sites prior to the construction of the bridge. The variables used in the field test model can be easily obtained from the SPT and CPT tests.

The linear model was found to be more effective in identifying problematic sites as well as estimating approach settlements compared to the nonlinear model. The field test model seems to be more efficacious in identifying the problematic sites than estimating the approach settlements. The coefficient of correlation, R^2 (adjusted), varied from about 0.88 to 0.94. It was observed that the site-specific SPT (N-value), cone resistance (q_c), friction ratio and the thickness of the foundation soil have significant impact on estimating the approach settlements as compared to other variables of the models, indicating that the site-specific embankment and foundation soil characteristics are the most influential causative factors.

17. KEY WORDS
Bridge approach, settlement, bump

18. DISTRIBUTION STATEMENT

19. SECURITY CLASSIF. (OF THIS REPORT)
None

20. SECURITY CLASSIF. (OF THIS PAGE)
None

21. NO. OF PAGES
110

22. PRICE

Final Report

STATISTICAL MODELS FOR IDENTIFICATION OF PROBLEMATIC BRIDGE SITES AND ESTIMATION OF APPROACH SETTLEMENTS

(STUDY 2188, ORA 125-6074)

Submitted to

Mr. Lawrence J. Senkowski, P.E.
Assistant Division Engineer
Research and Development
Oklahoma Department of Transportation
200 N.E. 21st Street
Oklahoma City, Oklahoma 73105

Prepared by

Musharraf Zaman, Professor
Joakim G. Laguros, David Ross Boyd Professor Emeritus
Ratan K. Jha, Graduate Research Assistant

School of Civil Engineering and Environmental Science
The University of Oklahoma
Norman, Oklahoma 73019
(405) 325-5911

From

The University of Oklahoma
Office of Research Administration
1000 Asp Avenue, Room 314
Norman, Oklahoma 73019

May 1995

DISCLAIMER

The contents of this report reflect the views of the authors who are responsible for the facts and the accuracy of the data presented herein. The contents do not necessarily reflect the official views of the Oklahoma Department of Transportation or the Federal Highway Administration. This report does not constitute a standard, specification, or regulation. While equipment and contractor names are used in this report, it is not intended as an endorsement of any machine, contractor, or process.

ACKNOWLEDGEMENTS

The authors would like to express their appreciation to Mr. C. Dwight Hixon, Mr. Curt Hayes, Mr. Larry Senkowski, and Ms. Jennifer Koscelny of the Research Division, Oklahoma Department of Transportation (ODOT) for their valuable assistance and suggestions throughout the course of this research. Special thanks are due to Dr. James Nevels, Jr. and the drilling crew of the Materials Division of the ODOT for their assistance in procuring field test data used in this study.

The financial support for this study provided by the Federal Highway Administration (FHWA) in cooperation with the ODOT is gratefully acknowledged.

Finally, the authors would like to thank Dr. Kenneth Fridley, former Assistant Professor of Civil Engineering and Environmental Science, University of Oklahoma for his invaluable comments and suggestions at various stages of this study.

TABLE OF CONTENTS

III METHODOLOGY

IV LINEAR STATISTICAL MODELING

V NONLINEAR REGRESSION MODELS

VI CONCLUSIONS AND RECOMMENDATIONS

APPENDIX I

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Bridge approach settlement or "bump" at the ends of a bridge, which can be technically defined as the differential settlement between the bridge deck and the approach pavement, is a well recognized and widespread problem encountered by the transportation agencies across the United States including Oklahoma. Passenger discomfort, traffic accidents and delay, premature failure of bridge (due to excessive impact loads) and excessive maintenance cost are a few examples demonstrating the severity of this problem. A study based on statistical methods was carried out to identify the influential causative factors/parameters contributing to bridge approach settlements. Data pertaining to total settlement and the various causative factors were obtained for 26 bridge sites at various locations in Oklahoma. The causative factors considered in this study are: height of embankment, age of approach, traffic volume, skewness of approach, foundation soil thickness and the site-specific embankment and foundation soil characteristics. The site-specific SPT and CPT values (e.g., N-value, tip resistance and friction ratio) were used for the quantitative representation of the embankment and foundation soil characteristics. The sites were done selected to encompass a wide variation of the aforementioned causative parameters. Soil properties evaluated in the laboratory are not included here because of the lack of uniformity in testing plans for different sites; however, they are used in deterministic analyses of site specific settlements.

Extensive statistical analyses were performed using a statistical package program, SAS. The following SAS Procedures were opted: REG, RSQUARE, STEPWISE and MODEL. Both linear and nonlinear multiple regression models were developed to identify problematic bridge sites and predict the approach settlements. The validity of the variables/parameters in the

models was judged based on their level of significance, partial R^2 (square of coefficient of correlation), t- and F- statistics; while the best models were decided on the basis of their overall $R^2_{(adjusted)}$, Mallow's C_p and their goodness of fit. The goodness of fit of the models was assessed based on their predictive capabilities and the analysis of the residuals of the predicted values. The special "Field Test Model" was also developed with the aim of identifying the problematic bridge sites prior to the construction of the bridge. The variables used in the field test model can be easily obtained from the Standard Penetration and Electric Cone Penetration tests.

The linear model was found to be more effective in identifying problematic sites as well as estimating approach settlements, compared to the nonlinear model. The field test model seems to be more efficacious in identifying the problematic sites than estimating the approach settlements. The coefficient of correlation, $R^2_{(adjusted)}$, varied from about 0.88 to 0.94. It was observed that the site-specific SPT (N-value), cone resistance (q_c), friction ratio and the thickness of the foundation soil have significant impact on estimating the approach settlements as compared to other variables of the models, indicating that the site-specific embankment and foundation soil characteristics together constitute the most influential causative factor. However, it should be noted that these characteristics did not include the assessment of dispersivity and erodability which leads to a void-domain matrix different than the one as-constricted, which in turn may account for significant portion of the settlement.

CHAPTER I

INTRODUCTION

1.1 General

Settlement of bridge approach pavements near abutments often leads to abrupt grade differences at the interface of abutments and the approach pavements (Fig. 1.1). These grade differences, also known as "bumps" at the ends of a bridge, give rise to driver discomfort and potentially unsafe driving conditions, premature failure of bridge deck and/or bridge itself due to repeated impact loads, and above all, costly and perpetual maintenance work that usually tends to impede the normal flow of traffic, especially in a high-speed, high-volume type highway. The bridge approach settlement has been a persistent and widespread problem facing transportation agencies, namely DOT's and FHWA, across the United States for many years. In order to reduce bumps at bridge ends frequent maintenance procedures in the form of mud-jacking (for concrete approaches) or patching (for both concrete and asphalt concrete approaches) are performed which provide only temporary remedies to the problem. Sometimes these procedures may result in detrimental consequences, as in the case of mud-jacking causing cracking of the approach slab. Although the bridge approach settlement has been a well recognized and widespread problem for a long time, only a few systematic and comprehensive studies have been undertaken in the past in this field. Most of the previous investigations focused only on certain specific aspects of the problem such as effect of some selected fill material on settlement, or of drainage conditions on approach settlement.

Cognizant of the extent of this problem in Oklahoma, the Department of Transportation (ODOT) commissioned the University of Oklahoma to undertake a systematic and comprehensive

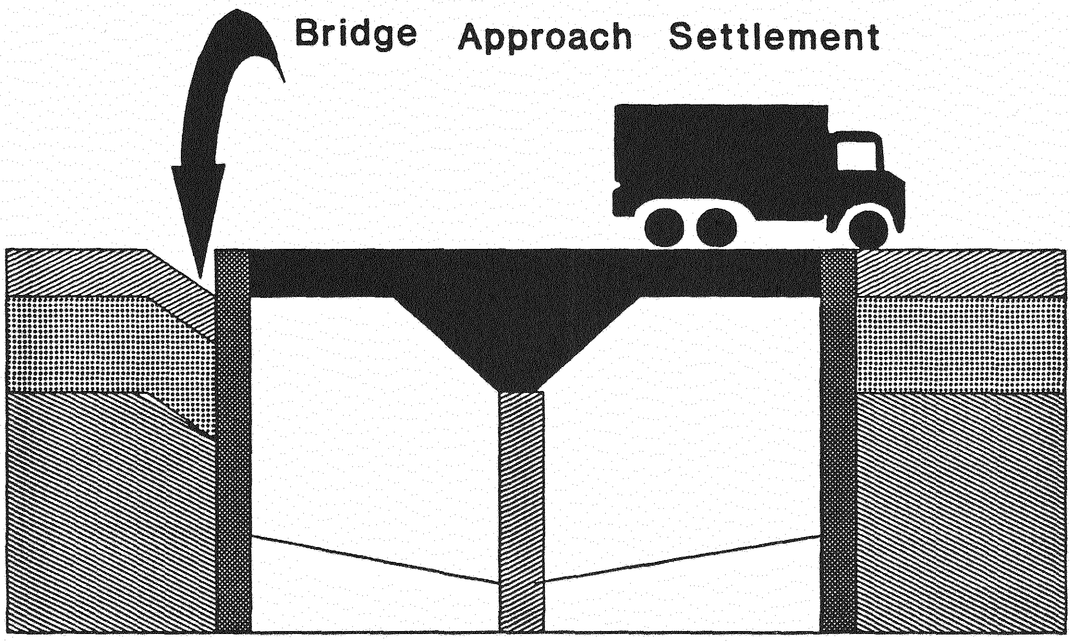


Fig. 1.1 Schematic diagram of bridge approach settlement.

and widespread problem for a long time, only a few systematic and comprehensive studies have been undertaken in the past in this field. Most of the previous investigations focused only on certain specific aspects of the problem such as effect of some selected fill material on settlement, or of drainage conditions on approach settlement.

Cognizant of the extent of this problem in Oklahoma, the Department of Transportation (ODOT) commissioned the University of Oklahoma to undertake a systematic and comprehensive research study to evaluate the causes of excessive approach settlement and to recommend appropriate remedial measures. To accomplish this goal, it was decided to conduct the study in five different phases as presented in Table 1.1.

1.2 Scope

The primary goal of this study is to develop statistical models for the quantitative characterization of the bridge approach settlement. The main objectives of developing such models are to

(i) predict approach settlement at a bridge site,

(ii) identify problematic bridge sites, and

(iii) examine the relative significance of various causative factors contributing to bridge approach settlement.

Both linear and nonlinear regression analyses were performed on the data obtained from level-one and level-two surveys. The relative significance of various research study to evaluate the causes of excessive approach settlement and to recommend appropriate remedial measures. To accomplish this goal, it was decided to conduct the study in five different phases as presented in Table 1.1.

Table 1.1 Various phases of the research.

Phase	Description
I	Literature review and survey of transportation agencies to investigate the extent of approach settlement problems in Oklahoma and other states.
II	Level-one survey of selected bridge approaches in Oklahoma for a qualitative assessment of causative factors.
III	Level-two survey of selected bridge approaches in Oklahoma for field testing and collecting samples for laboratory testing.
IV	Developing numerical and statistical models for prediction of approach settlements and identification of problematic sites.
V	Developing guidelines for design, construction, and maintenance of bridge approaches.

1.2 Scope

The primary goal of this study is to develop statistical models for the quantitative characterization of the bridge approach settlement. The main objectives of developing such models are to

(a) predict approach settlement at a bridge site;

(b) identify problematic bridge sites, and

(c) examine the relative significance of various causative factors contributing to bridge approach settlement.

Both linear and nonlinear regression analyses were performed on the data obtained from level-one and level-two surveys. The relative significance of various causative factors was examined, and statistical models were developed based on the significant causative factors. Laboratory test results are not incorporated in the current analysis, rather data collected from field tests (e.g., SPT and CPT tests) are exclusively utilized in developing the statistical models. Moreover, the regression analyses are restricted to the use of quantitative variables (causative factors) only. Use of qualitative (category) variables is considered to be beyond the scope of the current study. The site-specific SPT (N-value) values, tip resistance and friction ratio are considered to be the quantitative representation of the embankment and foundation soil characteristics.

1.3 Literature Review

As presented in Table 1.1, the current study belongs to Task (4) of the project. A brief overview of the previous phases is presented here for clarity and completeness.

1.3.1 Overview of Phases I, II and III

Phase I (Laguros et al., 1986) of this research work was conducted during the period of May 1985 - February 1986 and was mainly devoted to the collection of pertinent information associated with the bridge approach settlement problems. In order to achieve this goal, the following tasks were identified: (1) comprehensive literature search, (2) survey of various state, federal and private agencies involved in the construction and/or maintenance of bridges, and (3) analysis of the survey responses. In Task (1), an extensive literature search on the referenced problem was conducted by utilizing the computer search facilities of the Highway Research Information Service (HRIS) and other data bases such as DIALOG, ORBIT and BRS available at the University of Oklahoma. Manual searches were also conducted in this process.

In connection with Task (2), a questionnaire was prepared, in consultation with ODOT, and was sent to 52 state DOTs and 36 US Corps of Engineers Districts, as well as few other private agencies associated with the design construction and maintenance of bridges and highways. Responses were received from 61 agencies, out of which 42 agencies (approximately 70%) reported to have significant or very significant approach settlement problems. The responses of the questionnaire also revealed that only 6 states (California, Iowa, Kentucky, Louisiana, Ohio and Texas) had undertaken some sort of research to investigate this problem (Hopkins et al., 1969, 1985; Timmerman, 1976). However, several other states (viz. Colorado, North Carolina, Washington, Wyoming and Maryland) have recently been involved in investigating certain specific aspects of the approach settlement problems (Wolde-Tinasae et al., 1987; Kramer et al., 1991).

The findings of Phase I conclusively demonstrated that the bridge approach settlement problems are quite extensive in many areas of the United States, including Oklahoma. An in-depth understanding of the settlement process and the various causative factors are essential for finding any remedial measures.

Phase II (Laguros et al., 1990) started in February 1987 and was completed in December 1989. The major tasks in this phase of the research work involved: (1) selection of bridge sites for level-one survey, (2) survey of those selected bridge sites, and (3) characterization of approach pavement settlement in the state. A total of 381 bridge sites, scattered in 77 counties of Oklahoma, were selected in consultation with ODOT and were later surveyed to obtain data related to the following factors: (i) bridge, abutment and approach geometry, (ii) existing conditions of the approach, abutment headwall, slope protection structure, drainage, and embankment slope, and (iii) embankment material. In addition, information related to construction and maintenance history of the selected bridges was obtained by interviewing ODOT personnel and examining ODOT records. Based on an extensive analysis of survey data, the following conclusions were reported:

1. The problem of approach settlement is extensive in Oklahoma; and approximately 83% of the bridge approaches surveyed experienced settlement.
2. The problem seems to be more pronounced in the absence of any drainage in the fill behind abutment.
3. The long term performance of both rigid and flexible approaches are similar, however, in the short term performance, rigid approaches undergo lower differential settlements.
4. Pile supported abutments are more susceptible to approach settlements than the stub type.
5. The higher embankment height might be partly responsible for larger approach settlements.
6. Skewed approaches generally undergo larger settlements than nonskewed approaches.

7. A major portion of the settlement of approaches occurs within the first 20 years of the service life of the bridge approach.

Apart from these, regression analyses on data collected from the level-one survey were conducted to develop an empirical relationship between the bridge approach settlement and the various causative factors/parameters and are discussed in subsection 1.3.3 in detail. In addition, as a preliminary work for Phase III, level-two survey, a detailed field and laboratory testing program was conducted at two selected bridge sites with the aim of determining their site-specific embankment and foundation soil characteristics.

As the majority of data collected from the level-one survey was based on visual inspections of the selected bridge sites, only qualitative characterization of the approach settlement problem was possible in Phase II. For quantitative characterization, however, a level-two survey of the selected bridge sites were deemed necessary and it was conducted during Phase III of this research work.

Phase III (Laguros et al., 1991) started by conducting the level-two survey of some selected bridge sites in Oklahoma. The main objectives of the level-two survey were to:

(a) obtain comprehensive site-specific data, mainly pertaining to embankment and foundation soil characteristics, which can be used for a quantitative characterization of the causes and mechanisms of approach settlements at these sites;

(b) acquire data for validation of the numerical model for the prediction of approach pavement settlement; and

(c) build a data-base which can be used for assessment and/or estimate of approach settlement at similar sites.

The site-specific properties related to compaction, creep, consolidation and drainage, among many others, were considered to be necessary for the quantitative characterization of approach settlement. In order to achieve the objectives of Phase III, the following tasks were identified:

(1) Selection of bridge approaches for level-two survey.

(2) Collection of soil samples from the selected sites and laboratory testing.

(3) Instrumentation and monitoring of selected sites.

(4) Quantitative characterization of approach settlement causes and mechanisms based on the survey data.

As a part of the level-two survey, it was proposed to include field instrumentation and monitoring of approach settlements at selected sites, Task (3), so as to evaluate the significance of creep movements of alluvium soil deposits in the overall settlement problem. Owing to financial constraints, however, this task could not be carried out.

The quantitative characterization of approach settlement defined in Task (4) involved a comparison and correlation of data collected from the field and laboratory tests. It was originally planned to conduct statistical analysis to obtain the desired correlations. It was estimated that the test results from at least 20 bridge sites would be required to obtain meaningful correlations from the statistical analysis. However, in the course of that study, it was possible to obtain the test results from only 9 bridge sites. Since information from such a limited number of sites was judged to be inadequate, the statistical analysis could not be conducted. Nevertheless, the field and laboratory test results for each site were analyzed thoroughly to identify the site-specific causative factors of the bridge approach settlement. As a part of this analysis, the potential for consolidation settlement of the approach foundations was evaluated for sites having predominantly clayey soils. For this purpose, a numerical model was developed based on the nonlinear finite element method (FEM) and used to predict the consolidation settlement of the approach foundation due to embankment loads (Zaman et al., 1991).

Apart from the aforementioned tasks, an additional task of finding an appropriate definition for the terms "excessive settlement" or "bump" at the ends of a bridge was undertaken as a part of Phase III; and the findings of this task were reported separately to ODOT (Laguros et al., 1990). The report concluded that the approach settlement of 2 inches or more can be considered as an "excessive settlement" or "bump". Approach settlement of such amounts was considered hazardous to traffic safety and needed immediate maintenance. An approach settlement of 0 to 0.5 inch was considered as minor, while that of 0.5 to 1 inch was considered as moderate.

1.3.2 Investigations by Other Transportation Agencies

Several other transportation agencies across the United States have been involved in investigating bridge approach settlement problems. A brief review of these investigations are presented here in a chronological order. Moreover, the findings of certain investigations reviewed by Mahmood (1986, 1990) are also presented below in more detail.

Southern California Investigation (1959)

Jones (1959) conducted an investigation on four highway systems in the Los Angeles, California area to explore the relationship between bridge approach settlement and soil conditions at a site. Data pertaining to construction and approach characteristics were correlated to draw conclusions. The main causative factors contributing to the approach settlement were found to be:

(i) compression of the approach embankment;

(ii) consolidation of compressible foundation soils underlying the approach embankment; and

(iii) insufficient compaction of backfill or approach embankment, rapid construction, etc.

Jones (1959), on the basis of the observations of approach performance in California, suggested the following measures to reduce the approach settlement problem:

1) use of high quality backfill;

2) preloading of fill in case of rapid construction; and

3) removal and replacement of incompetent foundation soil, if economically feasible.

Kentucky Investigations (1969-1985)

Hopkins (1969) initiated a long term investigation of the bridge approach settlement problems in Kentucky. Seven hundred and eighty two bridge sites were surveyed of which six were selected for long-term study by installing instrumentations. The research continued for about a period of 20 years and the findings were published in reports by Hopkins (1969, 1973, 1985), Hopkins and Scott (1969), and Hopkins and Deen (1969). The following causative factors were identified to be the major contributors of approach settlements:

(i) consolidation settlement of approach foundation;

(ii) settlement of embankment due to the presence of compressible fill materials;

(iii) inadequate compaction of the approach embankment;

(iv) secondary compression and shear strain of the approach embankment and foundation;

(v) the local geology;

(vi) material erosion around abutment and approach pavement; and

(vii) lateral and vertical creep deformations of the bridge approach embankment and foundation soils.

The investigators also concluded that traffic volume has no significant effect on approach settlement and proposed the following mitigating measures:

a) Preconsolidation of compressible foundation soils by (i) using surcharge fill, (ii) the use of wick or sand drains acting alone or in combination with a surcharge, (iii) removal and replacement of incompetent material, and (iv) the use of lightweight fills.

b) Strict compliance of special compaction provisions.

c) Provision of proper drainage system.

Ohio Investigations (1976)

In a two-phase study sponsored by the Ohio Department of Transportation, Timmerman (1976) presented the results of an investigation of bridge approach design and construction practices in Ohio. The first phase basically involved a literature survey and interviews with bridge engineers, while the second phase dealt with the conditions of existing bridges throughout the state of Ohio based on three surveys conducted in 1961, 1974, and 1975.

The key factors of the bridge approach settlement identified by Timmerman were pavement thermal expansion, creep-induced lateral soil movements, abutment type, and fill characteristics. Regression analyses were performed on the data obtained from the bridge survey to find the correlation between approach embankment performance and various design-construction parameters; however no meaningful correlation was obtained. Some general remarks concerning the approach settlement were made based on the statistical analyses.

California Investigation (1985)

Stewart (1985) conducted an investigation, sponsored by the California Department of Transportation (CalTrans), in an attempt to relate the rough transition observed at bridge ends in California with various approach slab parameters such as age, fill height, abutment skewness, geographical region, average daily traffic, pavement type, and abutment type. Data from 820 bridges were obtained and statistical analyses were performed. Some of the important conclusions of the study were as follows:

a) Consolidation of embankment and foundation soils, poor quality embankment materials and improper compaction were key factors causing settlements of approach pavements.

b) Age of the approach and the geographical region in which the approach was situated were found to influence the approach settlement, whereas factors such as fill height, traffic volume, skewness of the approach, settlement period, ingress or egress approach end, and length of the approach slab were found to have negligible effects.

