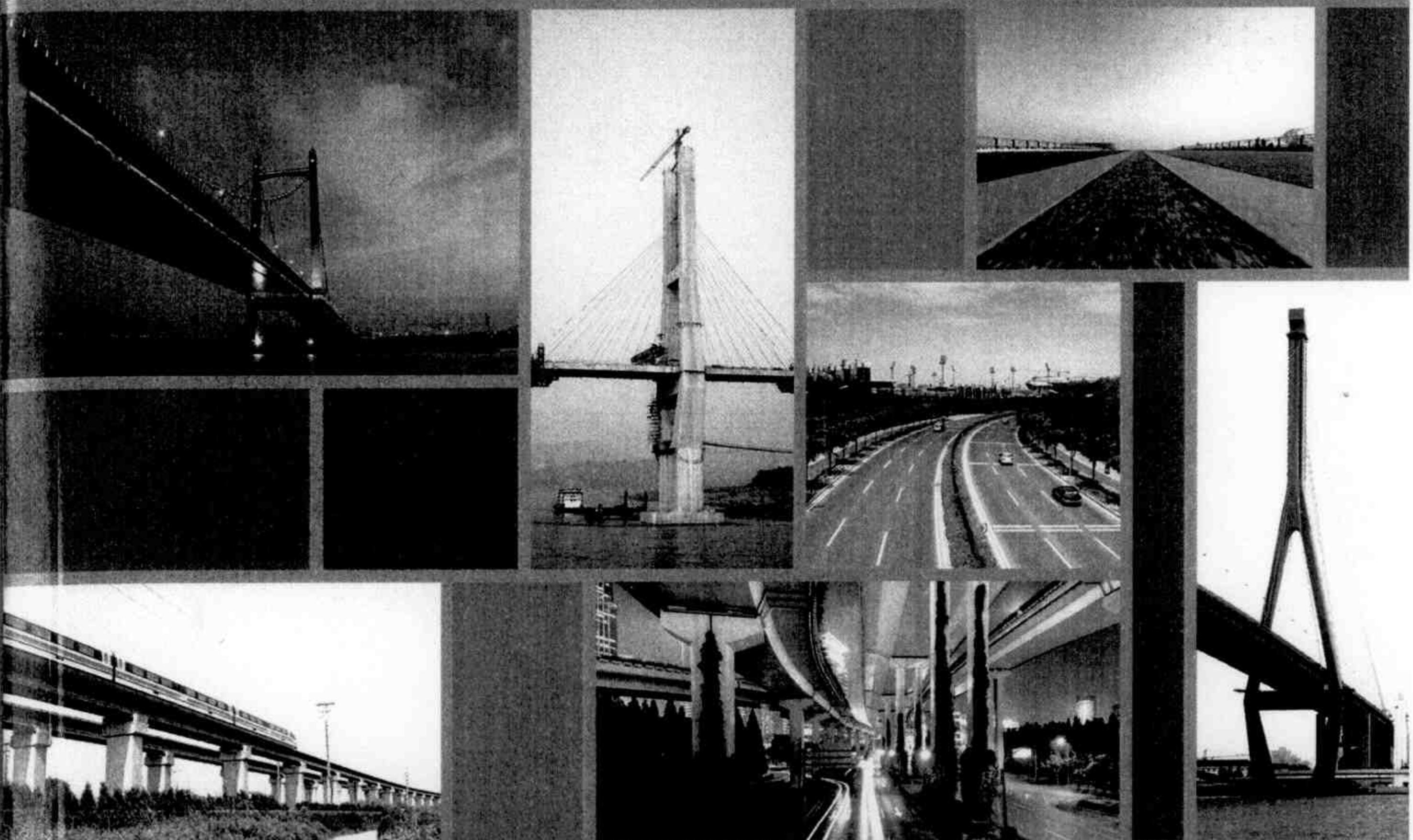

Challenges and Advances in Sustainable Transportation Systems

PLAN, DESIGN, BUILD, MANAGE, AND MAINTAIN

Edited by

Yong Bai, Ph.D., P.E.; Xiuli Du, Ph.D.;

Pei-Sung Lin, Ph.D., P.E.;

Wei-Chou Virgil Ping, Ph.D., P.E.;

T&DI TRANSPORTATION & DEVELOPMENT

CHALLENGES AND ADVANCES IN SUSTAINABLE TRANSPORTATION SYSTEMS: PLAN, DESIGN, BUILD, MANAGE, AND MAINTAIN

PROCEEDINGS OF THE 10TH ASIA PACIFIC TRANSPORTATION DEVELOPMENT CONFERENCE

Beijing, China
May 25–27, 2014

EDITED BY
Yong Bai, Ph.D., P.E.
Xiuli Du, Ph.D.
Pei-Sung Lin, Ph.D., P.E.
Wei-Chou Virgil Ping, Ph.D., P.E.
Endi Zhai, Ph.D., P.E.
Yilei Huang, Ph.D.

