SOFT COMPUTING TECHNIQUE IN PREDICTION OF PAVEMENT CONDITION

DEVINDER KAUR AND DEBARGHA DATTA
Department of Electrical Engineering and Computer Science and Department of Civil Engineering
University of Toledo
Mail Stop 308, W. Bancroft, Toledo, Ohio – 43606
UNITED STATES OF AMERICA

Abstract: – This paper presents a soft computing technique using neuro fuzzy approach to predict the future pavement condition based on the current pavement age and current pavement condition. The Ohio Department of Transportation (ODOT) database for the asphalt pavement sections of Interstates and US routes was used to build the prediction model. Both grid partitioning and subtractive clustering based pattern recognition followed by back propagation learning algorithm was followed to build and optimize the models. The performances of both these models were compared with the conventional Markov chain method of pavement performance prediction. The study reveals that grid partitioning based model outperforms both the Markov chain model and the subtractive clustering based model.

Key-Words: - ANFIS, grid partitioning, subtractive clustering, Markov chain, PCR, pavement age.

1 Introduction

Prediction of future pavement performance of a road network is a key step in a pavement management system. It has significant influence on investment decisions taken at both network and project levels for maintenance and repair actions. However, data based on which prediction of performance and other analyses are done are sometimes ambiguous and incomplete [1]. The parameters that affect the pavement performance are also not definite and stochastic approach is one way to deal with such uncertainties. Beside the most popular regression analysis, Markov chain process is another stochastic method, which is applied in prediction of pavement performance. The increasing popularity of use of soft computing in decision analysis and optimization problems encourages many to apply the knowledge in developing time-series prediction models [2] and prioritization models [3]. The concepts of fuzzy logic, adaptive artificial neural network, genetic algorithm and grammatical evolution come under the umbrella of soft computing techniques. Sometimes, a hybrid approach using all these techniques to discover the knowledge from complex databases is also followed [4]. All these artificial intelligence techniques help to take into account the partial truth, which underlies in the system and when human expertise either becomes increasingly difficult to find or they fails to identify such hidden knowledge from the database system [6].

The objective of this paper is to apply the concept of soft computing in building a model for prediction of pavement condition and check its efficacy with respect to the other conventional prediction model. Adaptive neuro fuzzy inference system (ANFIS) was used to develop the pavement performance prediction model. The initial membership functions and the fuzzy rules were generated from the data using both grid partitioning and subtractive clustering pattern recognition methods. This was followed by training the two different models separately by back propagation (BP) learning algorithm. Then these two different models were validated using different pavement condition dataset.

The asphalt pavement condition database for the interstates and US routes in Ohio, which is available from the Ohio Department of Transportation (ODOT), was used to train, check and evaluate the model. As asphalt pavement sections built prior to 1997 had the same design specification, therefore, 1991 – 1996 pavement condition data was used to train and check the model and 1986 – 1990 condition data was used to evaluate the model.

ANFIS toolbox available in Matlab 6.5 © was used to generate both the grid partitioning based model (henceforth we call it MODEL-I) and subtractive clustering based prediction model (henceforth we call it MODEL-II). After the models were built, Visual Basic program based macro was developed to invoke the models developed in Matlab from Microsoft Excel®. An initial pavement age at certain year (1986) and performance in that year (henceforth we call it as "Current Year") were entered in a spreadsheet and the macro was used to perform necessary prediction analysis using the FIS matrices within MS Excel. The output, i.e. the pavement condition for 1987-1990, was generated in the excel spreadsheet only.

In order to compare both MODEL-I and MODEL-II with a conventional prediction method of pavement performance, Markov chain model was developed and used to predict the condition for the same analysis period (1987-1990). The accuracy of prediction made by all these three models was checked against the actual pavement condition data for the year 1987 – 1990 available in the ODOT database. The estimated root mean square errors (RMSE) for the models were used to compare their efficiency in prediction. Results revealed that MODEL-I outperforms both the MODEL-II and Markov chain model. The detailed analysis and results are discussed in the following sections.