Colorado Investigations (1987)

In a research project sponsored by the Colorado Department of Highways (CDOH), Ardani (1987), based on the analysis of field and laboratory data from 20 bridge approaches in Colorado, concluded that the following factors were mostly responsible for pavement settlement at bridge approaches:

(i) consolidation of approach embankment and foundation soils;

(ii) inadequate compaction of backfill and embankment materials; and

(iii) poor drainage causing erosion at the abutment face.

Bridge approach settlement was correlated with SPT blow counts (N-values) of the backfill and the embankment in an attempt to attribute approach settlement to the poor compaction.

Maryland Investigations (1987)

Wolde-Tinsae et al. (1987) of the University of Maryland completed Phase I of a study sponsored by the Maryland Department of Transportation. The study included a literature search and a survey of state highway departments throughout the United States and overseas. The researchers concluded that the bridge approach was mainly caused by the following factors:

(1) consolidation of embankment foundation soils and volume change of the approach embankment;

(2) abutment movement due to slope failure, seepage, thermal forces, etc.; and

(3) type of abutment.

Wyoming Investigations (1989)

Edgar et al. (1989) reported the effectiveness of geotextiles in eliminating approach settlements of bridges in Wyoming. Geosynthetic Reinforced Walls (GRS) were erected behind bridge abutments and the performance of the approaches was monitored. It was found that none of the 90 approaches (45 bridges) required any maintenance in 5 years of usage. In addition, tests were conducted to evaluate the effects of lateral pressure on the abutment with GRS walls. It was concluded that the use of GRS limited the lateral deformations and consequently reduced approach settlement.

NCHRP Synthesis Report (1989)

This report was an updated version of the previous NCHPR Synthesis report (1969) and presented the new materials and techniques in the field of bridge embankment construction that had been developed over the last 20 years. The foundation and approach materials were identified to be the leading factors of bridge approach settlement, and the need for strict adherence to the standard specifications and procedures in embankment construction was emphasized.

Nebraska Investigations (1989)

Tadros et al. (1989), in a report to the Nebraska Department of Roads (NDR), presented the findings of Phase I of their research work concerning the approach settlement problems in Nebraska. Based on the information obtained from literature review, survey and inspections of 53 bridge sites situated in eastern Nebraska, the researchers proposed a number of potential causes of bridge approach settlement and their possible remedies. Some of the potential causes identified were:

a) differential settlement of two dissimilar systems viz. embankment and the abutment (the former being free to settle, while the latter being restrained to move);

b) consolidation of underlying foundation;

c) embankment stability; and

d) erosion on the abutment backwall.

Preconsolidation using surcharges, waiting periods, removal and replacement of weak soils, provisions of wick drains and benching were suggested as the mitigating measures to such settlements.

1.3.3 Statistical Modeling

During Phase II of the research work by Laguros et al. (1990), extensive statistical analyses were performed on the data obtained from Level I survey of 680 bridge approaches of Oklahoma (see Sec. 1.3.1). Various SAS procedures were used to develop an empirical relationship between the approach settlement and the various causative factors, namely, age of the approach, embankment height, traffic count, skewness of the approach, foundation soil depth, approach type, geologic unit, and foundation soil type. First the GLM (General Linear Model) procedure was used to assess the relationship of each factor/parameter separately with the settlement. All possible relationships, i.e., linear, quadratic, cubic, logarithmic, exponential, etc., were investigated. The significant relationships were retained for a multiple regression model using the STEPWISE procedure. The following generic model was used for the multiple regression analysis:

$$TSET = F\ (AGE, AGE2, AGE3, LAGE, EHGT, EHGT2, EHGT3, LEHGT, EAGE, EEHGT, SKEW, SKEW2, SKEW3, TRAFFIC, TRAFFIC2, ESKEW, LSKEW, LTRAFFIC)$$

where

TSET = total approach settlement in inches;

AGE = age of the approach in years;

AGE2 = AGE^2; AGE3 = AGE^3; LAGE = Log (AGE); EAGE = e^{AGE};

EHGT = embankment height in ft.;

EHGT2 = $EHGT^2$; EHGT3 = $EHGT^3$; LEHGT = Log (EHGT); EEHGT = e^{EHGT};

SKEW = skewness of the approach in degree;

SKEW2 = $SKEW^2$; SKEW3 = $SKEW^3$; LSKEW = Log (SKEW); ESKEW = e^{SKEW};

TRAFFIC = average daily traffic (number); and

TRAFFIC2 = $TRAFFIC^2$; LTRAFFIC = Log (TRAFFIC).

Out of the 18 regressors, only the significant variables were retained in the model. Based on the criteria of significance level of variables and the coefficient of correlation, R^2, the following models were proposed:

For Flexible Pavement: **($R^2 = 0.496$)**

$$TSET = .000011\ AGE^3 + .639760\ Log\ (AGE) - .0000378\ (EHGT)^3 + .323710\ Log\ (EHGT) - .004373\ (SKEW) + .008223\ Log\ (TRAFFIC) + .002497\ (AGE \times EHGT) \tag{1.1}$$

For Rigid Pavement with asphalt overlay: **($R^2 = 0.673$)**

$$TSET = .000150\ AGE^3 - 4.206597\ Log\ (AGE) - .000015\ (EHGT)^3 + 2.658108\ Log\ (EHGT) + .029693\ (SKEW) - .000913\ (AGE \times EHGT) + .606243\ (AGE \times EHGT) \tag{1.2}$$

For Rigid approach: ***($R^2 = 0.605$)***

$$TSET = .000032\ AGE^3 - .000003\ EHGT^3 - .079417\ Log\ (EHGT) + .010869\ (SKEW) + .069695\ Log\ (TRAFFIC) + .003868\ (AGE \times EHGT) - .022570\ Log\ (AGE \times TRAFFIC) \tag{1.3}$$

The models are based on the data from 680 bridge approaches. It is observed that in spite of incorporating such a large number of samples, a maximum R^2 of 0.673 could only be achieved. No details regarding the predictive ability of the above mentioned models were presented. The poor correlations of these models were attributed to the absence of any factors/variables related to the embankment and characteristics of the foundation soils in the models.

Mahmood (1990) covered the activities of Phases I and II (Laguros et al., 1986, 1990) of the research work. Various sophisticated statistical techniques were used to develop a probabilistic model to predict bridge approach settlement. After rigorous exploratory analysis and hypothesis testing on data pertaining to various causative factors, the following qualitative and quantitative variables were selected for model building processes:

Quantitative variables:

- age of the approach (years);
- embankment height (feet);
- thickness of foundation soil (feet);
- skewness of the approach (degrees); and
- traffic count (ADT);

Qualitative variables:

- approach type (Rigid, Flexible);
- abutment type (Stub, Pile-end-bent, Open-column);
- foundation soil type; and
- geologic unit (Gr.1-Gr.2-Gr.3-Gr.4).

Skewness of the approach was used as a qualitative variable because of the fact that 75% of the approaches had skewness of 0 degree.

Various SAS procedures were used in connection with the statistical analysis. GLM and RSQUARE procedures were used for exploratory analysis, whereas STEPWISE procedure was used for variables selection and model building. Both qualitative (dummy) and quantitative variables were incorporated into the model. The following regression model was proposed

$$TSET = -0.498X_1 - 0.445X_2 + 0.046\,(EHGT) + 1.186\,(LAGE) \qquad (1.4)$$

where

TSET = total settlement in inches

X_1 = 1, for skewness = 0

= 0, for skewness >0

X_2 = 1, for pile-end-bent abutment

= 0, otherwise.

The square of the coefficient of correlation, R^2, for the model was 0.5966, indicating that approximately 60% of the variation in settlement was explained by the model. Mahmood (1990) suggested that the 40% unexplained variation in the settlement might have been caused by not properly considering the embankment foundation soil characteristics.

It was suggested that variables pertaining to the embankment and foundation soil characteristics, obtained from level-two survey of the bridge sites, should be incorporated in order to develop a better regression model which was accomplished subsequently and is reported here.

1.4 Contents of the Report

The accuracy of the data plays an important role in any type of statistical analysis. A detailed description of data pertaining to the total settlement and its various causative factors, the process of acquisition, and the limitations of these data are presented in Chapter II. Chapter III deals with an overview of the multiple regression models and the various statistics used in connection with the model building. A brief description of the various SAS Procedures is also presented in Chapter III. A statistical based model can be very useful in identifying the problematic sites and/or estimating the approach settlement at a site. A linear multiple regression model developed to predict the bridge approach settlement is presented in Chapter IV, while Chapter V gives an account of various nonlinear models, including the "field test" model which best represent the data set under consideration. The conclusions and recommendations for further study are presented in Chapter VI.

CHAPTER II

ACQUISITION OF DATA

2.1 General

The data used in developing the statistical model(s) to identify problematic sites (i.e., sites with potential for excessive bridge approach settlements) were primarily obtained from the level-one and level-two surveys of the selected bridge sites in Oklahoma (Laguros et al., 1990, 1991). As mentioned in Chapter I, the primary objective of conducting the level-two survey was to obtain the desired site-specific embankment and foundation soil properties, which are suspected to be the dominant parameters/factors in causing the approach settlement problems. The sites were selected so that they are representative in terms of the various contributing factors, such as age of the approach, height of the embankment, skewness of the approach, traffic count, foundation soil characteristics. A description of the criteria used for site selection and field data collection (SPT and CPT tests) is presented in this chapter. Also, the extent of accuracy and the limitations of these data are discussed.

2.2 Criteria Used for Site Selection

As a part of Phase II of this research work (Laguros et al., 1990), a total of 361 bridges from all the 77 counties in Oklahoma were surveyed and information related to various features of the bridge was collected. A micro-computer oriented data-base (i.e., dBASE III Plus) was developed to store this information. The data was analyzed using a comprehensive statistical analysis package program (SAS) to establish the hierarchy of various causative factors of bridge approach settlement problems in Oklahoma. The data collected in the level-one survey were limited to field visits, review of records maintained by ODOT and interview of maintenance personnel, and did not involve any field and/or laboratory testing. The level-one data-base was utilized to select the sites for the level-two survey. The following factors were utilized: observed settlement, foundation soil type, composition of fill material, age of the bridge approach, embankment height, skewness of the bridge, approach pavement type, geographical distribution of the sites and maintenance frequency (Laguros et al., 1990). Based on the findings of the statistical analyses of level-one data, special attention was given to two selected factors (i.e., age of the approach and embankment height).

In the site selection process, all the bridge sites surveyed in Phase II were sorted with respect to different causative parameters and the critical sites were printed out for consideration. The computerized data-base was capable of searching the entries or input data based on an individual factor or a collective set of factors. This feature was very useful for ranking the bridge sites according to the causative factors.

Following the procedures outlined above, a total of 48 bridge sites (Laguros et al., 1991) in the eight Divisions in Oklahoma were initially selected as possible candidates for the level-two survey, in consultation with Oklahoma Department of Transportation (ODOT). However, later it was realized, in cooperation with ODOT, that 20 to 30 sites out of 48 selected sites might be realistic to achieve the stated objectives of the project within the time constraints. The present study is based on the results of the level-two survey of 29 sites. The 29 bridge sites considered in the study are listed in Table 2.1. It is important to note that owing to some practical difficulties like accessibility of the bridge sites, extent of traffic hampering etc., and also because of the preference for clayey sites over sandy sites, all the 29 sites, mentioned in Table 2.1 do not belong to the original list of 48 bridge sites.

Table 2.1 List of the 29 sites at which the level-two survey is completed.

Site #	Bridge #	County
1	I-40 20-02 X 1072E	Custer
2$^+$	US270 30-12 X 0849	Harper
3	SH152 75-10 X 0849	Washita
4	SH152 75-08 X 2077	Washita
5	SH9 14-11 X 1242	Cleveland
6$^+$	US270 04-20 X 1897	Beaver
7	SH102 41-38 X 1250	Lincoln
8	SH142 10-35 X 0376	Carter
9	SH003 64-12 X 0735	Pushmataha
10	I-35 36-25 X 1241E	Kay
11	SH17A 25-53 X 0087	Garvin
12	US69 18-10 X 0348	Craig
13	SH10 58-24 X 1287	Ottawa
14	US70 12-02 X 1078	Choctaw
15	US59 68-02 X 0000	Sequeah
16	SH20 66-08 X 0674	Rogers
17	US177 41-20 X 0611	Lincoln
18	US62 38-03 X 0213	Kiowa
19	US75 56-04 X 0113	Okmulgee
20	US183 75-06 X 0501	Washita
21	County Bridge (56th. st.)	Tulsa
22	SH15 23-20 X 0922	Ellis
23	US64 30-04 X 1825	Harper
24	US64 76-06 X 0545	Woods
25	SH8 47-18 X 1505	Major
26	SH51 70-85 X 0997	Tulsa
27$^+$	SH125 58-35 X 0942	Ottawa
28	SH10 58-24 X 0831	Ottawa
29	SH56 13-02 X 3314	Cimarron

$^+$*CPT test was not conducted at these sites.*

Both problematic (large settlements and frequent maintenance) and non-problematic (insignificant or zero settlement and no maintenance) sites were included to represent the "worse" and the "best" possible scenarios. Out of 29 sites selected, 3 sites have very high settlements (equal to or greater than 10 inches), 7 sites have experienced very little settlements (0 to 2 inches) and 19 sites have experienced settlements varying from 2 inches to 10 inches.

The embankment and foundation soil characteristics are represented by SPT (N- values), tip resistance, and friction ratio. The SPT value is obtained from the Standard Penetration Test (SPT) while the tip resistance and friction ratio are obtained from the Cone Penetration Test (CPT). The SPT values (weighted average value) range from as low as 3 to as high as 25 representing 'predominantly clayey' to 'predominantly sandy' sites. In a broader sense, tip resistance and friction ratio values obtained from a CPT test and the SPT-values from a SPT test have similar physical significance because they both represent resistance to penetration. To study the impact of embankment height, the selection included heights varying from 8 ft. to 36 ft. The thickness of foundation soil ranges from a low value of 5 ft. to a high of 72 ft. Similarly, bridges of different ages are included to consider the effect of age on approach settlements. The sites are selected from relatively new bridges (minimum 5 year old) to very old bridges (maximum of 57 year old). The effect of traffic count (average daily traffic) is taken into account in the selection by including sites experiencing ADT varying from 450 (low volume traffic) to 13,300 (high volume traffic). In other words, bridges falling under different categories namely interstate bridges, state highway bridges and county bridges were considered which indirectly attributed to the variation of the traffic count.

The process of acquisition of data for each of the above mentioned quantitative causative parameters as well as their extent of accuracy and limitations are discussed in detail in the following sub-sections.

2.3 Data Collection

There are mainly three sources from where the data used in the current study were obtained. They are:

- Level-one Survey
- Level-two Survey
- ODOT Bridge Division Records

2.3.1 Level-One Survey

The level-one survey of the 29 bridge sites was conducted concomitant with other bridge sites (a total of 361 bridge sites) during Phase II (Laguros et al., 1990) of this research project. The main strategies adopted to obtain the information were:

a) field visits to the bridge approach site;

b) interview of maintenance personnel; and

c) verification of the data from bridge records maintained at ODOT Bridge Division.

The following data used in this study were obtained from the level-one survey:

- total settlement;
- height of the embankment;
- traffic count; and
- age of the approach embankment.

Each of these factors is defined in context with the current study and their extent of accuracy is also discussed below.

1. Total Settlement

This is a quantitative estimate of the settlement which the bridge approach has experienced since its opening to traffic, expressed in inches. It has been estimated by one or more than one of the following methods:

- measuring movement of the curb (see Fig. 2.1);
- measuring thickness of cumulative overlay or thickness of cumulative patching (see Fig. 2.2);
- examining other noticeable evidences at the site; and
- interviewing maintenance personnel.

The differential settlement between the curb on the bridge and that on the approach was measured to obtain the total settlement the approach has experienced (Fig. 2.1). The thickness of level patching or overlay on the original pavement was also used to estimate the settlement at some sites (Fig. 2.2). For the sites where the pavement was mudjacked, the number of mudjacking operations was estimated from the evidence in the field and the settlement was estimated based on the total number of mudjacking operations. At the sites where the approach surface was nonuniform, the maximum settlement at any point on the approach pavement was recorded as the total settlement for that site. Interview of maintenance personnel of ODOT was conducted with the aim of obtaining the maintenance history of the bridge sites surveyed which

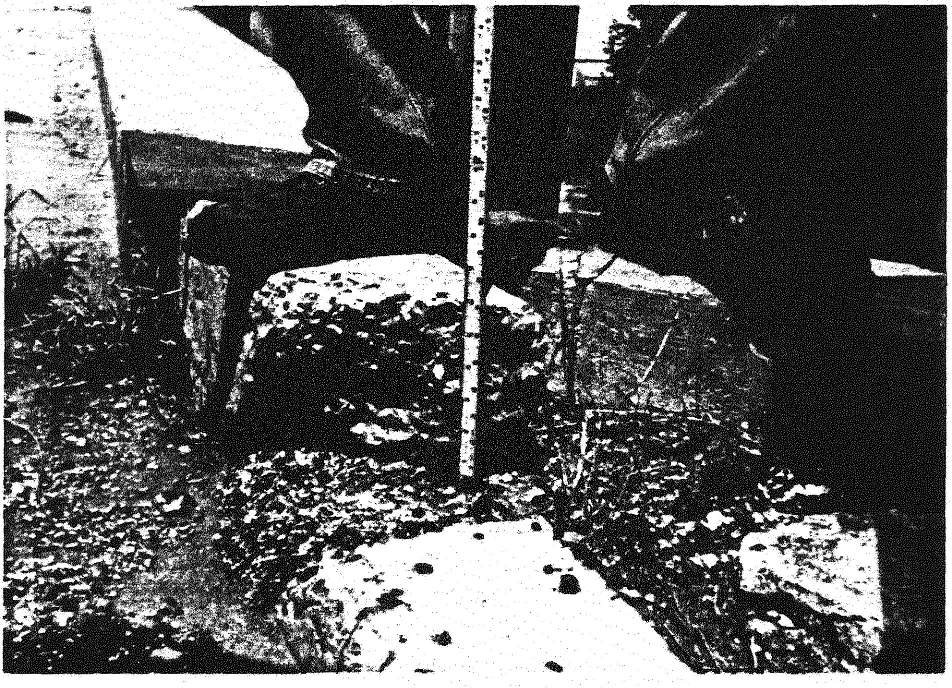


Fig. 2.1 Settlement estimate of bridge approach from the differential settlement of the curb.

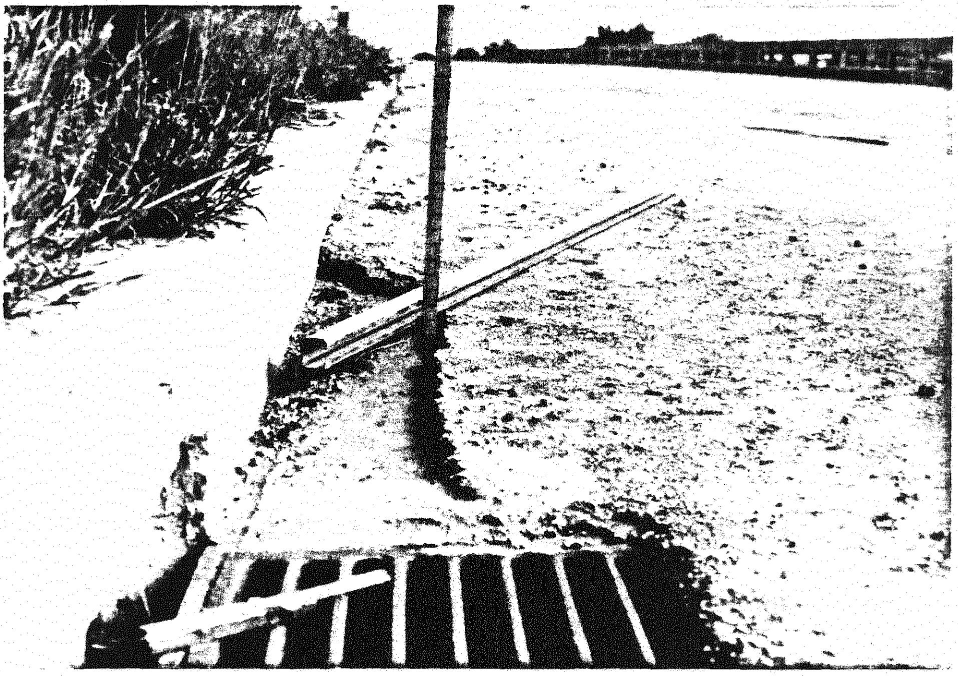


Fig. 2.2 Settlement estimate of bridge approach from the thickness of overlay or level patching.

proved very useful in estimating the total settlement. ODOT has 77 district maintenance units under its eight divisions. A special interview form (Zaman et al., 1993) was prepared and the district maintenance supervisors as well as the interstate unit supervisors along with their other related staff were interviewed. The estimate of settlements obtained through this process possess a number of limitations which are discussed in a separate section (see Sec. 2.4.1).

2. Height of Embankment

The height of embankment is the level difference between the top of the embankment and the original ground surface (see Fig. 2.3) expressed in feet. The estimated value was verified from the level-two survey as well as from the records maintained at the ODOT Bridge Division.

3. Age of the Approach/Embankment

This represents the time in years since the bridge was first opened to traffic. During the level-one survey types of the approach guardrail, general condition of the bridge and/or bridge structure type were used to roughly estimate the age of the approach. However, the accurate age of the bridge was obtained from the ODOT Bridge Division records.