SPONSORED BY
International Chinese Transportation Professionals Association
Beijing University of Technology
Transportation & Development Institute (T&DI) of the
American Society of Civil Engineers

ASCE

TRANSPORTATION
& DEVELOPMENT
INSTITUTE

Published by the American Society of Civil Engineers

Cataloging-in-Publication Data on file with the Library of Congress.

Published by American Society of Civil Engineers
1801 Alexander Bell Drive
Reston, Virginia, 20191-4382
www.asce.org/bookstore | ascelibrary.org

Errata: Errata, if any, can be found at http://dx.doi.org/10.1061/9780784413364.

ISBN 978-0-7844-1336-4 (paper)
Manufactured in the United States of America.

Applications of Artificial Neural Networks to Pavement Prediction Modeling: A Case Study

Ying-Haur Lee[1], Hsiang-Wei Ker[2], and Yao-Bin Liu[2]

[1]Department of Civil Engineering, Tamkang University, New Taipei City 251, Taiwan; email: yinghaur@mail.tku.edu.tw
[2]Department of International Trade, Chihlee Institute of Technology, New Taipei City 220, Taiwan; email: hker@mail.chihlee.edu.tw
[3]T.Y.Lin Taiwan, Inc., Taipei City 106, Taiwan; email: ybliu@tylin.com.tw

ABSTRACT

Artificial neural networks (ANN) have been used in many pavement prediction modeling analyses. However, the convergence characteristics and model selection guidelines are rarely studied due to the requirement of extensive network training time. Thus, the techniques and applications of back propagation neural networks were briefly reviewed. Three ANN models were developed using deflection databases generated by factorial BISAR runs. A study of the convergence characteristics indicated that the resulting ANN model using all dominating dimensionless parameters was proved to have higher accuracy and require less network training time and data than the other counterpart using purely input parameters. Increasing the complexity of ANN models does not necessarily improve the modeling statistics. With the incorporation of subject-related engineering and statistical knowledge into the modeling process, reasonably good predictions may be achieved with more convincing generalization and explanation yet requiring minimal amount of time and effort.

INTRODUCTION

Predictive models have been widely used in various pavement design procedures, evaluation, rehabilitation, and network management systems. Empirical and mechanistic-empirical approaches using statistical regression techniques have been utilized extensively in predicting extremely complicated pavement responses and performance indicators for more than four decades. Using purely empirical concepts to develop predictive models is not recommended. Lee (1993) proposed a systematic statistical and engineering modeling approach which strongly recommends to incorporate theoretical engineering knowledge, expert experience, heuristics, and statistical data analysis and regression techniques altogether into the framework to develop more mechanistic-based predictive models. In additional to the conventional "parametric" linear and nonlinear regression techniques, several ingenious iterative regression techniques in the area of "robust" and "nonparametric" regressions were also incorporated. The proposed approach has been successfully implemented in the development of many purely mechanistic (Lee, 1993; Lee & Darter, 1994) and purely empirical predictive models (Lee & Darter, 1995), as well as the mechanistic -empirical predictive models adopted in the early analyses of LTPP general pavement studies data (Simpson et al., 1993).

Significant progress has been reported in pavement prediction modeling of simulated data using artificial neural networks (ANN). However, the convergence characteristics and model selection guidelines are rarely studied due to the requirement of extensive network training time. As part of continuous research efforts in pavement design and analysis (Lee & Darter, 1994; Lee et al., 2004), the techniques and applications of back propagation neural networks were briefly reviewed. Factorial BISAR runs for different pavement systems are conducted to generate the deflection databases for the analysis. Artificial neural networks were utilized to improve the prediction accuracy of simulated pavement deflections (Liu, 2004). This study strives to illustrate the benefits of incorporating the principles of dimensional analysis, subject-related knowledge, and statistical knowledge into modeling process.