2 Background

2.1 Pavement Performance

Performance of pavement is assessed from the observable distresses in the pavement. The distresses are quantified with some weights, which are then converted to some scores like pavement condition rating (PCR) or Pavement Condition Index (PCI). ODOT follows the PCR scoring method following the mathematical expression as given below:

$$\text{PCR} = 100 - \sum_{i=1}^{n} Deduct_i \qquad (1)$$

where

n = number of observable distresses, and Deduct = (Weight for distress) x (Weight for severity) x (Weight for Extent)

The method of understanding the assessment of the weights is beyond the scope of this study and hence not discussed here. However, necessary guidelines regarding PCR calculation are available in the ODOT website [5]. Qualitative assessment of the pavement condition is done based on this PCR score. Generally, PCR scores are classified into five different categories, viz. “very good”, “good”, “fair”, “poor” and “very poor”. By definition, the higher is the PCR score, the better is the pavement condition. The PCR score drops with the drop in the pavement condition. However, in this study the condition was classified into three categories, viz. “good”, “fair” and “poor”. The PCR classification adopted for this present study is shown in Fig.1.

Pavement age, traffic volume and traffic axle loading, climate, design specification, type of pavement, i.e. asphalt, concrete, or composite type of pavement, and quality of materials used in the construction of the pavement are few parameters that affect the pavement performance. The extent of influence of each parameter on the pavement performance is somewhat ambiguous. In addition to these, there also exist some latent parameters, which sometimes have considerable influence on the condition of pavement. These parameters are very difficult to capture and quantify them in the real world. Therefore, adaptive neuro fuzzy approach may help to capture the ambiguous relationship between the parameters and the pavement condition. Back propagation learning algorithm helps the model to adapt in the prevailing deterioration system.

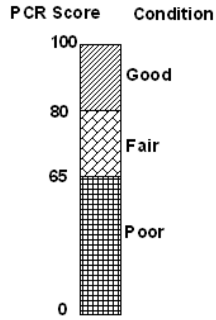


Fig.1 PCR vs. Pavement Condition classification

Sometimes, fuzzy logic based approach is also followed in modeling a deterioration system where the membership functions and fuzzy rules are determined using the knowledge base [2, 4, 6]. However, in order to simulate the actual deterioration system it is necessary to use the actual historical data. The advantage of using ANFIS over simple fuzzy logic approach lies here. The membership functions and the fuzzy rules are generated from the field data and after adequate training, the model approximately simulates the prevailing deterioration system.

2.2 Markov chain process

Markov process is a stochastic time dependant process, where the probability of transition from the current states i to the immediately future state j depends on the current state only [8]. The transition probability p_{ij}, the probability of transition from state i to state j, is given from every possible combination of i and j (including i = j) and the transition probabilities are assumed to be stationary (unchanging) over the time period of interest and independent of how state i was reached. Mathematically, a set of random variables $\{X_{t_n}\} = \{x_1, x_2, ...x_n\}$ in a given chronological times $\{t_0, t_1, ...t_n\}$ is said to be in Markov process if

$$P\{X_{t_n} = x_n \mid X_{t_{n-1}} = x_{n-1}, ...X_{t_0} = x_0\}$$
$$= P\{X_{t_n} = x_n \mid X_{t_{n-1}} = x_{n-1}\} \quad (2)$$

The probability of transition from any specific state i to any other state j over time t = 0, 1, 2….T is usually given by

$$p_{ij} = P\{X_t = j \mid X_{t-1} = i\} \quad (3)$$

where

(i, j) = 1, 2, 3,….n;
t = 0, 1, 2……T;
n = total no. of states that describes the system; and
T = the time period over which the state of the system is assessed

By definition, $\sum p_{ij} = 1, i = 1,2,....n$ and $p_{ij} \geq 0, (i, j) = 1,2,...n$

In matrix notation, Markov chain is expressed by the following transition probability matrix (TPM) **P**:

$$P = \begin{pmatrix} p_{11} & p_{12} & p_{13} & . & . & . & p_{1m} \\ p_{21} & p_{22} & p_{23} & . & . & . & p_{2m} \\ . & . & . & . & . & . & . \\ . & . & . & . & . & . & . \\ . & . & . & . & . & . & . \\ p_{n1} & p_{n2} & p_{n3} & . & . & . & p_{nm} \end{pmatrix} \quad (4)$$