4. Traffic Count

The traffic count represents the number of average daily traffic (ADT) on a specific approach and was obtained from the records maintained at the ODOT Bridge Division.

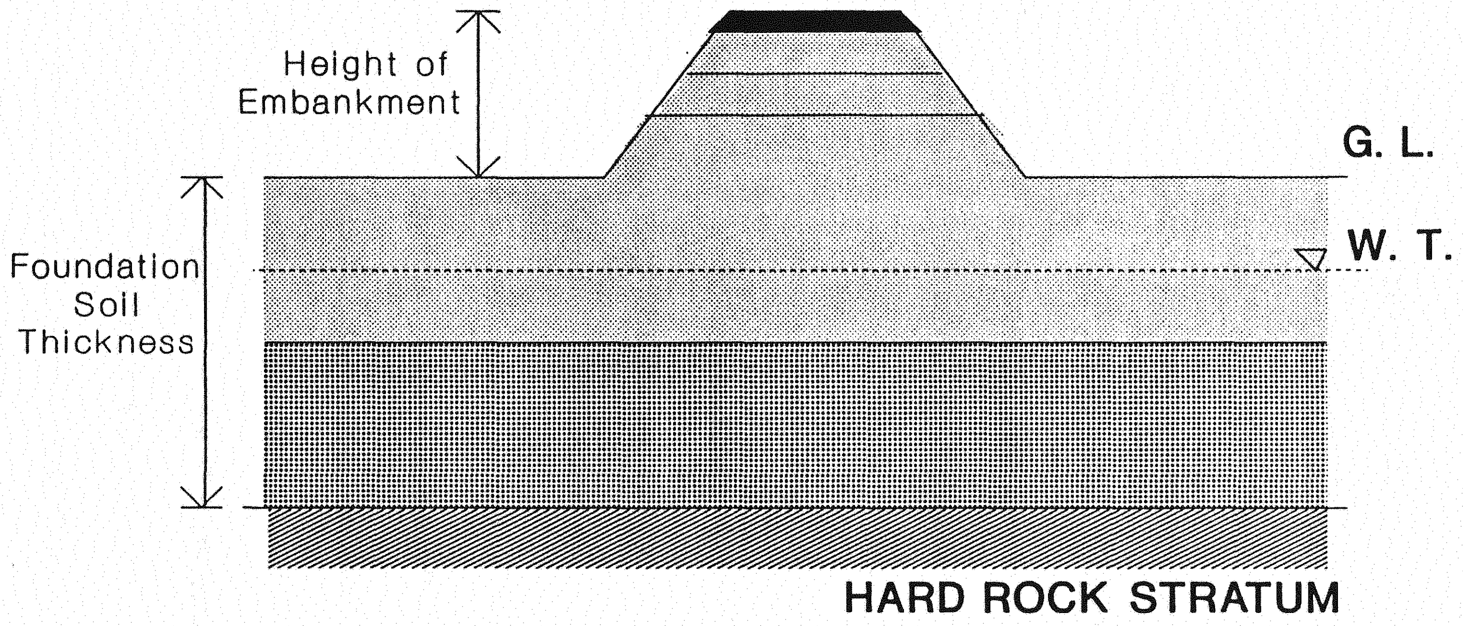


Fig. 2.3 Cross-section of the embankment and the underlying foundation soil near the bridge approach.

2.3.2 Level-Two Survey/Field Testing

The level-two survey involves a detailed field and laboratory testing program which can be utilized for the quantitative characterization of an approach site. It was suggested (Laguros et al., 1990) that the site-specific embankment and foundation soil characteristics, which is obtained from the level-two survey, can be used for a quantitative assessment or characterization of causes and mechanisms of approach settlements at a site. Accordingly, in the present study, results obtained from field testing are used exclusively to incorporate the effects of the embankment and foundation soil characteristics. Even though an extensive laboratory testing of soil samples obtained from all the 29 sites was completed as a part of Phase III of this research project, the results were not used in the present study. The main reasons being:

- all the tests could not be performed for all types of soils and
- push-tube samples could not be obtained for predominantly sandy sites.

Thus, it was not possible to form a uniform data set for all the sites needed for the statistical analyses. Moreover, the statistical model developed in the present study was not intended to be time-dependent, therefore the time-dependent tests, such as consolidation test were not considered to be very useful in the present context. Therefore, laboratory testing will not be discussed here onwards; however, field tests are discussed in detail. The site-specific data of SPT (N-value), tip resistance, friction ratio and thickness of the foundation soil were obtained from field testing which consists of Standard Penetration Tests and Electric Cone Penetration Tests. These tests are described below in detail. The thickness of foundation soil is the depth of soil layers extending from the ground surface level to the hard rock stratum (see Fig. 2.3) and was determined from the depth of borehole drilled on the ground surface in the vicinity of the bridge approach (Borehole #1). Subsequently, it was verified from the depth of another borehole, drilled on the embankment (Borehole #2), by subtracting the height of the embankment from it. However, the foundation soil thicknesses obtained from the two boreholes for Site #20 (Bridge #US183 75-03 X 0541, Washita Co.) were not identical. At Borehole #1, the hard bed rock was encountered at a depth of 40 ft., indicating that the thickness of foundation soil is 40 ft. for this site, while the other borehole (Borehole #2) indicated a foundation soil thickness of 75 ft. This indicates sloping type of ground in the vicinity of this bridge site. Therefore, the average of these two values (i.e., 55 ft.) was taken as the thickness of the foundation for this site. The heights of embankment initially obtained from field survey during Phase II of this project were also verified from the field test results.

The Standard Penetration Test (SPT)

Standard Penetration Test is a very common field test in Oklahoma. The SPT tests were conducted at all 29 sites. In general, two boreholes were drilled, one on the original ground surface (foundation soil) in the vicinity of the approach and the other on the approach slab close to the abutment (see Fig. 2.4). At one bridge site namely US64 30-40 X 1825 in Harper County, only one borehole was drilled on the embankment. The second borehole was not drilled because of the predominantly sandy nature of the site. At both boreholes, continuous soil sampling was done concomitant with the Standard Penetration tests (SPT). The SPT is a measure of standard penetration resistance measured in number of blows per foot (N) and consists of driving a split-spoon sampler into the soil using a 140 lb. of dynamic weight dropped from a height of 30 inches. The results of SPT were primarily used for qualitative interpretation

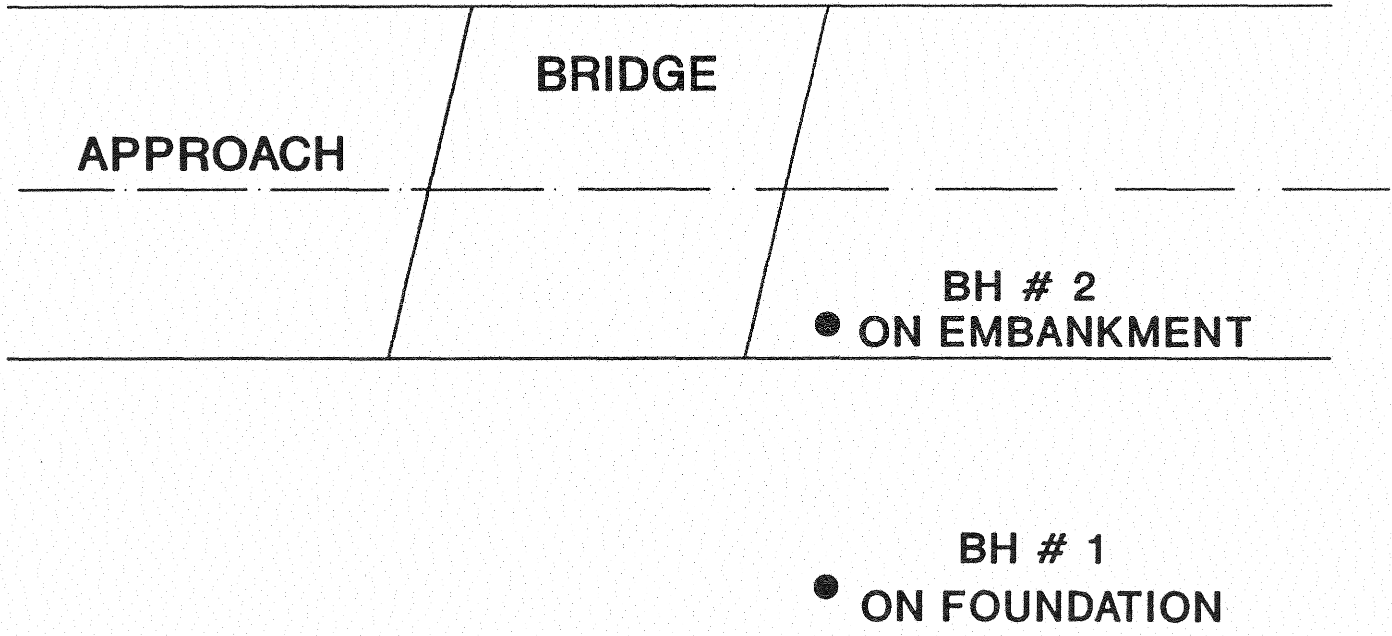


Fig. 2.4 Plan view showing the location of the boreholes.

of soil parameters; however, in the present study these results have been directly used as quantitative model variables to predict the bridge approach settlement (refer to Chapter III). The SPT results can also be used to determine the relative density of granular soils and the strength parameters of cohesive soils (Hunt, 1986), but the accuracy of such results are always questionable because of the fact that the accuracy of SPT results depends on several parameters including the overburden pressure, presence of water table, and pore pressure developed during testing. Many correction factors are, therefore, imposed on the raw SPT values to account for these parameters before proceeding for the qualitative assessment of the soils. Based on extensive field investigation (Hunt, 1986; Bowles, 1988), empirical relations have been developed for a qualitative characterization of both sands and clays based on SPT values (see Table 2.2).

Table 2.2 Empirical relationship of soil types with SPT values [after Hunt (1986)].

Sands		Clays	
N	Relative Density	N	Consistency
0 - 4	Very Loose	below 2	very soft
4 - 10	Loose	2 - 4	soft
10 - 30	Medium	4 - 8	medium
30 - 50	Dense	8 - 15	stiff
0ver 50	Very Dense	15 - 30	very stiff
		over 30	hard

In the present study, SPT values are used as quantitative measures of the embankment and foundation soil characteristics and hence the site-specific SPT value is calculated as the weighted average of the SPT values of various soil layers at a site. The layers were determined from the soil samples obtained during the SPT test which are usually referred to as 'disturbed samples'. The determination and segregation of layers were primarily based on a visual inspection (color, texture, etc.) of the soil samples and personal judgement. These samples were also used for laboratory testing for soil classification (Atterberg limits and grain size distribution) purposes. The average value of SPT is first determined for each layer and then the weighted average value is calculated from these average values. For an illustration, let there be five layers at a site of depths h1, h2, h3, h4, and h5 and the average SPT values for these layers be N1, N2, N3, N4 and N5, respectively (see Fig. 2.5). The weighted average value of SPT, denoted by SPT, for the site is calculated from the expression given below:

$$SPT = \frac{N1*h1 + N2*h2 + N3*h3 + N4*h4 + N5*h5}{(h1 + h2 + h3 + h4 + h5)} \tag{2.1}$$

In the process of statistical analysis, it was decided to explore the effect of SPT value of the embankment and that of the foundation soil separately because the soil characteristics of embankment might be completely different from that of the foundation soil. Consequently, the composite SPT value may represent the overestimated or underestimated value of SPT depending upon the nature of the site. Following the procedure outlined above, the SPT value for the embankment, denoted by SPTE, and the SPT value for the foundation soil, denoted by SPTF were calculated for each site. The splitting of the SPT value into SPTE and SPTF was also

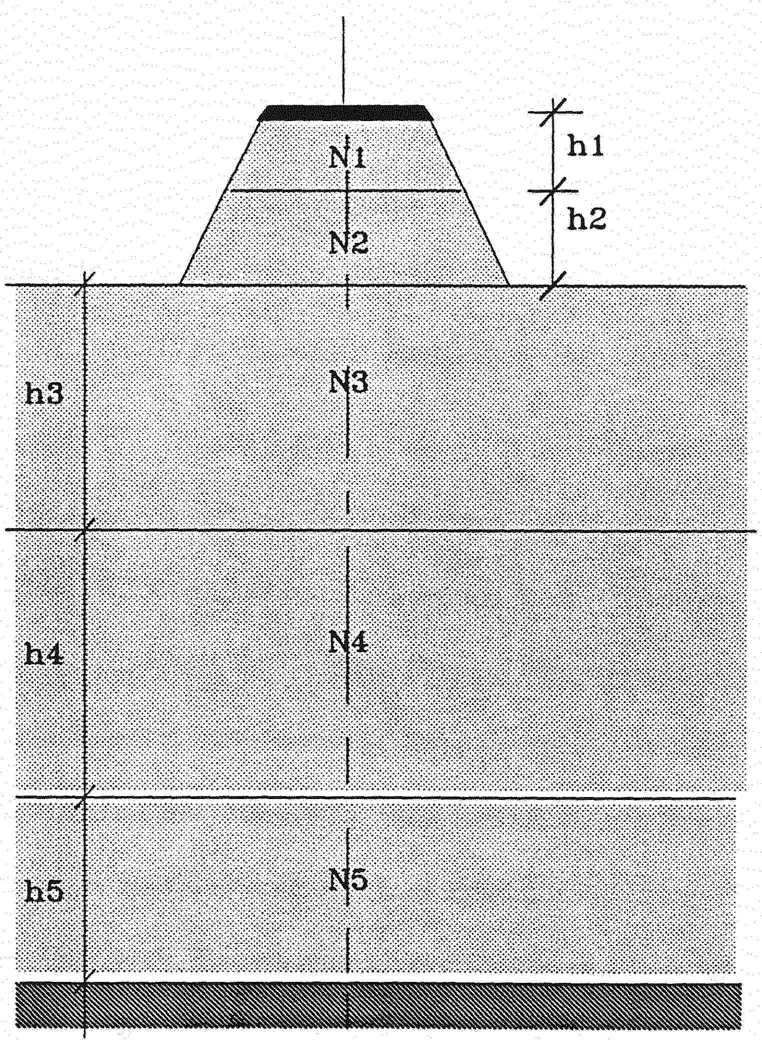


Fig. 2.5 Calculation of the weighted average value of SPT.

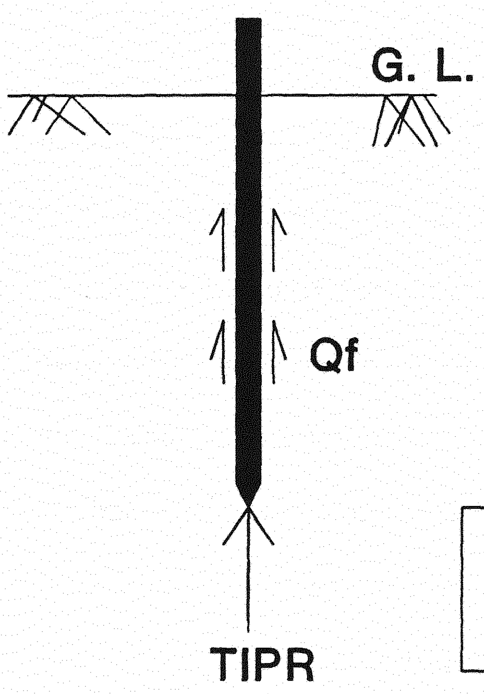
necessary to compare the relationship between the SPTF and the CPT values, since the CPT test was conducted only on the foundation soil and not on the embankment. Furthermore, in order to develop the "field test" model (refer to Sec. 5.3), the SPTF is required, since SPTE cannot be obtained at this stage.

During the SPT test, undisturbed samples were also collected at desired depths using a push-tube type sampler. Usually, undisturbed samples were collected for cohesive soils. These samples were sent to the ODOT Materials Division where they were extruded and coated with wax before being transported to the Geotechnical Laboratory at OU.

The Cone Penetration Test (CPT)

Out of 29 sites, Electric Cone Penetration Tests (CPT) were conducted at 26 sites only. The three bridge sites where CPT could not be conducted were: US270 30-20 X 0849 in Harper Co., US270 04-20 X 1897 in Beaver Co. and SH125 58-35 X 0942 in Ottawa Co. Thus, the statistical analysis is confined to the site-specific data available for the 26 bridge sites. At each site, a continuous electric cone penetration test (CPT) was conducted at the original ground (foundation soil) in the vicinity of the approach to obtain information pertaining to foundation soil characteristics viz. soil type, soil stratification, strength, and bed rock depth. This information proved very useful in planning the drilling activities. Figure 2.6 shows the schematic diagram of the CPT test.

The continuous cone penetration test (CPT test) basically involves pushing a steel cone and rods into the soil and monitoring the mobilized resistance to penetration in the soil (Robertson et al., 1984). The cone resistance, also called tip resistance (q_c), and the sleeve friction are monitored continuously during penetration via separate strain gauge and load cells



$$\boxed{FR = \frac{100 \times Qf}{TIPR}}$$

Fig. 2.6 Schematic diagram showing Cone Penetration Test CPT).

mounted on the cone and the friction sleeve. Signals from the strain gauges are relayed to the surface through electric cables running inside the Penetrometer rods. The friction ratio which is the ratio of sleeve friction to the cone resistance is interactively calculated by the computer using the expression represented by Eqn. (2.2).

$$FR = \frac{Sleeve\ Friction}{Cone\ Resistance}\ X\ 100 \tag{2.2}$$

The graphical variation of the aforementioned values (i.e., tip resistance, sleeve resistance and friction ratio) with depths is also produced at the site. The CPT results obtained from 26 sites and used in the present study are presented in a separate report (Zaman et al., 1993). The CPT test results can also be used to find the in-situ soil properties from available correlations (Robertson et al., 1984; Douglas et al., 1981). This test is economical , less time consuming and is capable of providing useful information related to foundation soil characteristics. For example, for clayey soils the CPT results can be used to determine soil classification, undrained shear strength C_u, sensitivity of clay S_t, volumetric compressibility n_w, consolidation or past pressure from the available correlations. However, in the present study, only the weighted average values of q_c and FR are used. To find the weighted average values, the soil profile is divided into a number of layers depending upon the variation of cone resistance (q_c) and the friction ratio (FR). The average values of q_c and FR are determined for each layer and then the weighted average values are calculated following a similar procedure as outlined for the calculation of weighted average SPT value.

Finally, the data acquired from the aforementioned methods are presented in Table 2.3.

Table 2.3 Data-set for 26 Bridge Sites of Oklahoma Used in the Current Study

SITE #	TST (in.)	TCN (ADT)	FND (ft.)	EHT (ft.)	AOE (yrs.)	SPT (wt.av.)	TIPR (tsf)	FR	SPT	SPTF	SKEW (deg.)
1	8.0	4800	12	25	34	11	33.78	1.41	14	6	0
3	18.0	1500	60	22	16	9	45.22	2.19	10	8	0
4	5.0	2500	72	12	21	5	17.91	2.30	9	4	40
5	6.0	5000	50	10	29	12	38.92	2.50	10	10	10
7	8.0	450	40	8	13	13	21.53	2.94	16	13	50
8	10.0	4700	15	32	34	23	44.32	6.56	9	48	0
9	1.0	3700	26	26	23	10	75.33	4.41	10	9	10
10	6.0	4550	31	35	32	11	31.66	2.16	8	13	50
11	2.8	650	45	20	22	16	72.65	1.53	9	20	40
12	0.0	3000	18	21	5	17	44.86	6.46	9	23	60
13	2.5	13300	5	33	35	11	37.80	3.88	11	9	30
14	2.5	1600	38	36	7	7	16.17	4.19	7	8	0
15	15.0	10000	46	30	24	25	147.51	0.98	33	18	0
16	3.0	4500	20	16	44	6	25.04	4.90	6	6	40
17	5.0	1900	50	13	17	6	24.70	1.93	6	5	60
18	2.0	2500	42	24	16	22	41.24	4.83	10	29	40
19	3.5	6100	9	27	27	10	74.63	4.50	10	11	60
20	2.5	1000	40	15	9	5	16.50	0.84	5	5	0
21	3.0	3100	28	15	56	13	42.45	3.90	8	14	0
22	5.0	1800	50	11	53	20	74.07	1.04	12	21	0
23	2.5	900	40	10	57	13	51.11	2.29	11	14	0
24	2.5	1300	21	13	18	9	62.50	3.72	5	11	0
25	1.5	1600	25	8	38	17	55.31	4.43	9	21	0
26	0.0	10000	30	12	12	13	121.24	1.65	5	16	0
28	6.0	2500	20	35	41	13	43.41	1.27	12	14	0
29	1.5	4100	25	26	57	14	88.00	3.42	10	21	45

2.4 Limitations of the Data Collected

Accuracy of data plays an important role in interpreting conclusions from a statistical study. Also, the reliability of a statistical model is highly dependent on the accuracy of the data used to develop such a model. Errors in data collection/representation may lead to unexpected and misleading results in the process of statistical analysis. This section discusses some of the limitations of the data used in the present study.

2.4.1 Limitations of Level I Survey Data

The data collected from the level-one survey are height of the embankment (EHT), traffic count (TCN), age of the bridge (AOE) and total settlement. The height of the embankment as well as the age of the approach were subsequently verified from the field testing and from the records maintained at the ODOT Bridge Division, respectively. Therefore, chances of inaccuracy for these data are minimal at best. Traffic count (ADT) was also obtained from the ODOT Bridge Division. However, the estimated value of total settlement could not be verified from other sources.