THE CONCEPT OF ARTIFICIAL NEURAL NETWORKS

The concept of 'neural network' was originated by the work on 'perceptrons' around 1960. There were pictured as networks with a number of inputs x_i and an output (or outputs) y, where the inputs are connected to one or more neurons in the input layer and they are further connected in one or more hidden layers until they reach the output neuron. Artificial Neural Networks (ANN) provides a flexible way to generalize linear regression functions. They are nonlinear regression models but with so many parameters extremely flexible to approximate any smooth function. The most commonly used rule is the generalized delta rule or back propagation algorithm. Ripley (Ripley, 1993) provided the detail definitions and brief derivation of a back propagation network (BPN). The learning procedure has to select the weights and the biases by presenting the training examples in turn several times, while striving to minimize the total squared errors. However, the questions of how many layers and how many neurons should be used were treated very lightly in the literature.

Ripley (1993) also discussed many statistical aspects of neural networks and tested it with several benchmark examples against traditional and modern regression techniques, such as generalized discriminant analysis, projection pursuit regression, local regression, tree-based classification, etc. Ripley concluded that in one sense neural networks are little more than non-linear regression and allied optimization methods. "That two-layer networks can approximate arbitrary continuous functions does not change the validity of more direct approximations such as statistical smoothers, which certainly 'learn' very much faster (Ripley, 1993)." Projection pursuit regression highlights the value of differentiated units and other training schemes and offers computation shortcuts through forward and backward selection. Statistical and subject-related knowledge can be used to guide modeling in most real-world problems and so enable much more convincing generalization and explanation, in ways which can never be done by 'black-box' learning systems (Ripley, 1993).

BRIEF REVIEW OF ARTIFICIAL NEURAL NETWORKS APPLIED IN PAVEMENT PREDICTION MODELING

a multi-layered, feed-forward neural network trained by using an error back propagation algorithm or an error minimization technique (Haykin, 1994; Hecht-Nielsen1990). Ceylan (2004) conducted a literature search summarizing recent ANN applications in pavement structural evaluation such as backcalculating pavement layer moduli and predicting primary pavement responses (e.g., stress and deflection). As with many ANN applications in the literature, original pertinent input parameters were used to generate the training and testing databases. This approach often requires tremendous amount of time and efforts in network training and testing. To reduce the size of the required factorial databases, researchers sometimes opt to fix certain input parameters to some prescribed values as a special case study, which may result in limiting the inference space of the resulting model.

Nevertheless, some earlier ANN literature has also illustrated that the incorporation of the principles of dimensional analysis lead to significant savings during the training set generation. Ioannides et al. (1996) trained a back propagation neural network (BPN) to determine the in situ load transfer efficiency of rigid pavement joints from Falling Weight Deflectometer (FWD) data. Khazanovich and Roesler (1997) developed an ANN-based backcalculation procedure for composite pavements. The multilayer elastic program DIPLOMAT was used to analyze a three-layer pavement system consisting of an AC surface layer over a PCC slab resting on a Winkler foundation. Ioannides et al. (1999) trained BPN models to predict the critical slab bending stress for loading-only, curling-only, and loading-and-curling cases. BPN predictions were compared against the Westergaard closed-form solutions as well as the statistical regression models developed by Lee and Darter (1994) using a small set of factorial data with dimensionless mechanistic variables. It was re-emphasized that mature engineering judgment and in-depth understanding of the mechanics of the phenomenon remain the most reliable guides in the formation of the targeted problems.

Attoh-Okine (1994) proposed the use of ANN models in predicting roughness progression of flexible pavements. Although the results were promising, some built-in functions including learning rate and momentum term which form key neural network algorithm were not investigated. Attoh-Okine (1999) used real pavement condition and traffic data and specific architecture to investigate the effect of learning rate and momentum term on BPN models for predicting flexible pavement performance. Sorsa et al. (1991) indicated that adding many hidden layers gets the network to learn faster and the mean square error becomes a little smaller, but the generalization ability of the network reduces.

BENEFITS OF INCORPORATING STATISTICAL AND SUBJECT-RELATED KNOWLEDGE INTO THE MODELING PROCESS

To illustrate the benefits of incorporating statistical and subject-related knowledge into the modeling process, the following case studies were conducted using a more complicated database (Liu, 2004). A neural network modeling software package called Qnet v2000 for Windows (Vesta Services, Inc., 2000) was adopted for this study.

Development of Flexible Pavement Deflection Databases for the Analysis. Based on the multi-layer elastic theory (Huang, 2004) and the principles of identified for a three-layer pavement system: E_1/E_2, E_2/E_3, h_1/h_2, and a/h_2. In which, a is the radius of the applied load, (L); h_1 and h_2 are the thickness of the surface and base layers, (L); E_1, E_2, and E_3 are the Young's moduli of the surface layer, base layer, and subgrade, respectively, (FL^{-2}). A series of factorial BISAR runs was conducted with the following ranges to cover most practical pavement data: $0.5 \leqq E_1/E_2 \leqq 170$, $0.5 \leqq E_2/E_3 \leqq 170$, $0.2 \leqq h_1/h_2 \leqq 2.4$, and $0.5 \leqq a/h_2 \leqq 5.0$. A BASIC program written by Dr. Alaeddin Mohseni was used to automatically generate the input files and summarize the results to avoid untraced human errors. A pavement response database including the aforementioned dimensionless variables, deflections at the center of load (D_0), horizontal strain (ε_t and vertical strain (ε_v) at the bottom of the surface layer was obtained. A training database with 3,600 data points and an independent testing database with 1,728 data points were used in this study (Liu, 2004).

Development of ANN Models for Flexible Pavements. The training database was randomly separated into 3,400 data points for actual training and the remaining 200 observations for monitoring the training process. Hyperbolic tangent activation function was chosen in this case study (Vesta Services, Inc., 2000). The learning rate was set as 0.01. At the first trial (NET1) as shown in Table 1, no transformation was made on both explanatory and response variables. In which, the dependent variable is the maximum deflection (D_0), whereas the explanatory variables include E_1/E_2, E_2/E_3, h_1/h_2, and a/h_2. Nevertheless, extreme difficulty was encountered in obtaining reasonable convergence. Several other attempts were also conducted using simply the pertinent input parameters (such as a, h_1, h_2, E_1, E_2, and E_3) as the explanatory variables, but the results of ANN prediction modeling were even worse.

Based on the basic assumptions of conventional regression techniques that the random errors are mutually uncorrelated and normally distributed with zero mean and constant variance, and additive and independent of the expectation function, it is desirable to check the normality of the response variable. The Box and Cox (1964) transformation procedure was adopted to find the approximate power transformation of the response variable (D_0). The S-PLUS Statistical Analysis Software (Mathsoft, Inc., 1997) was used for this analysis. The maximum likelihood estimator λ of various power transformations (Weisberg, 1985) was approximately 0 indicating that a logarithm transformation was appropriate for D_0. The normal Q-Q plot which graphically compares the distribution of $\log(D_0)$ to the normal distribution represented by a straight line. This indicates that the logarithm of D_0 is approximate to normally-distributed. In the second trial (NET2), convergence was obtained though the number of learning cycles and modeling time were still very high. The root mean squared (RMS) errors were computed accordingly.

According to general statistical principles or using the alternating conditional expectations (ACE) algorithm (Mathsoft, Inc., 1997; Breiman & Friedman, 1985) together with the Box-Cox power transformation technique proposed by Lee (1993), logarithm transformations of D_0, E_1/E_2, and E_2/E_3 were recommended for NET3 model. As shown in Table 1, with more statistical knowledge incorporated into the ANN modeling process, the resulting ANN model was proved to have higher accuracy and less network training time than the other counterpart using purely input parameters. Figure 1(a) and 1(b) depict the network convergence results for NET2 and NET3 during the training process. The goodness of the prediction of log(D_0) for NET2 and NET3 were also provided in Figure 1(c) and 1(d) during the testing phase. The goodness of the prediction of D_0 for NET2 and NET3 were also provided in Figure 1(e) and 1(f) during the testing phase. With more statistical knowledge incorporated into the modeling process, the resulting ANN model was proved to have higher accuracy and less network training time than the other counterpart using purely input parameters.

Table 1. Comparison of Three Different ANN Models.

ANN Type	NET1	NET2	NET3
Outputs	D_0	Log(D_0)	Log(D_0)
Inputs	E_1/E_2, E_2/E_3, h_1/h_2, a/h_2	E_1/E_2, E_2/E_3, h_1/h_2, a/h_2	log(E_1/E_2), log(E_2/E_3), h_1/h_2, a/h_2
Hidden Layer(s)	3	3	2
Neurons in Each Hidden Layer	20-10-5	15-10-5	12-6
Learning Cycle	Cannot converge	200,000	27,000
Modeling Time	> 24 hrs	10 hrs	26 min
RMS	---	Training: 0.0048 Monitoring: 0.0045	Training: 0.0040 Monitoring: 0.0039

CONCLUSIONS

A case study was conducted to illustrate the benefits of incorporating statistical and subject-related knowledge into pavement prediction modeling process. The resulting ANN model using all dominating dimensionless parameters was proved to have higher accuracy and require less network training time and data than the other counterparts using purely input parameters. Increasing the complexity of ANN models does not necessarily improve the modeling statistics. The results also showed that using higher number of neurons and hidden layers sometimes lead to even worse modeling statistics which was an indication of over training and should be avoided. Statistical and subject-related knowledge can be used to guide modeling in most real-world problems and so enable much more convincing generalization and explanation, in ways which can never be done by 'black-box' learning systems (7).