The number of mudjacking operations in case of rigid pavements was estimated by careful observation of evidences in the field. For some sites it was not possible to obtain a single estimate. Moreover, the movement of the approach slab, at some sites, could not be estimated because of mudjacking. The accuracy of the estimated settlements was checked from the information supplied by the ODOT Maintenance Personnel for some bridge sites but for the majority of the sites, particularly the older sites, the data remained unverified due to lack of information. The evidence of mudjacking, in some cases, might have been buried by the asphalt overlay thus leading to an underestimate of the total settlement; however, the effect of these uncertainties would have least impact on the accuracy of the data because very few bridges in Oklahoma are mudjacked.

So far as the flexible pavements are concerned, the chances of erroneous records regarding the total settlements are greater than in rigid pavements. A bridge approach that has been level patched a few times may not show such maintenance measures due to the burial of previous operations by a new overlay. This might have led to a greater underestimation of the total settlements. The settlements of the county bridges could not be verified because of lack of records for such bridges. Therefore, only one county bridge is included in the current study.

In order to minimize the errors involved in the estimate of the total settlements, maintenance information was obtained by interviewing the ODOT maintenance personnel and other related staff. But the information gathered through this process did not prove very useful for most of the sites under consideration because of the following reasons:

- nonexistence of proper maintenance records;
- information gathered were based on memory of the maintenance personnel;
- transfer of maintenance supervisors and replacement by new supervisors;
- change of agency responsible for bridge maintenance; and
- no maintenance records for county bridges.

2.4.2 Limitations of Level II Survey Data

The data obtained from the level-two survey of the sites pertain to foundation soil thickness, tip/cone resistance, friction ratio and SPT values. The foundation soil thickness for all the sites was obtained by measuring the borehole depth drilled on the foundation soil till hard rock stratum was encountered, and were subsequently verified from the depth of another borehole drilled on the embankment by subtracting the height of the embankment from it. However, in one case (refer to Sec. 2.3.2), the average of the two values was taken as the foundation soil thickness.

As mentioned in the previous sections, the weighted average values of SPT, SPTE, SPTF, q_c, and FR are used in this study. These are based on the average values of various layers from the soil profile at a site. The determination of the layer thickness in all the cases is primarily based on personal judgement. Determination of soil layers and a weighted average SPT value was also dictated by the boring logs and soil samples for laboratory testing. In case of CPT test results, it was extremely difficult to determine the boundaries of different layers precisely based on q_c and FR values. This was because the variation of these quantities within a short depth was sometime erratic. Thus, a large number of layers would be required in order to represent these values (q_c and FR) accurately. Therefore, the weighted average values may vary depending on the number of layers used and on individual judgement. In this study, all possible cases were taken to maintain consistency in interpretation of field data and in evaluating the weighted averages of the CPT and SPT values.

CHAPTER III

METHODOLOGY

3.1 General

This chapter deals with a brief overview of the various statistical principles, procedures and methodology employed in regression analysis. Both linear and nonlinear multiple regression models are described with respect to the estimation of parameters, statistical tests on the model parameters, and on the model adequacy. The parameter estimation for both models is based on the method of least squares (Montgomery et al., 1992) which is considered to produce the best possible estimate of the parameters. The t- and/or F-statistics are used to test the validity of the individual parameters as well as overall adequacy of the model. The R^2, C_p and MS_E statistics are used to determine the optimum number of variables in the model (Daniel et al., 1980). A brief description of these statistics is presented in this chapter. Various SAS procedures were used to perform the aforementioned statistical analyses. A brief description of these procedures, namely REG, RSQUARE, STEPWISE, and MODEL, is presented in Section 3.3.

3.2 Overview of Multiple Regression Models

Multiple regression analysis involves exploring relationships between one variable y, referred to as a response, dependent or exogenous variable, and p variables $x_1, x_2, \ldots\ldots, x_p (p > 1)$. The variables $x_1, x_2, \ldots\ldots\ldots, x_p$ are called independent, endogenous, or predictor variables, or most often simply the regressors. In the present context, total observed settlement of an approach pavement is the response (dependent variable), while the data pertaining to the various causative factors, such as age of the approach, height of the embankment, etc. (refer to Chapter II) are used as the regressors. The main aim of the regression analysis is to express the response variable y as a function of the predictor variables. The relationship between the response y and the predictor variables can be expressed in the following functional form:

$$y = f\{(x_1, x_2, \ldots\ldots, x_p), (\beta_1, \beta_2, \ldots\ldots, \beta_p)\} + \varepsilon \tag{3.1}$$

where $\beta_1, \beta_2, \ldots\ldots, \beta_p$ are unknown model parameters referred to as the regression coefficients and ε represents the error term which is inserted to account for the variability/fluctuation in y resulting from prediction by the regression function.

The multiple regression model can be classified as 'linear' or 'nonlinear' depending upon the way the unknown parameters appear in the model. The term "linear" refers to the linearity in the unknown parameters and not to the linearity in the predictor variables. Accordingly, the x_i's in Eqn. (3.1) can be squares, inverse, higher powers, cross-products and/or other transformations (e.g., logarithmic, exponential, etc.). For example, in the linear model developed in Chapter IV (see Sec. 4.3), the variables/regressors representing the embankment height, age of the approach, and traffic count are in their logarithmic forms, while those representing foundation soil thickness and SPT appear in the form of higher powers. A typical linear multiple regression can be expressed as

$$y = \beta_o + \beta_1 x_1 + \beta_2 x_2 + \ldots\ldots + \beta_p x_p + \varepsilon \tag{3.2}$$

In the same manner, a nonlinear multiple regression model is one in which the parameters appear nonlinearly. For example,

$$y = \beta_o + \beta_1 x_1^\theta + e^\alpha x_2 \tag{3.3}$$

where $\beta_o, \beta_1, \theta, \alpha$ are the unknown parameters and the model is nonlinear with respect to θ and α. Both linear and nonlinear multiple regression analyses have been carried out in the present study to obtain models which best represent the data set under consideration. The parameters (unknown) of the model are estimated by minimizing the sum of squared residuals, which is known as the method of least squares (LS). It is important to note that the statistical theory of multiple regression is based on some stringent classical assumptions; the accuracy/reliability of a model depends upon the extent to which these assumptions are incorporated into the model. Some of the important assumptions are:

- The response (dependent variable) is correct. However, the settlement data used in the study do suffer from some limitations (refer to Sec. 2.4.1).
- The form of the model is correct. The most appropriate functional forms for each of the regressors were determined before being included into the model, and hence this assumption is considered to be satisfied.
- Predictor variables are nonstochastic and are measured without error (see Sec. 2.4).
- The expected value of errors is zero.
- The errors are random quantities, independently distributed with zero mean, and constant variance σ^2 .

- The errors are uncorrelated across the observations.

Apart from these, an additional assumption that the errors are normally distributed, is made when hypotheses are tested. All these assumptions are critically reviewed in the context of the scope of the study.

3.2.1 Estimation of the Model Parameters

There are several methods available for estimating the values of the unknown parameters of the model. The Least Square (LS), Maximum Likelihood (ML) and Linear Approximation Method (LAM) are few examples which are widely used for this purpose (Seber et al., 1989). However, the computer algorithms used in the current statistical analysis are based on the method of LS only. Ratkowsky (1983) suggests that the method of least squares provides the best available estimates of the unknown parameters of the model, provided the assumptions outlined in Section 3.2 are satisfied. As the principle of LS method is same for linear as well as nonlinear model, the technique of LS is illustrated below in context with the linear multiple regression model represented by Eqn. (3.2)

Let there be $n>p$ observations, and let y_i denote the ith observed value of y (response) and x_{ij} denote the ith observation, then the model corresponding to Eqn. (3.2) may be written as

$$y_i = \beta_o + \beta_1 x_{i1} + \beta_2 x_{i2} + \ldots\ldots + \beta_p x_{ip} + \varepsilon_i \qquad (3.4a)$$

i.e.,

$$y_i = \beta_o + \sum_{j=i}^{p} \beta_j x_{ij} + \varepsilon_i \, , \qquad i=1,2,\ldots\ldots,n \qquad (3.4b)$$

The model in terms of the observations (3.4b) may be conveniently written in matrix form as

$$\{y\} = [x]\ \{\beta\} + \{\varepsilon\} \qquad (3.4c)$$

where

$$\{y\}=\begin{Bmatrix} y_1 \\ y_2 \\ \cdot \\ \cdot \\ \cdot \\ y_n \end{Bmatrix}, \quad [X]=\begin{bmatrix} 1 & x_{11} & x_{12} \ldots x_1 p \\ 1 & x_{21} & x_{22} \ldots x_{p2} \\ & \ldots\ldots\ldots\ldots & \\ 1 & x_{n1} & x_{n2} \ldots x_{np} \end{bmatrix}$$

$$\{\beta\}=\begin{Bmatrix} \beta_o \\ \beta_1 \\ \cdot \\ \cdot \\ \cdot \\ \beta_n \end{Bmatrix} \quad \text{and} \quad \{\varepsilon\}=\begin{Bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \cdot \\ \cdot \\ \cdot \\ \varepsilon_n \end{Bmatrix}$$

The least square function is

$$S(\beta_o, \beta_1, \ldots, \beta_p) = \Sigma\varepsilon^2 \qquad (3.5a)$$

In a more explicit form Eqn. (3.5a) can be written as,

$$SS_E = \Sigma(y_i - \beta_o - \Sigma\ \beta_j . x_{ij})^2 \qquad (3.5b)$$

where SS_E is called the sum of square errors. The unknown parameters $\overline{\beta_i' s}$ are obtained by minimizing the functions S with respect to $\beta_o, \beta_1, \ldots, \beta_p$. In the matrix form

$$SS_E = \{\varepsilon\}^T \{\varepsilon\} \tag{3.5c}$$

i.e.,

$$SS_E = (\{y\} - [x]^T \{\beta\})^T (\{y\} - [x]^T \{\beta\}) \tag{3.5d}$$

The vector of least squares estimators, $\{\overline{\beta}\}$, is obtained by solving the normal equations

$$[x][x]^T \{\overline{\beta}\} = [x] \{y\} \tag{3.6}$$

The solution is given by

$$\{\overline{\beta}\} = ([x][x]^T)^{-1} ([x]\{y\}) \tag{3.7}$$

3.2.2 Partitioning the Sum of Squares

The partitioning of the sum of the squares is basically done to derive many important statistics, such as, t- and F-statistics, and R^2-statistic. The method of least squares leads to the following identity (refer to Chatterjee et al.,1977):

$$\Sigma (y_i-\underline{y})^2 = \Sigma (\overline{y_i}-\underline{y})^2 + \Sigma (y_i - \overline{y_i})^2 \tag{3.9}$$

where $\underline{y}$ is the mean value of y_i's and $\overline{y}$ is the model predicted value. The above identity shows that the total sum of squares S_{yy} can be partitioned into a sum of squares due to regression (or model) and the residual (or error) sum of squares, i. e.,

$$S_{yy} = SS_R + SS_E \tag{3.10}$$

The SS_R is computed by the expression

$$SS_R = \{\beta\}\ [x]^T\ \{y\} \tag{3.11}$$

and SS_E by

$$SS_E = \{y\}^T\ \{y\}\ -\ (\{\beta\}^T[x]^T\{y\}) \tag{3.12}$$

The unbiased estimate of the error variance σ^2 (also called MSE) is obtained by dividing SS_E by its degree of freedom (d.o.f.), which is (n-p-1). In the same way, if SS_R is divided by its d.o.f. p, mean square due to regression (MSR) is obtained. Thus,

$$MSE = \frac{SS_E}{(n-p-1)} \tag{3.13}$$

$$MSR = \frac{SS_R}{p} \tag{3.14}$$

The quantities, namely MSE and MSR, are used to test the significance of individual parameters estimates (regression coefficients) and also to measure the adequacy of a model by calculating the t-statistic (t-value) and F-statistic (F-value). This is commonly known as 'Hypothesis Testing' in statistics. In the present study, t-values are used as a governing criterion to check the validity of the individual parameters; while F-values are mainly used to assess the overall adequacy of a model. In some cases, F-values have been used to measure the significance of an individual parameter.

The F-and t- statistics being the governing criteria in all the stages of model building, are described in detail below.

The t-value

The t-value, also called student's t, associated with the parameters estimate is a very useful criterion for examining the behavior of a model in estimation, and is defined as

$$t = \frac{\overline{\beta}_j}{\sqrt{(MSE)C_{jj}}} \tag{3.15}$$

where C_{jj} is the diagonal element of $([x]^T\,[x])^{-1}$ corresponding to b_{ij}. A high t-value associated with a parameter estimate is indicative of the fact that the estimate is well determined in the model; whereas a low t-value is a sign of poorly determined parameters. However, a low t-value may arise due to high correlation of the parameters with other parameters in the model (Ratkowsky, 1983). The value of t calculated by using Eqn. (3.15) is compared with the appropriate critical t-value (for n-p d.o.f.) from a standard t- distribution table (refer to any Standard text-book on Statistics, e.g., Montgomery et al., 1992), and the corresponding probability is evaluated at a specified level of significance. The t-value is used for testing the hypothesis that the individual parameter is zero. Thus, this is only a partial test as it does not incorporate the influence of other regression coefficients in a model.

The F-value

The overall significance of a model is tested by calculating the F-statistic for the model, which is defined as

$$F = \frac{MSR}{MSE} \tag{3.16}$$

The F-value is used for testing the hypothesis that all parameters in the model are zero. Usually the F-value obtained from Eqn. (3.16) is compared with the critical value of F(p, n-p-1) from a F-distribution table and the probability of the relationship being by chance is evaluated. However, in the present study, the appropriateness of the F-value (or t-value) is measured based on the significance probability (denoted by Prob > F, or Prob > |t|) associated with these values. In order to be a F-value significant at 5% level (e.g., 95% confidence level), the value of Prob > F is required to be 0.05 (also see Sec. 4.4.1). The F-value is also used to measure the relative significance of the regressors. Thus, a variable (causative factor) with high F-value indicates that the factor has significant influence on the approach settlement and vise-versa. Further details on t-and F- values and the related hypothesis testings may be obtained in any standard text book on regression analysis (Montgomery et al., 1992; Caulcult, 1983; Afifi et al., 1972; Chatterjee et al.,1977).

Apart from these two statistics, there are several other statistics which are employed to test the specific aspect of a model. Of these, Mallow's C_p, residual mean square error and the square of the coefficient of correlation, R^2, are widely used in the present study to determine the optimum sets of variables for a model. The R^2-statistic is also used as a measure of goodness of fit of the model. These statistics are described in detail in the following subsections.

3.2.3 Mallow's C_p-Statistic

The C_p-statistic is used as a criterion for selecting a model, and is defined as

$$C_p = \frac{(SS_E)_p}{s^2} + (2p-n) \qquad (3.17)$$

where s^2 is the MSE for the full model and $(SS_E)_p$ is the sum-of-squares errors for a model with p variables plus the intercept (Mallows, 1973). According to this criterion, a right model is the one which has C_p value equal to p (also see Sec. 4.2.3). In practice, C_p is plotted against p, and the model where C_p first approaches p is selected. The parameter estimates for such models are unbiased. A very thorough treatment of the C_p-Statistic may be obtained in Daniel et al. (1980).

3.2.4 Coefficient of Correlation, R^2

The coefficient of multiple correlation R, or most commonly the square of the multiple correlation coefficient R^2, is the most widely used measure of the adequacy of fit of a model, defined as

$$R^2 = \frac{SS_R(p)}{S_{yy}} = 1 - \frac{SS_E(p)}{S_{yy}} \qquad (3.18)$$

where $SS_R(p)$ and $SS_E(p)$ are the regression sum of squares, errors sum of squares, respectively, for a model with p variables. The values of R^2 can vary from 0 to 1. A low value of R^2 indicates a poor fit, whereas a value of R^2 close to 1.0 indicates that the model fits the data well. It is the measure of variability of response variable. Thus, for example, an R^2 value of 0.9656 indicates that 96.56 percent of the variability in the response variable is explained by the presence of the predictor variables in that model.

A modified version of R^2, called adjusted R^2-statistic, is commonly used in place of R^2. This is the value of R^2 after adjustment for degree of freedom has been done.

Thus,

$$(R^2)_{(adjusted)} = 1 - \frac{n-1}{n-p}\,(1-R^2) \qquad (3.19)$$

For a well specified model, R^2 and $R^2_{adjusted}$ are very similar.

3.2.5 Residual Mean Square, MS_E

The residual mean square for a subset regression model is given by

$$MS_E(p) = \frac{SS_E(p)}{n-p} \tag{3.20}$$

This is very useful in determining the optimum subset regression model. Usually MS_E (p) is plotted against the number of variables, p (see Fig. 4.2), and the optimum number of variables (optimum subset) is determined based on

- the minimum $MS_E(p)$ and
- a value of p near the point where the smallest $MS_E(p)$ turns upward (Montgomery et al., 1992).

3.3 Analysis Procedure Adopted

The statistical analysis and subsequent model building have been carried out by using various SAS regression analysis packages available at the University of Oklahoma, Norman mainframe computer system. SAS package software has many regression analysis options viz. REG, RSQUARE, STEPWISE, NLIN, RSREG, GLM, AUTOREG, SYSLIN, SYSNLIN, MODEL and PDLREG. In the present study, the REG, RSQUARE and STEPWISE procedures were used for linear regression modeling, while MODEL procedure was used for nonlinear modeling. In the subsequent subsections, the SAS procedures adopted in the current analysis are described in brief.

3.3.1. The REG-Procedure

The REG Procedure is a general-purpose procedure for regression, while other aforementioned procedures have more specialized applications. The REG procedure fits the least-squares estimates to linear regression models. REG uses the principle of Least Square to produce estimates that are best linear unbiased estimates (BLUE) under classical statistical assumptions mentioned in Section (3.2). In the present study, this procedure is extensively used for checking the validity of the parameter estimates, residual analysis, influence diagnostics, and generating various plots.

3.3.2 The RSQUARE-Procedure

The RSQUARE procedure selects optimal subsets of independent variables in a multiple regression analysis. This procedure is a useful tool for exploratory model building. The largest and the smallest number of independent variables in a subset and number of subsets of each size can be specified. The R^2 and/or $(R^2)_{adjusted}$ statistic is the criterion for selecting subsets. The RSQUARE procedure can also be effectively used to perform all possible subsets regressions in a decreasing order of model R^2 within each subset size.

3.3.3 The STEPWISE Procedure

The STEPWISE procedure is very useful in selecting significant variables for a model from a set of a large of number of independent variables. STEPWISE has five different methods for developing a regression model. These are

1) FORWARD - forward selection

2) BACKWARD - backward selection

3) STEPWISE - stepwise regression, backward and forward

4) MAXR - forward selection with pair switching

5) MINR - forward selection with pair searching

The forward-selection (FORWARD) technique begins with no variables in the model. For each of the independent variables, the procedure (FORWARD) calculates the F statistics reflecting the variables contribution to the model if it is included. FORWARD adds the variables that has the largest F statistic to the model. It then calculates the F statistic for the variables still remaining outside the model, and the evaluation process is repeated. Variables are added one by one to the model until no remaining variables produces a significant F statistic.

The backward elimination technique (BACKWARD) begins by evaluating the F statistics for a model, including all of the independent variables. Then the variables are deleted from the model one by one until all the variables remaining in the model produce F statistics significant at the specified level. At each step, the variables showing the smallest contribution to the model is deleted. The STEPWISE technique is a modification of the forward selection technique and differs in that the variables already present in the model do not necessarily stay there. After a variable is needed, stepwise method looks at all the variables already included in the model and deletes a variable that does not produce a significant F statistic.

The MAXR method begins by finding the one-variable model producing the highest R^2. Then another variable, the one that yields the greatest increase in R^2, is added. Once the two-variables model is obtained, each of the two variables in the model is compared to the variables not present in the model. For each comparison MAXR determines if removing one variable and replacing it with another increases R^2. Comparisons begin again, and the process continues until MAXR finds that no switch could increase R^2. Thus, the two-variables model achieved is considered the "best" two-variable model the technique can find. Another variable is then added to the model, and the comparing and switching process is repeated to find the best three-variables model, and so forth.

The minimum R^2 improvement technique (MINR) closely resembles MAXR, but the switch chosen is the one that produces the smallest increase in R^2. For a given number of variables in the model, MARX and MINR usually produce the same "best" model.

3.3.4 The MODEL Procedure

The MODEL Procedure is basically used for the nonlinear regression modeling. The MODEL procedure can be used to define, analyze the structure and estimate unknown parameters of nonlinear models. It has also been found useful in simulation and forecasting. It has the capability to deal with any complex types of nonlinear models. The parameter estimates are based on nonlinear least squares method. However, linear modeling can also be done using this procedure.

CHAPTER IV

LINEAR STATISTICAL MODELING

4.1 General

Prediction of bridge approach settlement can be of utmost importance in (i) identifying the problematic sites prior to the construction of bridge structures, (ii) selecting proper construction and/or maintenance techniques, and (iii) adopting suitable remedial measures. Deterministic approaches based on the principles of soil mechanics are generally used for these purposes. As a part of this overall project, a deterministic approach based on the Finite Element Method was used to evaluate the bridge approach settlement due to consolidation of the foundation soils resulting from embankment construction (Gopalasingam, 1989; Laguros et al., 1990; Mahmood, 1990; Vavarapis, 1991). Although deterministic models are desirable for a comprehensive evaluation of bridge approach settlements at specified sites, empirical models based on statistical methods have also proved to be very useful in assessing the settlement problems.

In this chapter, systematic statistical analyses are conducted on the settlement data from 26 selected bridge sites (refer to Chapter II) using various SAS Procedures. Although the level-two survey included a total of 29 sites, field tests at three sites remained incomplete, and as such these sites were not incorporated in statistical modeling. The various procedures used for multiple linear regression analyses are: RSQUARE, STEPWISE (e.g., FORWARD, BACKWARD and MAXR), and REG Procedure. Multiple linear regression analyses are performed to develop empirical relationships between bridge approach settlement and various causative factors. The efficacy of the best model is discussed with respect to its effectiveness as estimating approach settlements and identifying problematic bridge sites.