ACKNOWLEDGMENTS

The financial support provided by the National Science Council, Taiwan is gratefully acknowledged.

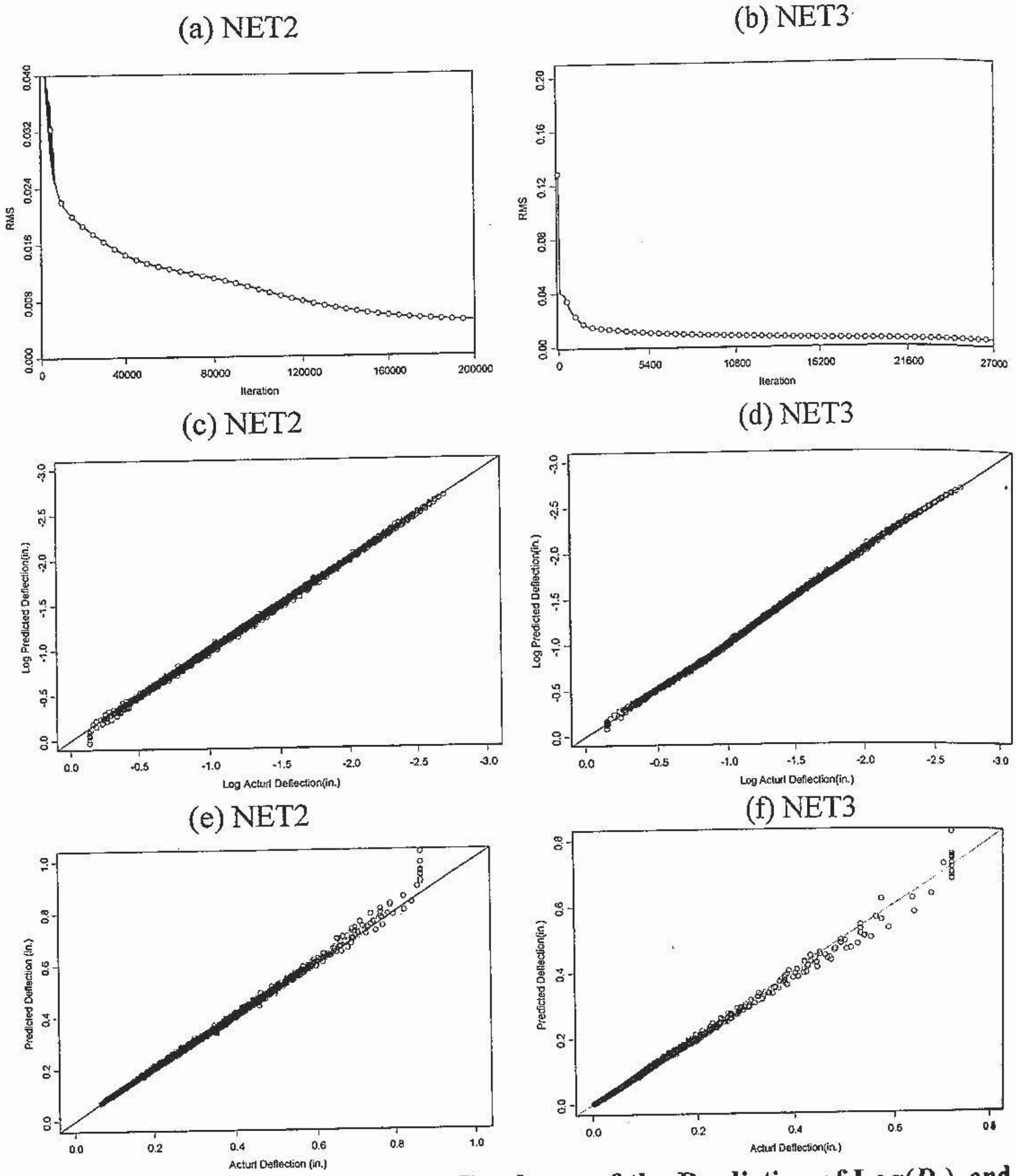

Figure 1. Convergence Results, Goodness of the Prediction of Log(D_0), and Prediction of D_0 for NET2 and NET3, Respectively.