4.2 Variable Selection

There are a number of causative factors which contribute to the bridge approach settlement. It may not be possible to include all the factors in a model; however, it is necessary that at least the influential factors be included in the model either in a direct or an indirect manner. In this way, building a regression model that includes only a subset of the available regressors (causative factors) involves two conflicting objectives. First, the model should be such that it includes as many regressors as possible so that the "information content" in these factors can influence the predicted value of the approach settlement (TST). Secondly, the model should be such that it includes as few regressors as possible so that the variance of the prediction is minimal. In the following subsections, various variable selection strategies and the resulting subsets of models are presented which are, in essence, based on a compromise between the two aforementioned objectives.

4.2.1 Preliminary Investigation

In the present study, the following factors are considered to be influential in causing approach settlements. Hence they are used as candidate regressors or variables in developing a statistically-based model to predict bridge approach settlement.

Regressors/Variables	Description
SPT	SPT blow count (N-value)
TIPR	Tip resistance (tsf)
FR	Friction ratio (number)

FND	Foundation soil thickness (ft.)
AOE	Age of the approach (years)
EHT	Embankment height (ft.)
TCN	Traffic count (ADT)
SKEW	Skewness of the approach (degree)

Dependent Variable

TST	Total settlement/bridge approach settlement (inches)

As a preliminary investigation, the total settlement (TST) was plotted against each of the regressor separately to detect any specific relationship/pattern that may be established between approach settlement and a particular regressor. Best fit curves of various degree (linear, quadratic and cubic) were generated to represent the data set (Zaman et al., 1993). As an example, these plots are presented for the regressor SPT in Figs. 4.1a through 4.1c. These plots are very useful in detecting outliers and deciding the appropriate functional form of a regressor. The outliers are the unusual observations in a data set, the presence of which lead to poor correlation and bad prediction of the response. As indicated in Figs. 4.1a through 4.1c, the outliers are represented by a circle around the data point. It is observed that the data point representing the site with 18 inches of settlement might be an outlier responsible for poor correlation. Subsequently, the settlement data for that site was verified and it was found that the unusual value was correct. The outliers may result from many sources, such as erroneous observations in the field and data recording procedure, faulty data editing, etc. However, other tests such as residual analysis should be performed to fully ascertain these outliers. Also, no specific pattern is obvious from these plots to decide the functional form of a regressor. The

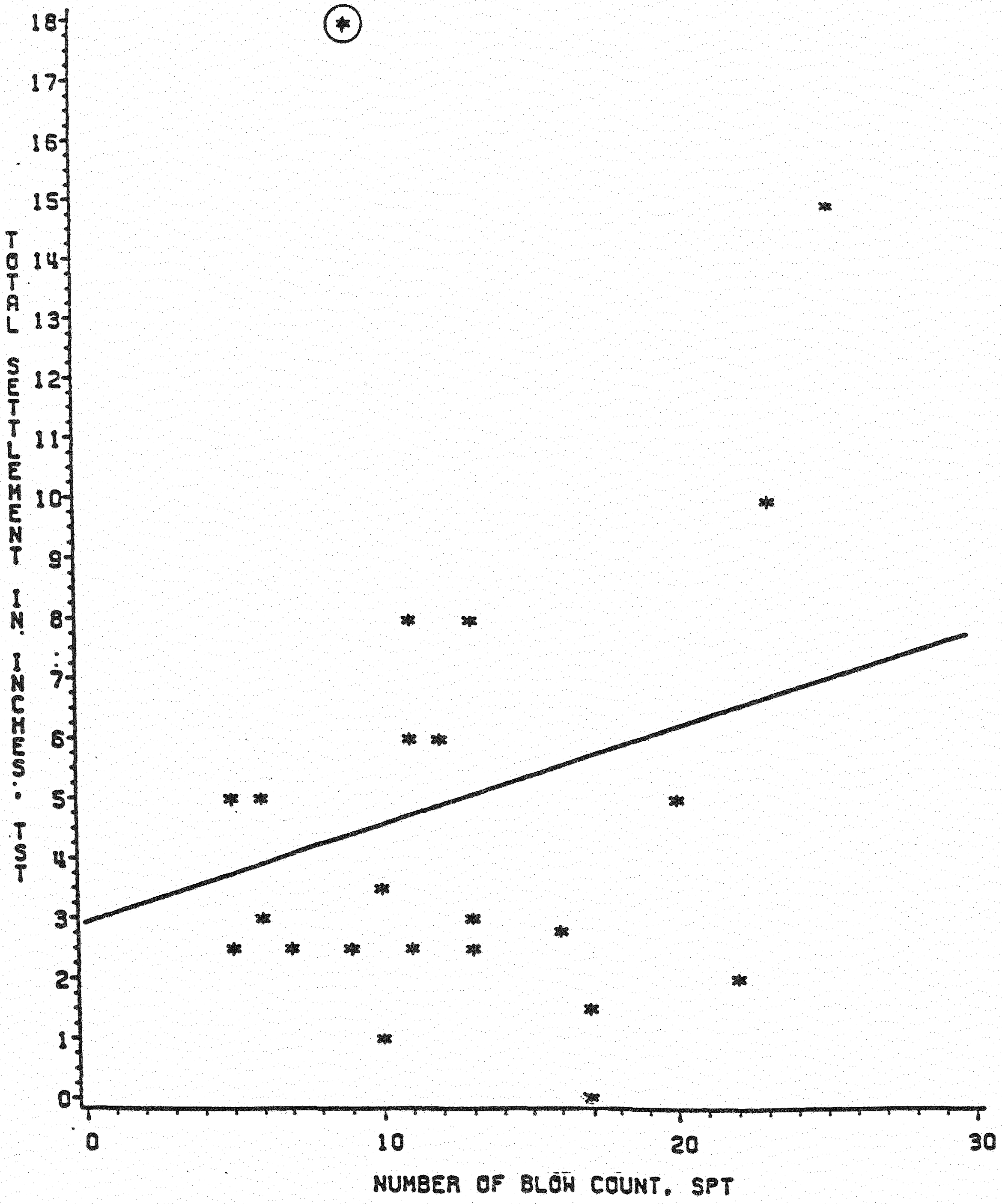


Fig. 4.1a Preliminary investigation and detection of outlier (best linear fit).

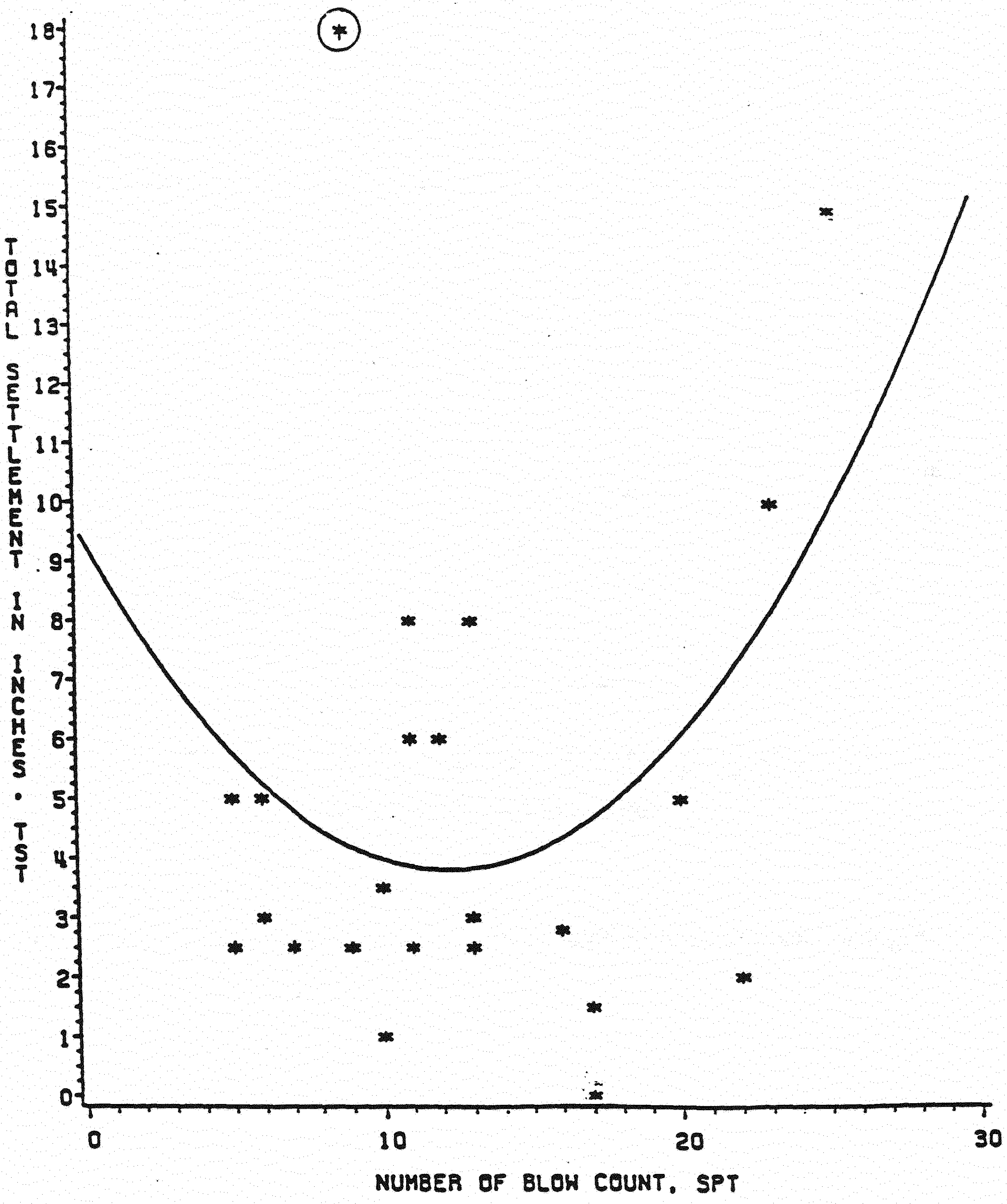


Fig. 4.1b Preliminary investigation and detection of outlier (best quadratic fit).

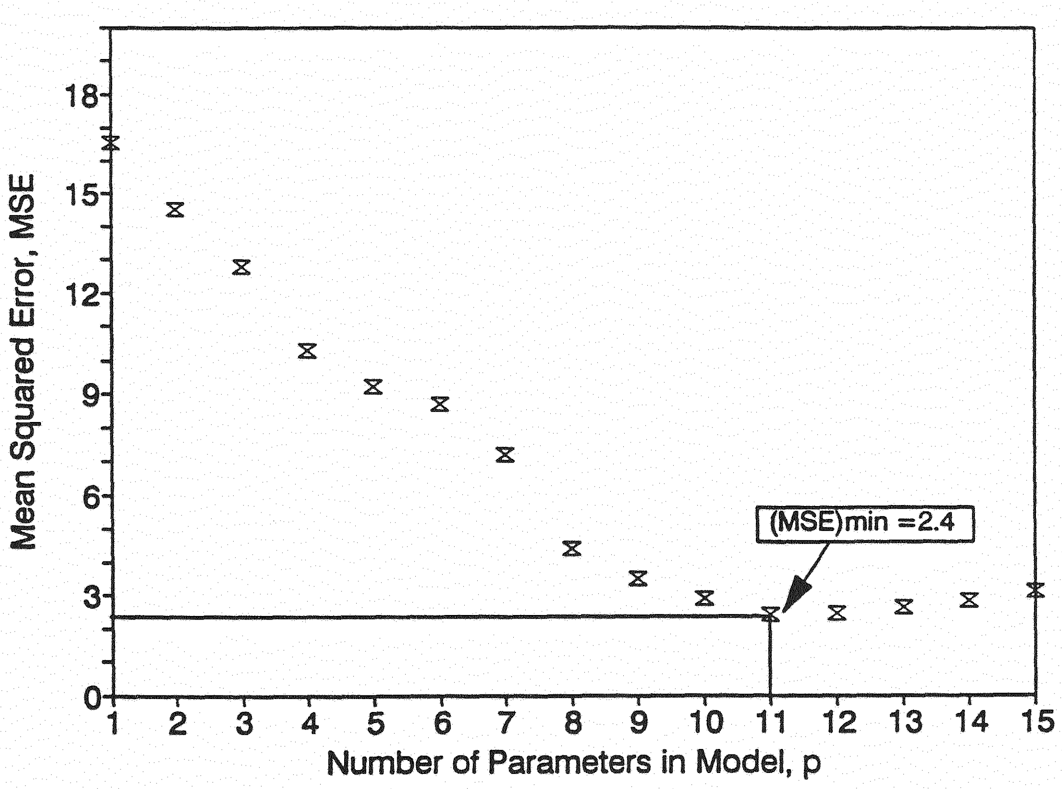


Fig 4.2 Determination of Number of Variables Based on Mean Square Error

need for correct functional form and its effect on the model are described in the next subsection.

4.2.2 Appropriate Functional Form

In many instances, it is found that a variable in a transformed form shows better measure of association with the dependent variable. In order to determine the correct functional form of a regressor, the regressor is transformed to different forms such as logarithmic, exponential, higher powers, inverse, etc. and the significance of each of these terms is examined. The F-statistics associated with the transformed terms are used as the "measure of significance". The terms with high F-statistic are retained in the model, while those with comparatively low F-statistic are deleted.

The transformation of a particular variable is as follows:

$$X2 = X^2 \tag{4.1a}$$

$$X3 = X^3 \tag{4.1b}$$

$$X4 = X^4 \tag{4.1c}$$

$$LX = \mathrm{Log}_{10} X \tag{4.1d}$$

$$INX = 1/X \tag{4.1e}$$

and so on.

The SAS Backward Elimination (BE) Procedure is used to generate the F-statistics for each regressor. In Table 4.1, it is observed that LTCN is the most significant term ($F_{LTCN} = 0.94$) as compared to other forms of the variable TCN ($F_{TCN}=0.4$, $F_{INLTCN}=0.15$). Hence in the subsequent model building process, only LTCN will be used.

Table 4.1 Determination of appropriate functional form for the regressor TCN.

Regressor/Variable	F-statistic
SPT2	1.58
SPT3	1.01
LAOE	2.53
LEHT	1.12
FND2	2.57
SPTFND	1.50
TCN	0.40
LTCN	0.94
INLTCN	0.15

Following the procedure outlined above, a pool of candidate regressors was determined which are listed below:

SPT
SPT3
LTIPR
FND
FND3
SPT2
SPT4
FR
FND2
FND4
SKEW
SPTFND
LAOE
LEHT
LTCN

The numerical values of the above mentioned variables are shown in Appendix I. The next step is to determine an optimum subset of variables for a model. There are several criteria for evaluating and comparing subset regression models (Cox et al., 1974; Myers, 1990). In the present study, the coefficient of correlation ($R^2_{adjusted}$), residual mean square (MS_E) and Mallows' C_p-statistic are used as criteria to determine the optimum subset regression models.

4.2.3 All Possible Regressions

The sample computer output obtained from the SAS RSQUARE Procedure (Zaman et al., 1993) shows the details of all possible subset regression models involving one-candidate regressor, two-candidate regressors, and so on. The optimum subsets models are presented in Table 4.2. It is observed that foundation soil thickness (FND) yields the best 1-variable model with $R^2_{adjusted}$ of 0.5884 (See Table 4.2). The best two variable ($R^2_{adjusted}$= 0.6391) is obtained by using the variables SPT4 and FND, and so on. It is observed that addition of more variables does not significantly increase the $R^2_{adjusted}$. Also, after a certain stage, inclusion of variables yields to decreased $R^2_{adjusted}$. In general, it is concluded that the 11-variable model ($R^2_{adjusted}$= 0.9402) including LTCN, FND4, SPT4, LTIPR, LEHT, LAOE, FND, SPT2, SPT3, FND2 and FND3 is the best subset of regression models. However, further tests are needed to determine the best model as discussed subsequently.

The optimum number of variables in a model could also be visualized from a plot of residual mean square error, $MS_E(p)$ (refer to Sec. 3.2.4) versus the number of variables, p (see Fig. 4.2). The minimum residual mean square model is the 11-regressor model (LTCN, FND4, SPT4, LTIPR, LEHT, LAOE, FND, SPT2, SPT3, FND2, FND3) with MS_E (11)= 2.40. It is also noted that the model which minimizes MS_E also maximizes the $R^2_{adjusted}$ (0.9402). Thus, from mean square error and $R^2_{adjusted}$ point of view, the 11-variable model can be considered as the optimum subset model.

Table 4.2 Determination of optimum subsets of model using SAS RSQUARE procedure.

No. of Variables	Variables in Model	$R^2_{adjusted}$
1	FND	0.5884
2	SPT4 FND	0.6391
3	SPT SPT4 SPT3	0.6818
4	SPT4 SPT2 SPT3 FND3	0.7436
5	SPT4 LTIPR SPT2 SPT3 FND3	0.7698
6	FND4 SPT4 SPT2 SPT3 FND2 FND3	0.7828
7	SPT LTCN FND4 SPT4 SPT2 FND2 FND3	0.8208
8	FND4 SPT4 LTIPR FND SPT2 SPT3 FND2 FND3	0.8904
9	LTCN FND4 SPT4 LTIPR FND SPT2 SPT3 FND2 FND3	0.9130
10	LTCN FND4 SPT4 LTIPR LAOE FND SPT2 SPT3 FND2 FND3	0.9283
11	LTCN FND4 SPT4 LTIPR LEHT LAOE FND SPT2 SPT3 FND2 FND3	0.9402
12	LTCN FND4 SPT4 LTIPR FR LEHT LAOE FND SPT2 SPT3 FND2 FND3	0.9391
13	SPT LTCN FND4 SPT4 LTIPR FR LEHT LAOE FND SPT2 SPT3 FND2 FND3	0.9348
14	SPT LTCN FND4 SPT4 LTIPR FR LEHT LAOE FND SPT2 SPT3 FND2 FND3 SPTFND	0.9294
15	SKEW SPT LTCN FND4 SPT4 LTIPR FR LEHT LAOE FND SPT2 SPT3 FND2 FND3 SPTFND	0.9230

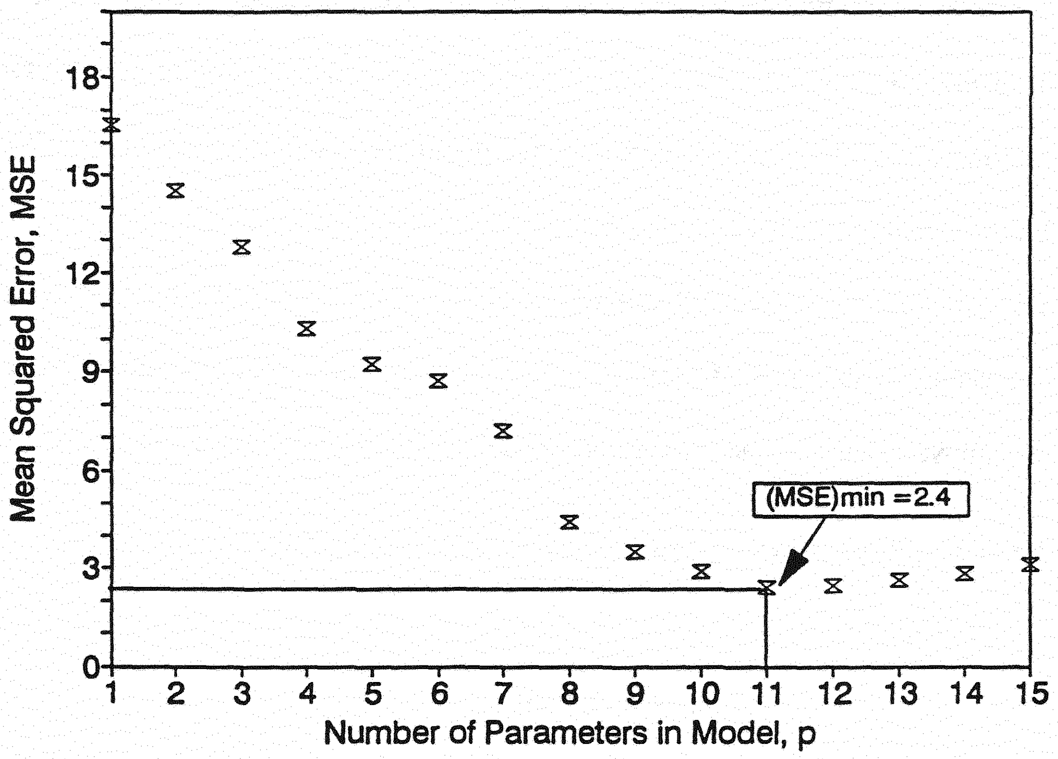


Fig 4.2 Determination of Number of Variables Based on Mean Square Error

A C_p plot (See Sec. 3.2.3) is shown in Fig. 4.3. On examining the plot, it is evident that five models, namely the 10-, 11-, 12-, 13- and 14-variable models

could be acceptable. For all of these models, C_p is very close to p (number of variables) and also falls below the C_p = p line. Also, as evidenced from Fig. 4.3, the C_p decreases in the beginning, reaches a minimum value, and then starts increasing again. Daniel et al. (1980) recommend using the "turning point" (e.g. the point after which C_p begins to increase) as a guide to determining the basic set of regressors for the models; therefore the 11-variable model may be chosen according to the C_p criterion.

Based on the aforementioned criteria, it is found that there can be several optimum models comprising various sets of variables. However, the best optimum subset model based on all of the above criteria is the 11-variable model which is presented below a generic form.

$$TST = F\,(LTCN,\ FND4,\ SPT4,\ LTIPR,\ LEHT,\ LAOE,\ FND,\ SPT2,\ SPT3,\ FND2,\ FND3) \tag{4.2}$$

(11-variable model)

It is observed that the variables SKEW and the cross-product term SPTFND do not significantly add to the accuracy of the above models, indicating that skewness of the approach does not have any influence on the approach settlement. A similar finding was reported by Mahmood (1990).