REFERENCES

Attoh-Okine, N. O. (1994)."Predicting roughness progression in flexible pavements using artificial neural networks." *Proceedings of the Third International Conference on Managing Pavements*, San Antonio, TX, Vol. 1: 55-62.

Attoh-Okine, N. O. (1999). "Analysis of learning rate and momentum term in backpropagation neural network algorithm trained to predict pavement performance." *Advances in Engineering Software*, 30(4): 291-302.

Box, G. E. P., and Cox, D. R. (1964). "An analysis of transformations (with discussion)." *Journal of the Royal Statistical Society*, B, 26: 211-246.

Breiman, L. and Friedman, J. H. (1985). "Estimating optimal transformations for multiple regression and correlation (with discussion)." *Journal of the American Statistical Association*, 80: 580-619.

Ceylan, H. (2004). "Use of artificial neural networks for the analysis & design of concrete pavement systems." *Proceedings of the 5th CROW Workshop on Concrete Pavements*, Istanbul, Turkey.

Haykin, S. (1994). *Neural networks: A comprehensive foundation*. Prentice-Hall, Inc., New Jersey.

Hecht-Nielsen, R. (1990). *Neurocomputing*. Addison-Wesley, New York.

Huang, Y. H. (2004). *Pavement analysis and design*, 2nd Ed., Prentice Hall, New Jersey.

Ioannides, A. M., Alexander, D. R., Hammons, M. I., and Davis, C. M. (1996). "Application of artificial neural networks to concrete pavement joint evaluation." *Transportation Research Record, 1540*: 56-64.

Ioannides, A. M., Davis, C. M., and Weber, C. M. (1999). "Westergaard curling solution reconsidered." *Transportation Research Record, 1684*: 61-70.

Khazanovich, L., and Roesler, J. (1997). "DIPLOBACK: Neural-network-based backcalculation program for composite pavements." *Transportation Research Record, 1570:* 143-150.

Lee, Y. H. (1993). *Development of pavement prediction models*, Ph.D. Dissertation, University of Illinois, Urbana, IL.

Lee, Y. H., and Darter, M. I. (1994). "New predictive modeling techniques for pavements." *Transportation Research Record, 1449:* 234-245.

Lee, Y. H., and Darter, M. I. (1995). "Development of performance prediction models for Illinois continuously reinforced concrete pavements." *Transportation Research Record, 1505:* 75-84.

Lee, Y. H., Wu, H. T., and Yen, S. T. (2004). "Parameter studies on three-dimensional finite element analysis of rigid pavements." *Proceedings of the 5th CROW Workshop on Concrete Pavements*, Istanbul, Turkey.

Liu, Y. B. (2004). *Application of modern regression techniques and neural network on rigid pavement backcalculation*, M. S. Thesis, Tamkang University, Taipei, Taiwan (In Chinese).

Mathsoft, Inc. (1997). *S-PLUS for Windows (Ver. 4.0) User's manual, reference manual, and guide to statistics.*

Ripley, B. D. (1993)."Statistical aspects of neural networks." *Networks and chaos - statistical and probabilistic aspects*, edited by Barndorff-Nielsen, O. E., Jensen, J. L., and Kendall, W. S., Chapman & Hall, London, 41-123.

Simpson, A. L., Rauhut, J. B., Jordahl, P. R., Owusu-Antwi, E., Darter, M. I., Ahmad, R., Pendleton, O. J., and Lee, Y. H. (1993). *Early analyses of LTPP general pavement studies data, Volume 3 - Sensitivity analyses for selected pavement distresses*, Strategic Highway Research Program, Contract No. P-020, Report No. SHRP-P-393, National Research Council, Washington, D.C.

Sorsa, T., Koivo, H. N., and Koivisto, H. (1991). "Neural networks in process fault diagnosis." *IEEE Transactions on Systems, Man and Cybernetics*, 21(4): 815-825.

Vesta Services, Inc. (2000). *Qnet v2000 for Windows.*

Weisberg, S. (1985). *Applied linear regression.* Second Edition, Wiley Series in Probability and Mathematical Statistics, John Wiley & Sons, Inc.