It is important to note that the evaluation of optimum subsets of models are based on certain assumptions, namely correct functional forms of regressors and presence of no outliers or influential observations in the data set (Montgomery et al., 1992). Therefore, before arriving at the final model, a model should be checked for residual analysis, collinearity, outliers

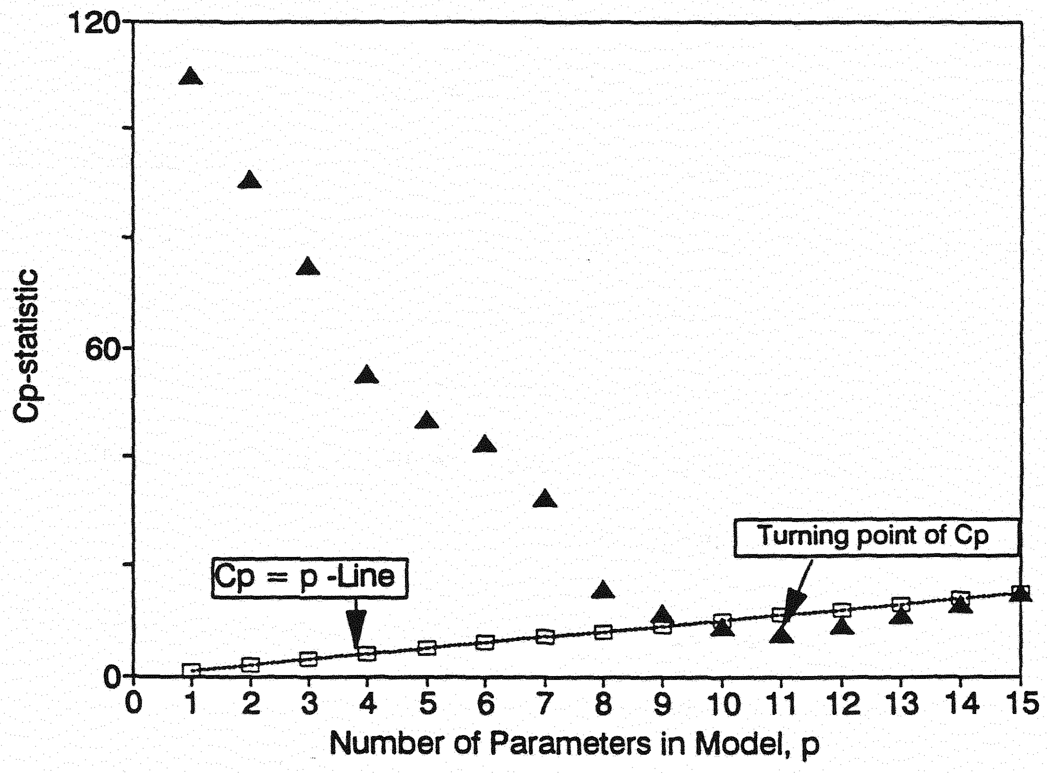


Fig. 4.3 Determination of Number of Variables Based on C_p-Statistic

It was observed that almost all of the aforementioned models failed to meet the convergence criteria (see Table 5.1) indicating that these models are not suitable for the data-set under consideration (Belsley et al., 1980). Further trials were made using relatively simpler models, i.e., close-to-linear models. It was observed that as the model approaches linear (linear with respect to parameters), the parameter estimates always converge in one iteration. This observation leads to the conclusion that a linear or close-to-linear model would be more viable option for the data-set under consideration. Moreover, Ratkowsky (1983) suggests that the model that comes closest to behaving as a linear model should be the preferred choice. While formulating such a nonlinear model/equation, it is highly desirable that the model comprises of as few variables as possible without sacrificing much accuracy of the model. Following these guidelines, the following nonlinear model was found to be the best model representing the data set under consideration (i.e., for 26 bridge sites):

$$TST = 0.016219\ (SPTE2) - 0.096970\ (TIPR) + 2.492993\ (LTIPR) - 0.751326\ (FR) + 0.571589\ (LFR) \tag{5.1}$$

$$(\mathbf{R^2_{adjusted} = 0.3309})$$

where SPTE is the SPT value for the embankment only (refer to Sec. 2.4.2), while the remaining variables have same meaning as defined in Section 4.4.

The square of the coefficient of correlation, $R^2_{adjusted}$ for this model is 0.3309, which indicates that the above model does not fit the data set well. In other words, only about 33% of the variability in total settlement is accounted for by this model. In order to investigate the causes of such a poor correlation, residual plot (refer to Sec. 4.4.2) of the model was generated (Fig. 5.1). The figure indicates that the predicted settlements of the sites having settlements of

Table 4.3a Summary of the SAS forward selection (FS) procedure at significance level $\alpha = 0.25$.

Step	Variable Entered	Partial R^2	Model R^2	F	Prob > F
1	FND	0.6042	0.6042	38.16	0.0001
2	SPT4	0.0627	0.6669	4.51	0.0441

Table 4.3b Parameter estimates, F-ratios and their significance probability (prob > f) produced by the SAS FS procedure (Eqn. 4.4).

Variable	Parameter	Parameter Estimates	F	Prob > F
SPT4	b1	0.00001618	4.51	0.0441
FND	b2	0.10990724	21.28	0.0001

The backward elimination (BE) method was then carried out at $\alpha = 0.10$ so that any variable having a confidence level of 90% would be retained in the model. The results of BE method are presented in Tables 4.4a and 4.4b. This method yields the following 11-variable model at $\alpha = 0.10$:

$$\begin{aligned} TST = &- 4.2056\ (LTCN) - .00002467\ (FND4) + 0.000991\ (SPT4) \\ &- 8.6133\ (LTIPR) + 3.53568\ (LEHT) + 3.0401\ (LAOE) \\ &+ 2.7808\ (FND) + 0.425376\ (SPT2) - 0.0400405\ (SPT3) \\ &- 0.1685713\ (FND2) + 0.00362526\ (FND3) \end{aligned} \tag{4.5}$$

Table 4.4a Summary of the SAS backward elimination (BE) procedure at significance level $\alpha = 0.10$.

Step	Variable Removed	Partial R^2	Model R^2	F	Prob > F
1	SKEW	0.0000	0.9674	0.0001	0.9917
2	SPTFND	0.0000	0.9674	0.0070	0.9349
3	SPT	0.0002	0.9672	0.0802	0.7815
4	FR	0.0017	0.9655	0.7299	0.4073

Table 4.4b Parameter estimates, F-ratios and their significance probability (Prob > F) produced by the SAS BE procedure (Eqn. 4.5).

Variable	Parameter	Parameter Estimates	F	Prob > F
LTCN	b1	-4.205655	16.17	0.0011
FND4	b2	-0.00002467	55.77	0.0001
SPT4	b3	0.00099054	55.88	0.0001
LTIPR	b4	-8.613293	18.43	0.0006
LEHT	b5	3.53568449	4.18	0.0587
LAOE	b6	3.04015704	6.40	0.0231
FND	b7	2.78076844	39.20	0.0001
SPT2	b8	0.42537576	43.56	0.0001
SPT3	b9	-0.04004052	50.10	0.0001
FND2	b10	-0.16857125	48.41	0.0001
FND3	b11	0.00362526	54.09	0.0001

It is important to note that the FS method, in spite of using a more conservative significance level ($\alpha = 0.25$) ended with a 2-variable model, while the BE method yielded 11-variable model at a higher confidence level ($\alpha = 0.10$). This indicates that there exist intercorrelations between the regressors. In other words, an individual causative factor is possibly linked with or related to other factor(s) in contributing to approach settlement problems. The BE method is often less adversely affected by the correlative structure of the regressors than the FS method (Mantel, 1970). Hence, the model obtained from the BE method (Eqn. 4.5) is considered as the possible candidate for the final model. Moreover, the best models with optimum subsets of variables, based on the C_p- statistic, minimum MS_E, and maximum $R^2_{adjusted}$ (Sec. 4.2.3) contain the same sets of variables which also appear in the model specified by Eqn. 4.5.

The BE method starts with Step 0 (Zaman et al., 1993) which shows the results of fitting the full model (e.g. all variables entered). The smallest partial F value is zero (0.0), and it is associated with SKEW and SPTFND (Tables 4.4a and 4.4b). Out of these two variables, the significance probability of the F value for SKEW is more (Prob > F = 0.9917) than that of SPTFND (Prob > 0.9349); hence, the variable SKEW is removed from the model at Step 1 and the parameter estimates and the F values are calculated again. The SPTFND is removed from the model in Step 2, and so on, until all the variables retained in the model are significant at 0.1 level.

The SAS MAXR Procedure (maximum R^2 improvement method) is then used (See Sec. 3.3.3) to check the statistical significance of a model. According to this method, the best model will have the highest possible R^2. Parameter estimates are also produced by this procedure. Table 4.5 presents the parameter estimates and other related statistics for the best 11-variable model produced by the SAS version of MAXR (for detailed computer output refer to Zaman et al., 1993).

Table 4.5 Parameter estimates, F-ratios and their significance probability (Prob > F) for the best 11-variable model produced by the SAS MAXR procedure (Eqn. 4.6).

Variable	Parameter	Parameter Estimates	F	Prob > F
LTCN	b1	-4.205655	16.17	0.0011
FND4	b2	-0.00002467	55.77	0.0001
SPT4	b3	0.00099054	55.88	0.0001
LTIPR	b4	-8.613293	18.43	0.0006
LEHT	b5	3.53568449	4.18	0.0587
LAOE	b6	3.04015704	6.40	0.0231
FND	b7	2.78076844	39.20	0.0001
SPT2	b8	0.42537576	43.56	0.0001
SPT3	b9	-0.04004052	50.10	0.0001
FND2	b10	-0.16857125	48.41	0.0001
FND3	b11	0.00362526	54.09	0.0001

In the context of the present study, this procedure is used for the purpose of comparing the models obtained from BE and FS methods. It is found that the best 2-variable model obtained by the FS method (Eqn. 4.4) as well as the 11-variable model obtained from the BE method (Eqn. 4.5) are similar to the ones obtained by this method (e.g., MAXR method). The best 11-variable model (R^2 =0.9655) resulting from the MAXR Procedure is as follows:

$$\begin{aligned} TST = & - 4.2056\ (LTCN) - .00002467\ (FND4) + 0.000991\ (SPT4) \\ & - 8.6133\ (LTIPR) + 3.53568\ (LEHT) + 3.0401\ (LAOE) \\ & + 2.7808\ (FND) + 0.425376\ (SPT2) - 0.0400405\ (SPT3) \\ & - 0.1685713\ (FND2) + 0.00362526\ (FND3) \end{aligned} \tag{4.6}$$

4.4 Critical Assessment of the Adequacy of the Model

The final model represented by Eqn. (4.6) is further assessed to establish the adequacy of the model. The validity of the individual parameter estimates are judged based on the |t-statistic| associated with each parameter (Sec. 3.2.2), while residual plots are used to assess the goodness of fit of the model. Finally, the influential observation(s) having conspicuous effect on the model are investigated based on certain influence measuring statistics such as Cook's D, DFFITS and DFBETAS statistics. These statistics are discussed in detail in Sec. 4.4.3.

4.4.1 Validity of Parameters

The t-statistics associated with the parameters estimates are obtained using the SAS REG Procedure. Table 4.6 presents an extract from the computer output generated by the REG Procedure (Zaman et al., 1993) comprising of t-statistics corresponding to individual parameters concomitant with other useful results. Usually, the t-value of each of the parameter is compared with the critical value of t obtained from t-distribution Table (refer to any standard text on statistics,e.g., Montgomery et al., 1992). For example, for total number of observations, n (i.e., 26) and number of variables, p (i.e., 11), the t-distribution Table gives $t_{critical} = 1.34$ at 0.1 level of significance (Montgomery et al., 1992). It is observed that the |t| value for each of the estimates is greater than the critical value of t. However, in the present study, the significance probability of t (e.g., Prob > |t|) is used to assess the appropriateness of the t-values. A |t| value with Prob > |t| equal to or less than 0.05 is generally preferred. This means the probability of getting such t value is 95%. As can be seen from Table 4.6, the significance probability associated with each t value is within or near the range of 0.05 (95% confidence level).

Table 4.6 Determination of validity of the parameter estimates. Parameter estimates, t-ratios and their significance probability (Prob > t) produced by the SAS REG procedure (Eqn. 4.5).

Variable	Parameter	Parameter Estimates	t-value	Prob > |t|
LTCN	b1	-4.205655	-4.021	0.0011
FND4	b2	-0.00002467	-7.468	0.0001
SPT4	b3	0.00099054	7.475	0.0001
LTIPR	b4	-8.613293	-4.293	0.0006
LEHT	b5	3.53568449	2.046	0.0587
LAOE	b6	3.04015704	2.529	0.0231
FND	b7	2.78076844	6.261	0.0001
SPT2	b8	0.42537576	6.600	0.0001
SPT3	b9	-0.04004052	-7.078	0.0001
FND2	b10	-0.16857125	-6.958	0.0001
FND3	b11	0.00362526	7.355	0.0001

4.4.2 Residual Plot and Prediction by the Model

The residuals from the multiple regression model play an important role in judging model adequacy. Generally, the residuals are plotted against predicted value of response and individual regressors. Figure 4.4 presents the residual versus predicted value of the bridge approach settlement (TST). No strong unusual pattern is evident from the plot. It is observed that the settlements of 14 sites are slightly overpredicted, while those of the other 10 sites are underpredicted by the model; however, the overall fluctuation for any site does not exceed 2 inches. Figure 4.5 shows a similar plot in which actual and the predicted settlements are plotted against site numbers.

4.4.3 Identification of Problematic Sites

Previous studies suggest that a problematic bridge site is the one which has experienced a settlement of 2 inches or more (Zaman et al., 1990). In the current study, however, any site with a settlement of 1 inch or more is considered as problematic. Following this criterion, it is observed that the model successfully identifies the majority of the problematic sites (about 90%) under consideration (Fig. 4.5). However, in the case of two sites, namely Site #20 and Site #25 which are problematic, the model predicts them as nonproblematic. The nonproblematic sites are correctly predicted by the model. Thus, the model is found to be relatively effective in identifying both problematic and nonproblematic bridge sites.

4.4.4 Influence Diagnostics

Occasionally it is found that a small subset of the data exerts a disproportionate influence on the fitted regression model. In other words, parameter estimates and/or prediction may depend more on the influential subset than on majority of the data. In order to locate these influential points and assess their impact on the model, the REG Procedure is used to generate influence measuring statistics such as Cook's D, DFFITS and DFBETAS statistics (see Table 4.7).

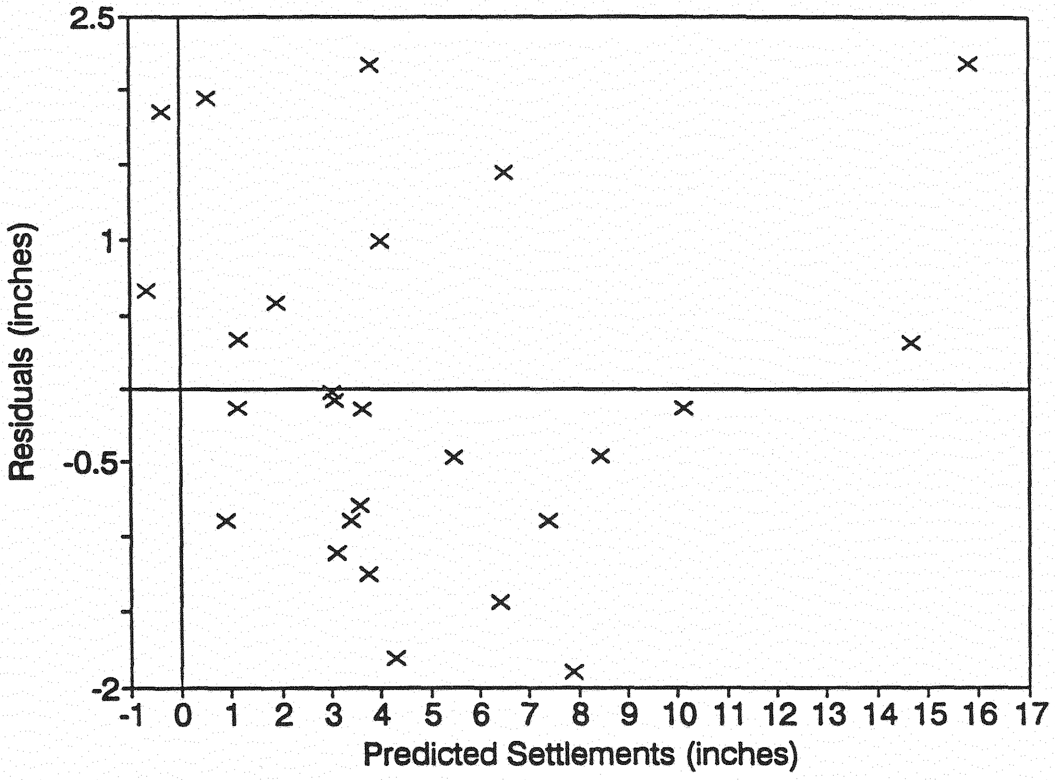


Fig. 4.4 Assessment of the adequacy of the model based on residual analysis.

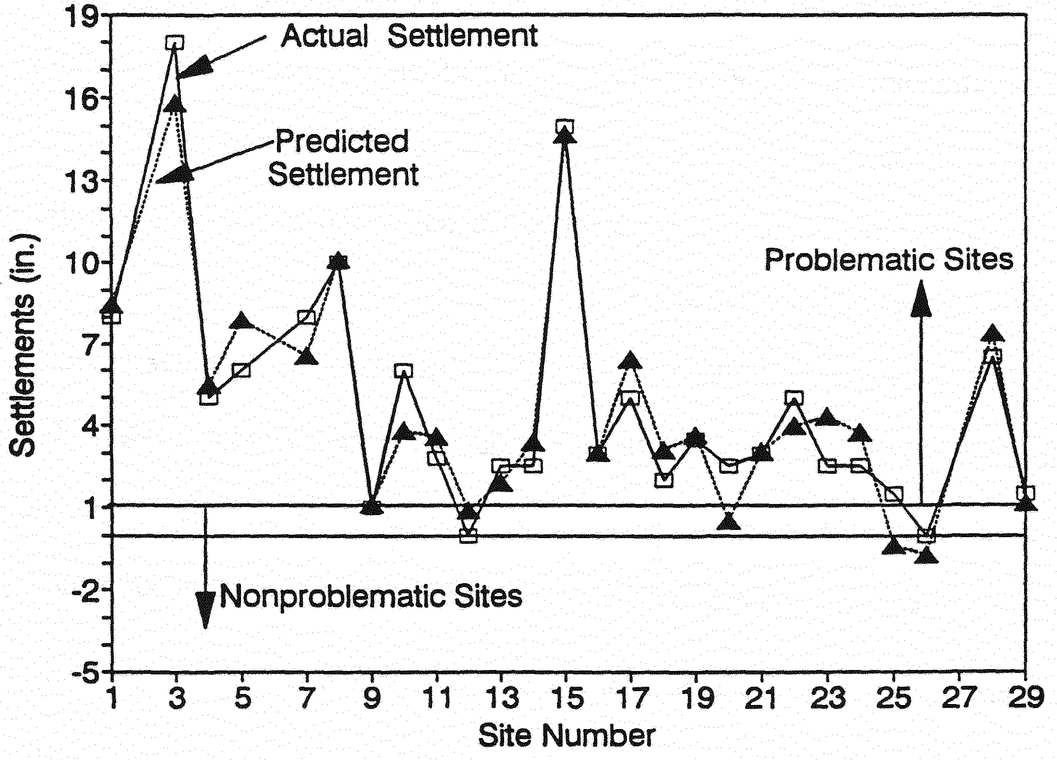


Fig. 4.5 Prediction obtained by the linear model.

Cook's D statistic is a measure of influence representing the change to the estimates that results from deleting each observation (Cook; 1977, 1979). The points/observations for which D > 1 are considered influential (Montgomery et al., 1992). It is observed that Observation 3 (i.e. Site #4) has exceptionally high value of Cook's D (=15.139), while all other observations have this value well below the cutoff value (see Fig. 4.6). This indicates that observation 3 has remarkable influence on the model. Back examining the original data set for Site #4 (refer to Table 2.3), it is found that it has very large foundation soil thickness (FND = 72 ft.). Thus, it may be inferred that foundation soil thickness is one of the most dominant causative factors responsible for approach settlement.

DFFITS-statistic is similar to Cook's D statistic and is used to investigate the influence of each of the observation on the predicted or fitted value (Belsley et al., 1980). For the model under consideration, these statistics have been obtained using the REG Procedure and are shown in Table 4.7 under the column "Dffits". In general, any observation for which $|\text{DFFITS}| > 2\sqrt{p/n}$, should be considered influential observation (Montgomery et al., 1992). As reflected by Cook's D-statistic, this statistic also indicates that the model is highly influenced by the Observation 3 ($|\text{DFFITS}| = 14.5228$). From Table 4.7, it is also observed that some of the other sites have $|\text{DFFITS}| > 1.3831$ (i.e. $> 2\sqrt{p/n}$) as well, indicating that they might be influential. However, in comparison with the highest DFFITS associated with Observation 3, these values are too small; and hence do not warrant attention.

Table 4.7 Detection of influential observation using the SAS REG procedure.

Site #	Cook's D-statistic	|Dffits-statistic|
1	0.004	0.2049
3	0.813	3.5936
4	18.679+	16.4941+
5	0.209	1.6228
7	0.077	0.9367
8	0.002	0.1330
9	0.000	0.0518
10	0.074	0.9571
11	0.029	0.5506
12	0.087	0.9700
13	0.084	0.9404
14	0.036	0.6157
15	0.191	1.4167
16	0.000	0.0158
17	0.059	0.8135
18	0.041	0.6684
19	0.000	0.0676
20	0.114	1.1823
21	0.000	0.0280
22	0.072	0.8858
23	0.066	0.8802
24	0.040	0.6661
25	0.073	0.9305
26	0.038	0.6313
28	0.017	0.4191
29	0.001	0.1133

+*Exceptionally high values indicate influential observation.*

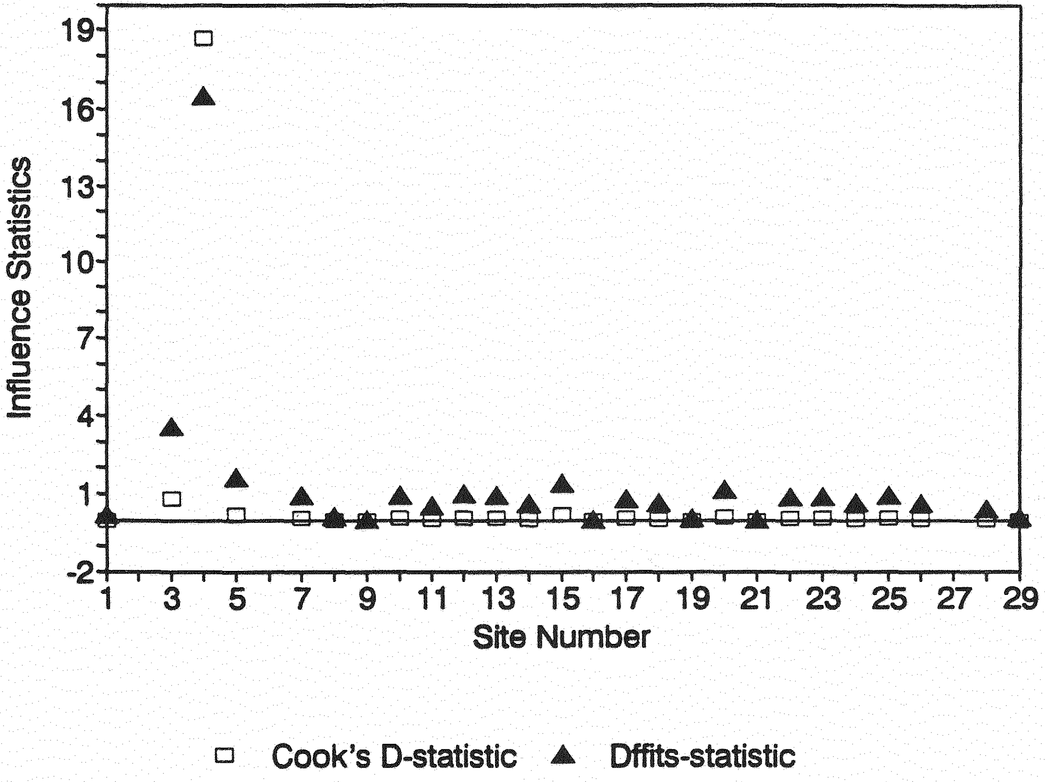


Fig. 4.6 Detection of influential observation (influence diagnostics).

The DFBETAS statistic basically measures the influence based on the changes in parameter estimates after deleting each observation (Belsley et al., 1980). Since prediction capability of the model is the prime focus of the present study, this statistic is considered to be of secondary importance and is not used in the present context.

4.5 Concluding Remarks

Based on the extensive statistical analyses discussed in the preceding sections, the following linear multiple regression is proposed

$$\begin{aligned} TST = & - 4.2056\ (LTCN) - .00002467\ (FND4) + 0.000991\ (SPT4) \\ & - 8.6133\ (LTIPR) + 3.53568\ (LEHT) + 3.0401\ (LAOE) \\ & + 2.7808\ (FND) + 0.425376\ (SPT2) - 0.0400405\ (SPT3) \\ & - 0.1685713\ (FND2) + 0.00362526\ (FND3) \end{aligned} \tag{4.7}$$

$$(R^2_{adjusted} = 0.9402)$$

In the equation above, the variables representing significant causative factors affecting bridge approach settlement (TST) are present in their most appropriate functional forms. The optimum number of variables (e.g., 11) in the model has been selected based on $R^2_{adjusted}$, residual mean square error (MS_E), and C_p-statistic, so that the resulting model has maximum $R^2_{adjusted}$ and minimum residual error in the predicted settlement. Also, all the variables in the model are significant at 90% confidence level. The parameter estimates are well determined as reflected by the high $|$t-ratio$|$ associated with each of the parameters. The predictive ability of the model is accurate to a reasonable level. The errors associated with the predicted settlements seem to be normally distributed. The model also seems to be very effective in identifying problematic and nonproblematic sites. The influence diagnostics conducted on the data set confirm that Site #4, with exceptionally high value of foundation soil thickness (FND), has proportionately more influence on the model as compared to other observations.

CHAPTER V

NONLINEAR REGRESSION MODELS

5.1 General

Usually nonlinear models are opted for when linear models fail to produce the desired results. This is, however, not the supporting reason in the present context. In the course of the present study, it was decided to explore possible nonlinear relationships between the approach settlement and the individual factors potentially responsible for such settlement so that a nonlinear regression model, comprised of a relatively fewer number of variables, could be developed to predict the bridge approach settlement. The significant variables used in these models have been chosen based on the results of the linear regression analyses (refer to Chapter IV). The SAS MODEL Procedure is extensively used to generate parameter estimates and other desired statistics to assess the validity of the model.

5.2 Nonlinear Model

Nonlinear regression analysis is primarily intended to explore the existence of a specific nonlinearity, if any, between the approach settlement and the individual causative factors. The absence of prior knowledge of any such relationship led to attempts of several standard nonlinear models to fit the data-set under consideration. Table 5.1 presents the standard nonlinear models used in the present study. For detailed descriptions of these models and the general mechanisms involved in nonlinear modeling, refer to any standard text on nonlinear regression models, e.g., Belsley et al.(1980), Ratkowsky (1983). The efficacy of the model is judged by the $R^2_{adjusted}$-statistic and the prediction obtained by the model.

TABLE 5.1 Standard nonlinear models.

Trial #	**Nonlinear Equation Used**	**Remarks**
1	Michaelis-Menten Model: $f(x, \theta) = \frac{\theta_1 x}{\theta_2 + x}$	No Convergence
2	Asymptotic Model: $f(x, \theta) = \theta_1 + \theta_2 e^{(\theta_3 x)}$	No Convergence
3	Logistic Model: $f(x, \theta) = \frac{\theta_1}{1 + \theta_2 e^{(\theta_3 x)}}$	$R^2 = 0.0$ θ_2 and θ_3 - Biased
4	Gompertz Growth Model: $f(x, \theta) = \theta_1 e^{-e^{(\theta_2 - \theta_3 x)}}$	No Convergence
5	Log Logistic Growth Model: $f(x, \theta) = \theta_1 - \ln\left[1 + \theta_2 e^{-\theta_3 x}\right]$	$R^2 = 0.0$ Parameters Biased
6	Morgan-Mercer-Flodin Growth Model: $f(x, \theta) = \frac{\theta_2 \theta_3 + \theta_1 x^{\theta_4}}{\theta_3 + x^{\theta_4}}$	No Convergence
7	Richards Growth Model: $f(x, \theta) = \frac{\theta_1}{\left[1 + \theta_2 e^{(-\theta_3 x)}\right]^{1/\theta_4}}$	No Convergence

It was observed that almost all of the aforementioned models failed to meet the convergence criteria (see Table 5.1) indicating that these models are not suitable for the data-set under consideration (Belsley et al., 1980). Further trials were made using relatively simpler models, i.e., close-to-linear models. It was observed that as the model approaches linear (linear with respect to parameters), the parameter estimates always converge in one iteration. This observation leads to the conclusion that a linear or close-to-linear model would be more viable option for the data-set under consideration. Moreover, Ratkowsky (1983) suggests that the model that comes closest to behaving as a linear model should be the preferred choice. While formulating such a nonlinear model/equation, it is highly desirable that the model comprises of as few variables as possible without sacrificing much accuracy of the model. Following these guidelines, the following nonlinear model was found to be the best model representing the data set under consideration (i.e., for 26 bridge sites):

$$TST = 0.016219\,(SPTE2) - 0.096970\,(TIPR) + 2.492993\,(LTIPR) - 0.751326\,(FR) + 0.571589\,(LFR) \tag{5.1}$$

$$(\mathbf{R^2_{adjusted} = 0.3309})$$

where SPTE is the SPT value for the embankment only (refer to Sec. 2.4.2), while the remaining variables have same meaning as defined in Section 4.4.

The square of the coefficient of correlation, $R^2_{adjusted}$ for this model is 0.3309, which indicates that the above model does not fit the data set well. In other words, only about 33% of the variability in total settlement is accounted for by this model. In order to investigate the causes of such a poor correlation, residual plot (refer to Sec. 4.4.2) of the model was generated (Fig. 5.1). The figure indicates that the predicted settlements of the sites having settlements of

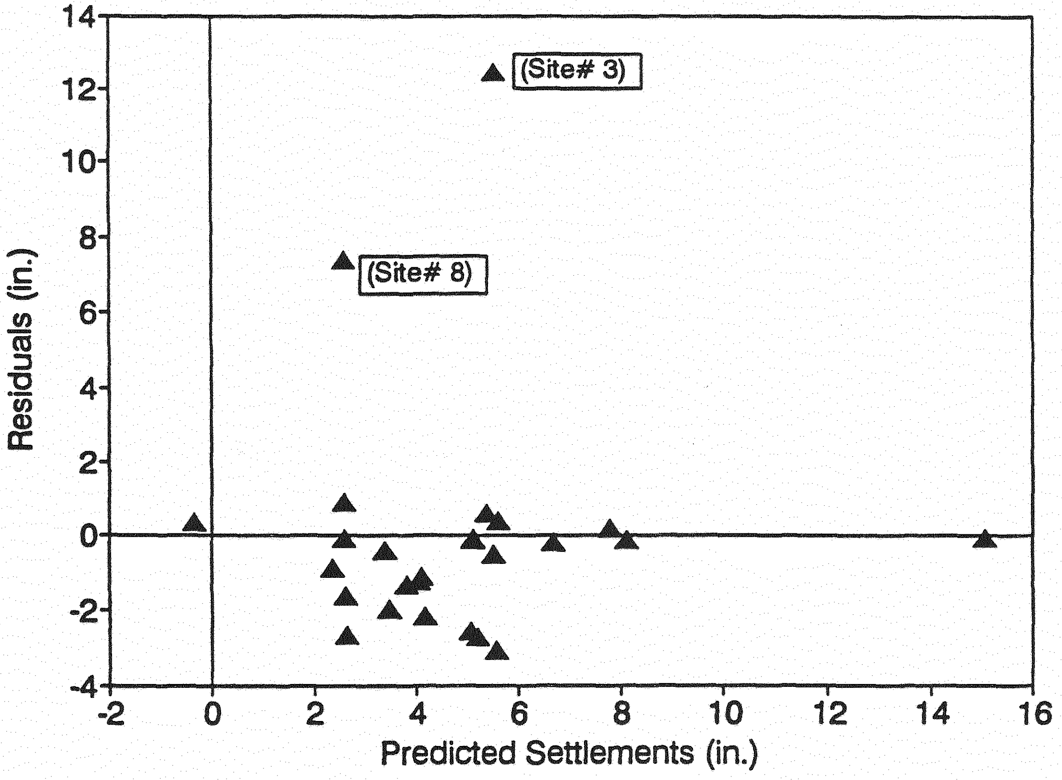


Fig. 5.1 Residual analysis for determination of problematic sites.

18 in. and 10 in. (e.g., Site #3 and 8) are associated with very large amount of errors. Therefore in the subsequent model building process, these two sites were removed from the data set. As a result of deleting these two sites, it was observed that the $R^2_{adjusted}$ increased from 0.3309 to 0.8816. Such a drastic improvement in correlation definitely provides rationale for removing these two sites from the analysis. Moreover, the t-values associated with the parameter estimates also improved significantly (Table 5.2 and 5.3). It is seen that the $|t|$ values associated with each of the parameters is fairly high, and also the significance probability of these t-values fall within and/or near the acceptable range of .05 (i.e. 95% confidence level). It was observed that inclusion of additional parameter(s) and/or other functional forms of the variables results in a poor estimation of the parameters (i.e., low $|t|$ value) and also deteriorated prediction. The new nonlinear model based on the data from 24 bridge sites is presented below:

$$\begin{aligned} TST = \; & 0.016545\,(SPTE2) - 0.08886\,(TIPR) + 2.413741\,(LTIPR) - 1.909181\,(FR) \\ & + 3.132881\,(LFR) \qquad (5.2) \end{aligned}$$

$$(\mathbf{R^2_{adjusted} = 0.8816})$$

The predictions by this model are fairly accurate as shown in Fig. 5.2. As can be observed from this graph, the model is able to predict the extreme values of the settlements (e.g., 15 inches and 0 inches); also all other intermediate values are well approximated by the model. The problematic and the nonproblematic sites are also well identified by the model (Fig. 5.2) quite satisfactorily. Detailed computer output produced by the SAS MODEL Procedure are given in a separate report (Zaman et al., 1993).

CHAPTER VI

CONCLUSIONS AND RECOMMENDATIONS

6.1 Summary

This report presents the results of statistical analyses conducted on the data set of 26 selected bridge approach sites in Oklahoma (See Table 2.1). The majority of the data were obtained from the field tests (e.g. CPT and SPT tests) which were conducted as a part of level-two survey of these sites. However, some of the data such as age and skewness of the approach, traffic count (ADT) and height of the embankment were obtained from the level-one survey and/or Bridge Division, ODOT. The selected sites encompass a wide range of variation with respect to the total settlement and the various causative factors/parameters such as age of the approach, traffic count, skewness of the approach, height of the embankment, thickness of the foundation soil underlying the embankment, and characteristics of the embankment and the foundation soil. The site-specific SPT and CPT results are used for the quantitative representation of the embankment and foundation soil characteristics. Extensive statistical analyses were performed using a statistical package program, called SAS. The following SAS Procedures were opted: REG, RSQUARE, STEPWISE and MODEL. Both linear and nonlinear multiple regression models were developed to predict the approach settlements. The multiple regression analyses were restricted to the use of quantitative variables only. The validity of the variables/parameters in the models is judged based on their level of significance, partial R^2 (square of coefficient of correlation), t- and F- statistics; while the best models are decided on the basis of their overall $R^2_{(adjusted)}$, Mallow's C_p-statistic, mean square error (MS_E), and their

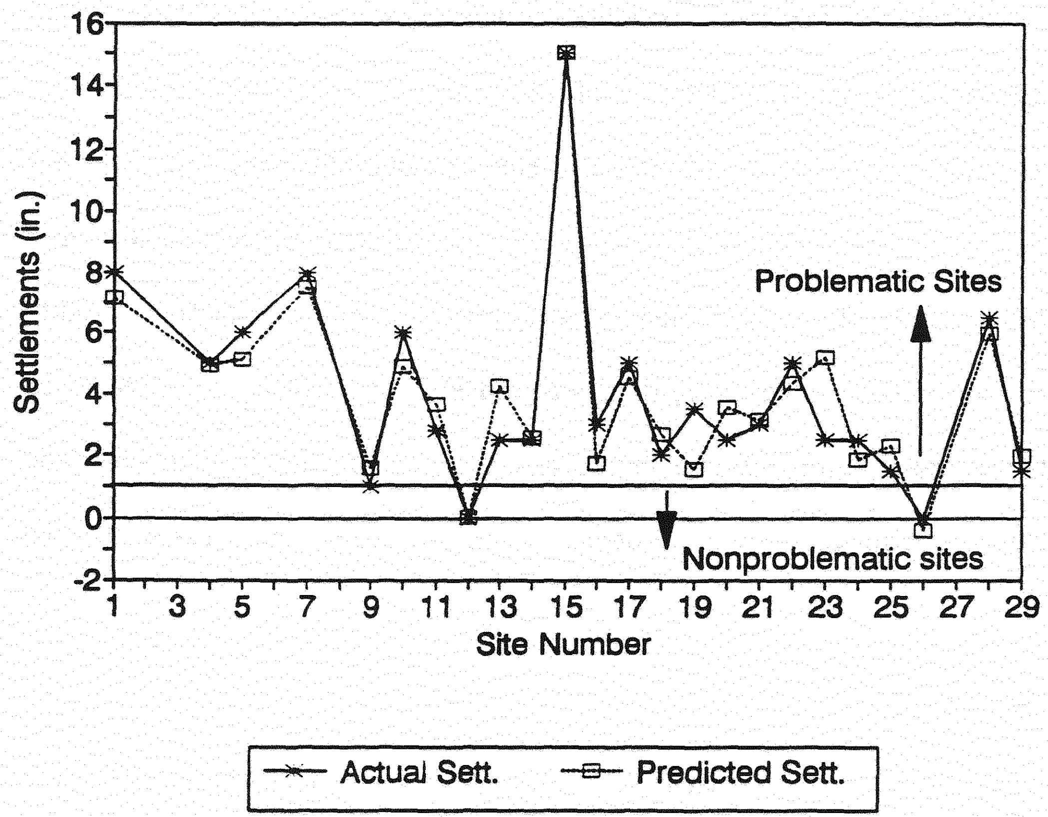


Fig. 5.2 Prediction obtained by the nonlinear model based on 24 sites.

5.3 Critical Assessment of Adequacy of the Model

It is observed that in comparison with the 11-variable linear multiple regression model (Eqn. 4.6), this model is comprised of less variables (only 5-variables) while maintaining a comparable $R^2_{adjusted}$. However, when this model was used to predict the settlements of the two excluded bridge sites (i.e., Site #3 and 8), the predicted settlements were obtained as -1.49 inch and -8.59 inches against their actual settlements of 18 inches and 10 inches, respectively (Fig. 5.3). Also, the model predicts both of these two sites as nonproblematic sites, whereas these sites are the most problematic sites of all. Such a poor prediction by the nonlinear model, with respect to identification as well as estimation, certainly poses a question mark on its validity as compared to the linear model. Therefore, the linear model (Eqn. 4.6) is concluded to be superior to the nonlinear model (Eqn. 5.2) in both the aspects (i.e., estimation of settlements and identification of problematic sites). Figure 5.3 depicts a comparative effectiveness of the two models.

5.4 Field Test Model

One of the main objectives of this study is to develop a model which can be utilized to identify problematic bridge sites prior to the construction of bridge structures. Accordingly, a multiple regression model based on the nonlinear algorithm of parameter estimation, referred to as the "Field Test" model is developed. As the name of the model suggests, this model utilizes data obtained from the field tests (SPT and CPT tests) as the only candidate regressors. The basic variables of the model are tip resistance, TIPR and friction ratio, FR (obtained from CPT test), and SPTF and foundation soil thickness, FND (from SPT test). It is important to note that in the case of linear regression model presented in Chapter IV, the variable SPT

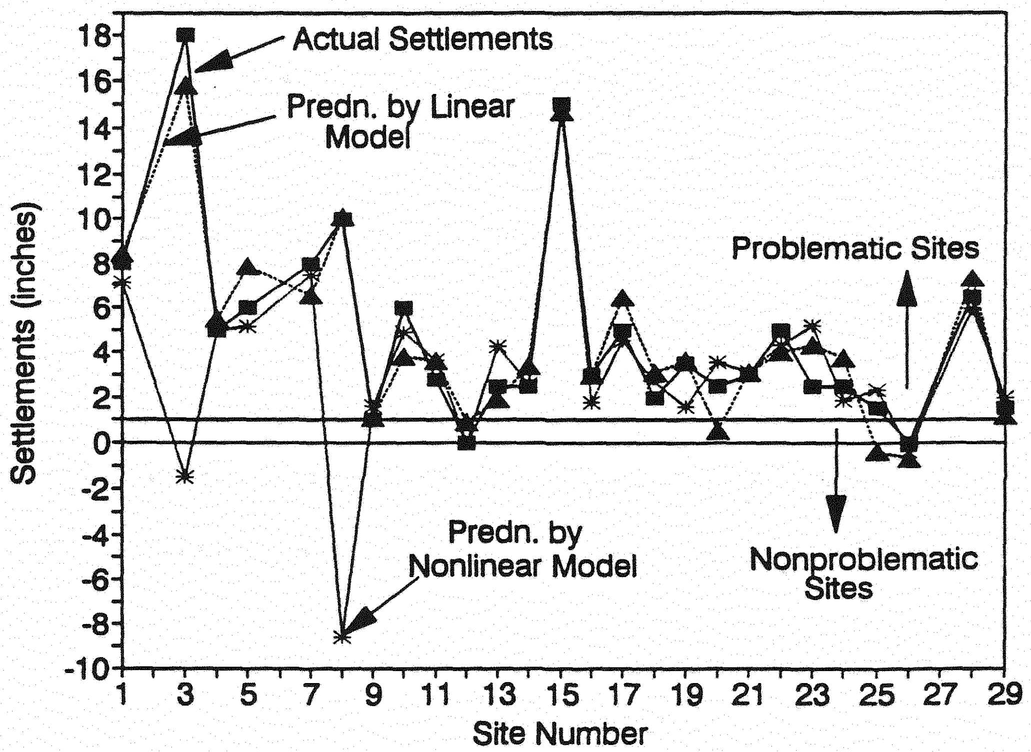


Fig. 5.3 Comparison of the predictive Ability of the Linear and the Nonlinear Models

represents the composite weighted average value of the SPT for both the embankment and the foundation, while SPTF, used in this model, represents the weighted average value of the SPT for the foundation portion only (refer to Sec. 2.4.2). The reason for using SPTF is further justified because there would be no embankment in place at this stage (e.g. a proposed bridge site). For the same reasons, the dependable variable in the model (left hand side in Eqn. 5.3) is denoted by TSTF (instead of TST used in chapter IV) to indicate that the predicted settlement represents the settlement contributed by the foundation only.

Following the general considerations outlined in the previous section (see Sec. 5.2), a series of trials were made using the SAS MODEL Procedure before arriving at the following best Field Test model:

$$\begin{aligned} TSTF = \; & 4.454\,(SPTF) - 0.2696\,(SPTF2) + 0.005\,(SPTF3) + 0.2152\,(TIPR) - \\ & 12.263\,(LTIPR) - 1.4267\,(FR) + 14.05\,(LFND) - 0.0782\,(FND2) + 0.002 \\ & (FND3) - 0.000014\,(FND4) \end{aligned} \tag{5.3}$$

$$(\mathbf{R^2_{adjusted} = 0.6738})$$

The parameter estimates, t-ratio, and the significance probability of the t-values are presented in Table 5.4. The detailed computer outputs obtained from the SAS MODEL Procedure are presented in Zaman et al., 1993. The $R^2_{adjusted}$ for this model is 0.6738, which is rather low as compared to the previous models. A lower value of $R^2_{adjusted}$ is expected here for the following reasons. The field settlement values used in Fig. 5.4 for comparison include the combined response of the foundation and the embankment, while the predicted response represents the foundation settlement only.

Table 5.4 Parameter estimates, t-values, and their significance probability for the field test model (Eqn. 5.3).

Variable	Parameter	Estimate	t-Ratio	Prob > |t|
SPTF	A1	4.45407	3.90	0.0025
SPTF2	A2	-0.269586	-3.79	0.0030
SPTF3	A3	0.00502156	3.57	0.0044
TIPR	A4	0.215150	4.73	0.0006
LTIPR	A5	-12.262695	-4.17	0.0016
FR	A6	-1.426732	-3.25	0.0077
LFND	A7	14.053215	4.00	0.0021
FND2	A8	-0.078197	-3.88	0.0026
FND3	A9	0.00204066	3.75	0.0032
FND4	A10	-0.00001453	-3.59	0.0042

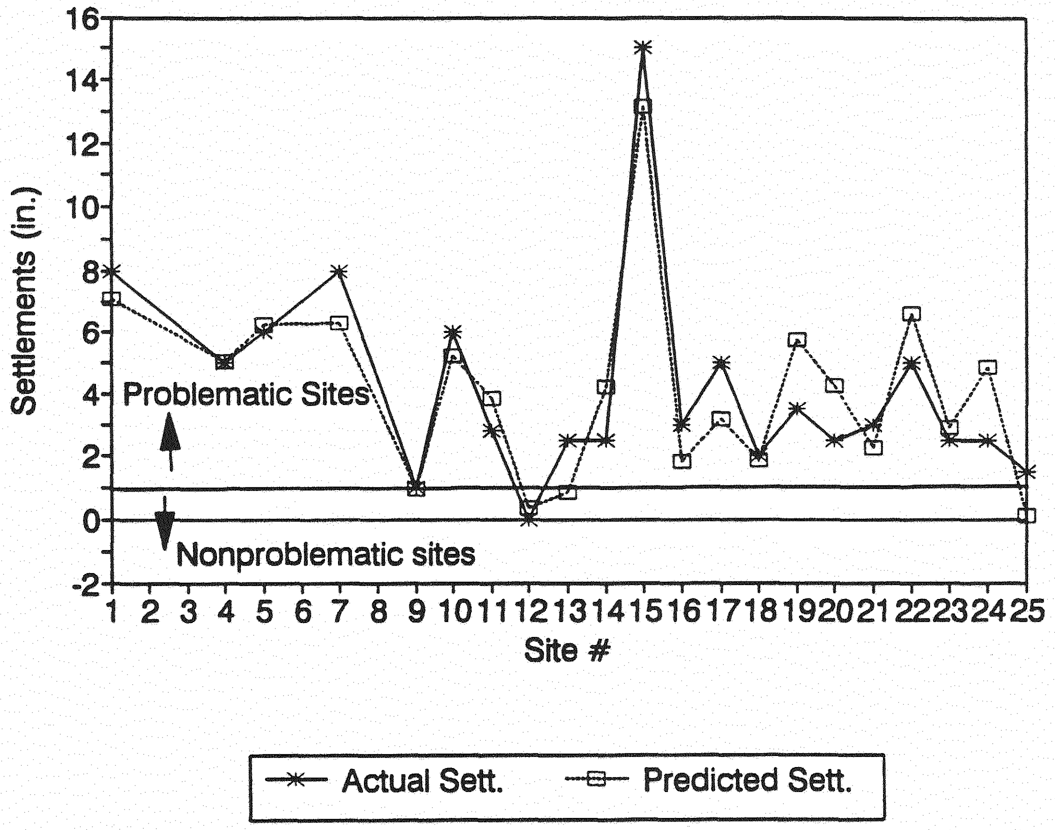


Fig. 5.4 Prediction obtained by the field test model.

Arguably, this is not a fair comparison, but other recourse is not feasible because of the lack of separate field data (settlement) for the embankment and the foundation. Nonetheless, the field test model accurately predicts the problematic and nonproblematic sites to a fairly reasonable level. As can be seen from Fig. 5.4, only one site, namely Site #25, was predicted nonproblematic, while in reality it is a problematic site. Conclusively, the field test model may not be very useful in estimating approach settlements; however, this model can be very reliable and effective in identifying problematic bridge sites. An important advantage of this type of statistical model is that it can be used to estimate the expected approach settlement at a site before the construction of the bridge. If the estimated settlement is found to be large, appropriate remedial measures could be undertaken such as preloading, ground improvement and non-conventional design. Further, this model can be used to identify if the settlement of embankment itself might be one of the significant contributors to the bridge approach settlement by comparing the TST and TSTF values.

CHAPTER VI

CONCLUSIONS AND RECOMMENDATIONS

6.1 Summary

This report presents the results of statistical analyses conducted on the data set of 26 selected bridge approach sites in Oklahoma (See Table 2.1). The majority of the data were obtained from the field tests (e.g. CPT and SPT tests) which were conducted as a part of level-two survey of these sites. However, some of the data such as age and skewness of the approach, traffic count (ADT) and height of the embankment were obtained from the level-one survey and or Bridge Division, ODOT. The selected sites encompass a wide range of variation with respect to the total settlement and the various causative factors/parameters such as age of the approach, traffic count, skewness of the approach, height of the embankment, thickness of the foundation soil underlying the embankment, and characteristics of the embankment and the foundation soil. The site-specific SPT and CPT results are used for the quantitative representation of the embankment and foundation soil characteristics. Extensive statistical analyses were performed using a statistical package program, called SAS. The following SAS Procedures were opted: REG, RSQUARE, STEPWISE and MODEL. Both linear and nonlinear multiple regression models were developed to predict the approach settlements. The multiple regression analyses were restricted to the use of quantitative variables only. The validity of the variables/parameters in the models is judged based on their level of significance, partial R^2 (square of coefficient of correlation), t- and F- statistics; while the best models are decided on the basis of their overall $R^2_{(adjusted)}$, Mallow's C_p-statistic, mean square error (MS_E), and their goodness of fit. The goodness of fit of the models is assessed based on their predictive capabilities and analysis of the residuals of the predicted values.

The statistical models developed in this study can be used to evaluate approach settlements at predominantly clayey sites. However, it is expected that these models may be applicable to sandy sites but the level of confidence is likely to be lower than that for the clayey sites because only a limited number of sandy/silty sites have been included so far in developing these models. Based on the estimate of settlements obtained from these models, a bridge site which is likely to present problems, referred to as problematic site, can be identified. In the present study, any site having a settlement of 1 inch or more is considered to be problematic. The special "Field Test Model" was also developed with the aim of identifying the problematic bridge sites before the construction of the bridge structure, so that appropriate precautionary and remedial measures such as ground improvement, preconsolidation (preloading) of the foundation, and use of select embankment materials can be implemented before or during the construction phase of the bridge. Moreover, the variables (SPT and CPT values) used in the field test model can be easily obtained from the Standard Penetration and Electric Cone Penetration tests.

6.2 Conclusions

Based on the data obtained and the observations made during the course of the extensive statistical analyses, the following conclusions are drawn:

1. The linear multiple regression model is the most reliable model in identifying problematic bridge sites as well as estimating approach settlements with a reasonable degree of accuracy. It has the

highest coefficient of correlation ($R^2_{adjusted}$=0.94) and fits the data set under consideration very well .

2. The coefficient of correlation ($R^2_{adjusted}$=0.88) of the nonlinear model based on 24 sites (e.g., after excluding 2 sites) and taking into account the number of parameters in the model, is quite comparable with that of the linear model. However, the model exhibits extremely poor prediction, in identifying problematic sites as well as estimating the approach settlement, when it was used for the two excluded sites (Fig. 5.3). Therefore, the linear model is considered to be superior to the nonlinear model.

3. Moreover, the nonlinear regression analyses indicate that the best nonlinear model is the one which tends to behave like a linear and/or close-to-linear model. Also, there seems to be absence of any specific nonlinear relationship between the approach settlement and the individual causative factors. Therefore, for the data set under consideration, the linear approach seems to be more reasonable.

4. The "Field Test Model" developed in this study can be used to identify the problematic bridge sites before the construction of the bridge so that appropriate precautionary and remedial measures such as ground improvement, and the preconsolidation (preloading)

of the approach foundation soil can be implemented before or during the construction phase.

5. The relatively low correlation coefficient ($R^2_{adjusted} = 0.6738$) in the case of the "Field Test Model" indicates that the embankment settlement is one of the major contributors to the approach settlement (about 33 percent).

6. The factors which significantly affect the approach settlement are: age of the approach, height of the embankment, traffic count (ADT), foundation soil thickness, and the embankment and the foundation soil macro characteristics.

7. There is no strong correlation between the bridge approach settlement and a single individual causative factor (variable), indicating that bridge approach settlement is a complex problem and several factors need to be considered in order to fully explain the phenomenon.

8. The site-specific SPT value, thickness of the foundation soil (FND), and cone resistance (q_c) are more significant variables as compared to the other variables (e.g., age of the approach, height of the embankment and traffic count) indicating that the site-specific embankment and the foundation soil characteristics are the most influential causative factors responsible for the approach settlement problems.

9. SPT value (N-value), tip resistance and friction ratio are found to be effective in representing the site-specific embankment and the foundation soil characteristics. However, the SPT value gives better correlation with the approach settlement as compared to the tip resistance (q_c).
10. Skewness of the approach embankment (SKEW) is found to have negligible effect on the approach settlement.

6.3 Recommendations

In view of the present study, the following studies are further recommended:

1. Data from additional sites, especially from nonproblematic sites, may be included for the enhanced reliability of the model.
2. A variable representing the effect of creep settlement on the approach embankment should be investigated.
3. Embankment soil characteristics have to be considered at a microscale; these include dispersity properties of clays, especially in the unsaturated condition, and their tendency under hydraulic action to erode and create voids.
4. The effect of slope erosion of the embankment may be incorporated into the model.

REFERENCES

1. Amick, D.J. and Walberg, H.J. (1975), Introductory Multivariate Analysis, Mc Cutchan Publishing Corporation.

2. Ardani, A. (1987), Bridge Approach Settlement, Colorado Department of Highways, Report No. CDOH-DTP-R-87-6.

3. Bates, D.M. and Watts, D.G. (1988), Nonlinear Regression Analysis and its Application, John Wiley & Sons, New York.

4. Belsley, D.A., Kuh, E., and Welsch, R.E. (1980), Regression Diagnostics: identifying influential data and sources of collinearity, John Wiley & Sons, New York.

5. Bendel, R.B. and Afifi, A.A. (1974), "Comparison of stopping rules in forward stepwise regressions," presented at the Joint Statistical Meeting, St. Louis, Missouri.

6. Bowles, J. E. (1988), Foundation Analysis and Design, 4th ed., McGraw-Hill, Oklahoma City.

7. Caulcutt, R. (1983), Statistics in Research and Development, Chapman and Hall, New York.

8. Chatterjee, S. and Price, B. (1977), Regression Analysis by Example, John Wiley & Sons, New York.

9. Cook, R.D. (1977), "Detection of influential observation in linear regression," Technometrics, **19**, pp. 15-18.

10. Cook, R.D. (1979), "Influential observations in linear regression," J. Am. Statist. Assoc., **74**, pp. 169-174.

11. Cox. D.R. and Snell, E.J. (1974), "The Choice of Variables in Observational Studies," Appl. Statist., **23**, pp. 51-59.

12. Daniel, C. and Wood, F.S. (1980), Fitting Equations to Data, 2nd. ed., Wiley, New York.

13. Das, B. M. (1990), Principles of Foundation Engineering, 2nd edition, PWS-KENT, Boston.

14. Douglas, B. J., and Olsen, R. S. (1981), Soil Classification Using Electric Cone Penetrometer, Symposium on Cone Penetration Testing and Experience, Geotechnical Engineering Division, ASCE, St. Louis, pp. 209-227.

15. Edgar, T. V., Puckett, J. A., and D'Spain, R. B. (1989), Effects of Geotextiles on Lateral Pressure and Deformation in Highway Embankments, Report No. FHWA-WY-89-001, The University of Wyoming and Wyoming Highway Department, Laramie.

16. Gopalasingam, A. (1989), Analysis of Consolidation Settlement of Bridge Approach Foundation by Nonlinear Finite Element Method including Infinite Elements, M. S. Thesis, University of Oklahoma.

17. Gunst, R.F. and Mason, R.L. (1980), Regression Analysis and its Application (A data oriented approach), Marcel Dekker, Inc.

18. Hopkins, T.C. and Scott, G.D. (1969), Estimated and Observed Settlements of Bridge Approaches, Highway Research Record #302, Division of Research, Kentucky Department of Highways.

19. Hopkins, T.C. and Deen, R.C. (1969), The Bump at the End of the Bridge, Highway Research Record #302, Division of Research, Kentucky Department of Highways.

20. Hopkins, T.C. (1973), Settlement of Highway Bridge Approaches and Embankment Foundations, Kentucky Department of Highways, Rep. No 356.

21. Hopkins, T. C. (1985), Long-term Movements of Highways Bridge Approach Embankments and Pavements, Report UKTRP-85-12, Kentucky Transportation Research Program, Lexinton, Kentucky.

22. Hunt. R. E. (1986), Geotechnical Engineering Analysis and Evaluation. McGraw-Hill, New York.

23. Jones. C.W. (1959), "Smoother Bridge Approaches," Civil Engineering, Vol. 29. June, pp. 407-409.

24. Kennedy, W.J. and Bancroft, T.A. (1971), "Model-building for prediction in regression using repeated significance tests," Ann. Math. Statist., **42**, pp. 1273-1284.

25. Kramer, S.L. and Sarej,P. (1991), Bridge Approach Slab Effectiveness, Final Report, TRAC Research Project GC 8286, Task 35, University of Washington.

26. Laguros, J. G., Zaman, M. M. and Mahmood, I. U. (1986), Evaluation of Causes of Excessive Settlements of Pavements Behind Bridge Abutments and Their Remedies-Phase I, ODOT Study 84-12-2, ORA 155-857, University of Oklahoma.

27. Laguros, J. G., Zaman, M. M. and Mahmood, I. U. (1990), Evaluation of Causes of Excessive Settlements of Pavements Behind Bridge Abutments and Their Remedies-Phase II, ODOT Study 86-04-2, ORA 157-293, University of Oklahoma, Report No. FHWA/OK 89 (06).

28. Laguros. J. G.. Zaman. M. M. and Alvapillai. A., Vaparapis, K.E. (1991), Evaluation of Causes of Excessive Settlements of Pavements Behind Bridge Abutments and Their Remedies-Phase III. ODOT Study 2163, ORA 157-829, University of Oklahoma.

29. Mahmood, I.U. (1990). Evaluation of Causes of Bridge Approach Settlement and Development of Settlement Prediction Models, Ph. D. dissertation, University of Oklahoma.

30. Montgomery, D.C. and Peck, E.A. (1992), Introduction to Linear Regression Analysis, 2nd. ed., John Wiley & Sons, Inc., New York.

31. Myers, R.H. (1990), Classical and Modern Regression with Applications, 2nd. ed., PWS-Kent Publishers, Boston.

32. Ratkowsky, D.A. (1983), Nonlinear Regression Modeling, Marcel Dekker, Inc., New York.

33. Robertson, P.K. and Campanella, R.G. (1984), Guidelines for Use and Interpretation of the Electric Cone Penetration Test, Hogentogler & Company, Inc., Maryland.

34. SAS/ETS User's Guide Version 5 Edition, SAS Institute Inc., Cary, North Carolina, 1984.

35. SAS/ETS User's Guide Version 6, SAS Institute Inc., Cary, North Carolina, 1988.

36. SAS User's Guide: Statistics Version 5 Edition, SAS Institute Inc., Cary, North Carolina, 1985.

37. SAS System for Linear Models, SAS Institute Inc., Cary, North Carolina, 1986.

38. Seber, G.A.F. and Wild. C.J. (1989), Nonlinear Regression, John Wiley and Sons, New York.

39. Stewart, C.F. (1985), Highway Structure Approaches, Report No. FWHA/CA SD-85-05, California Department of Transportation, Sacramento, California.

40. Tadros, M.K. and Benak, J.V. (1989), Bridge Abutment and Approach Slab Settlement, Phase I, Final Report, University of Nebraska-Lincoln.

41. Takeuchi, K., Yanai, H., and Mukherjee, B.N. (1982), The Foundation of Multivariate Analysis, Wiley Eastern Limited, New Delhi, India.

42. Timmerman, D.H. (1976), An Evaluation of Bridge Approach Design and Construction Practices, Report No. OHIO-DOT-03-77, University of Akron, Akron.

43. Vavarapis, K.E. (1991), Factors Influencing Bridge Approach Settlement, M.S. Thesis, University of Oklahoma.

44. Wahls, H.E. (1990), Design and Construction of Bridge Approaches, NCHRP, Synthesis of Highway Practice 159.

45. Wolde-Tinasae, A.M., Aggour, M.S., and Chini, S.A. (1987), Structural and Soil Provisions for Approaches to Bridges, Phase I, Final Report, University of Maryland.

46. Zaman, M.M. and Laguros, J.G. (1990), Progress Report on the Definition of Bump, ODOT, University of Oklahoma.

47. Zaman, M.M., Gopalasingam, A. and Laguros, J.G. (1991), "Consolidation Settlement of Bridge Approach Foundation," ASCE, Geotechnical Engineering Division, Vol. 117 No.2, 1991, pp. 219-240.

48. Zaman, M.M., Laguros, J.G. and Jha, R.K. (1993), Statistical Models for Identification of Problematic Bridge Sites and Estimation of Approach Settlements, Progress Report, ODOT Study 2188, ORA 125-6074, University of Oklahoma.

APPENDIX I

NUMERICAL VALUES OF THE VARIABLES USED IN MODEL BUILDING

SITE#	TST	LTCN	FND4	SPT4	LTIPR	LEHT	LAOE
1	8.0	3.681	20736	14641	1.53	1.398	1.531
3	18.0	3.176	12960000	6561	1.66	1.342	1.204
4	5.0	3.398	26873856	625	1.25	1.079	1.322
5	6.0	3.699	6250000	20736	1.59	1.000	1.462
7	8.0	2.653	2560000	28561	1.33	0.903	1.114
8	10.0	3.672	50625	279841	1.65	1.505	1.531
9	1.0	3.568	456976	10000	1.88	1.415	1.362
10	6.0	3.658	923521	14641	1.50	1.544	1.505
11	2.8	2.813	4100625	65536	1.86	1.301	1.342
12	0.0	3.477	104976	83521	1.65	1.322	0.699
13	2.5	4.124	625	14641	1.58	1.519	1.544
14	2.5	3.204	2085136	2401	1.21	1.556	0.845
15	15.0	4.000	4477456	390625	2.17	1.477	1.380
16	3.0	3.653	160000	1296	1.40	1.204	1.643
17	5.0	3.279	6250000	1296	1.39	1.114	1.230
18	2.0	3.398	3111696	234256	1.62	1.380	1.204
19	3.5	3.785	6561	10000	1.87	1.431	1.431
20	2.5	3.000	2560000	625	1.22	1.176	0.954
21	3.0	3.491	614656	28561	1.63	1.176	1.748
22	5.0	3.255	6250000	160000	1.87	1.041	1.724
23	2.5	2.954	2560000	28561	1.71	1.000	1.756
24	2.5	3.114	194481	6561	1.80	1.114	1.255
25	1.5	3.204	390625	83521	1.74	0.903	1.580
26	0.0	4.000	810000	28561	2.08	1.079	1.079
28	6	3.398	160000	28561	1.64	1.544	1.613
29	1.5	3.613	390625	38416	1.94	1.415	1.756

FND	SPT2	SPT3	FND2	FND3
12	121	1331	144	1728
60	81	729	3600	216000
72	25	125	5184	373248
50	144	1728	2500	125000
40	169	2197	1600	64000
15	529	12167	225	3375
26	100	1000	676	17576
31	121	1331	961	29791
45	256	4096	2025	91125
18	289	4913	324	5832
5	121	1331	25	125
38	49	343	1444	54872
46	625	15625	2116	97336
20	36	216	400	8000
50	36	216	2500	125000
42	484	10648	1764	74088
9	100	1000	81	729
40	25	125	1600	64000
28	169	2197	784	21952
50	400	8000	2500	125000
40	169	2197	1600	64000
21	81	729	441	9261
25	289	4913	625	15625
30	169	2197	900	27000
20	169	2197	400	8000
25	196	2744	625	15625