Pavement deterioration system is also modeled by using the Markov chain process. The model requires initial pavement condition data and the Markov TPM generated by observing, at least for two consecutive years, the transition of actual pavement condition [9]. For an ideal deterioration system, $p_{ij} = 0, \forall i > j$. Therefore, given an initial pavement condition distribution p (0),

P (1) = p (0). **P**
P (2) = p (1).**P** = p (0).$\mathbf{P}^2$….

thus, after a time period of n,
p (n) = p (0).$\mathbf{P}^n$ (5)

The pavement reaches steady state after certain time period k, beyond which the overall network condition becomes almost constant.

Although Markov process is popular for its simplicity in calculation, but the effect of parameters, which are time variant like, traffic, weather and pavement age are not considered in this prediction process.

3 Project Details

3.1 Data

A gigantic Microsoft Access® based historical pavement database was available from ODOT. The database contains mileage-wise information regarding the type of pavement, type of road network, i.e. Interstates, US routes and state routes, its yearly performance in terms of PCR between 1985 to 2006, year of construction, weather, traffic load in terms of equivalent single axle load (ESAL) and several other attributes necessary for pavement management system. The pavement age, condition, traffic axle load and snowfall information that was used as input parameters in building and checking the ANFIS models were queried from this database. The spectrum of axle load data (ESAL) is comparatively much wider than its effect on pavement condition. Therefore, in order to scale down its range, log transformation of the data was used as input parameter of the model. The model was built only for the Interstates and US routes, which has asphalt type of pavement sections. Specification of construction of pavement was changed from 1997 onwards. Therefore, assuming that all the pre 1997 data come from same population, 1991 – 1996 data were used to train and check the model and 1986-1990 data were used to evaluate all the three different models.

3.2 Training and Checking of ANFIS Model

1991 – 1996 data were used in training and checking both MODEL-I and MODEL-II. Randomly chosen 75% of the data was considered to train the individual model and the remaining 25% was used to check the model. Initially, four input parameters, viz. pavement age, logarithm of ESAL, snowfall and current year PCR were considered. The output of the model was the predicted PCR for the following year (henceforth we call it "succeeding PCR"). Ideally, with increase in traffic load and snowfall the pavement condition should deteriorate and thus, there should be some drop in the PCR score. Therefore, both the "SuccPCR" vs. $\log_{10}$(ESAL) and "SuccPCR" vs. Snowfall graphs should show some decreasing trend. However, results obtained and shown in Fig. 2 and Fig. 3 indicate that for Interstates and US routes, the traffic load and snowfall do not have any significant effect on the deterioration system. The checking error obtained in this 4-input model was also significantly high.

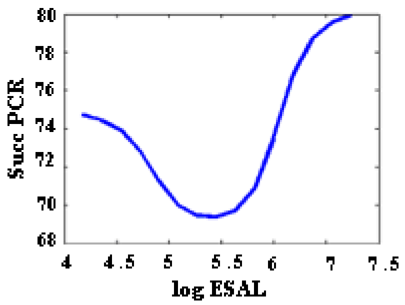


Fig.2 Succeeding Year's PCR vs. $\log_{10}$(ESAL)

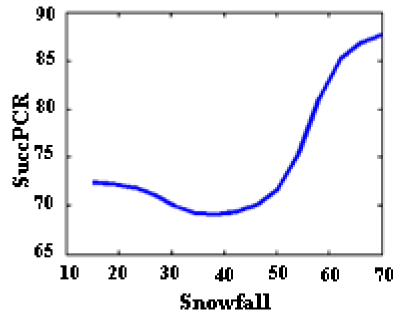


Fig.3 Succeeding Year's PCR vs. Snowfall

Therefore, finally two input parameters, viz. pavement age and current year's pavement condition were considered in both the MODEL-I and MODEL-II. In MODEL-I, each input was assumed to have 3 nos. of membership functions, thus evolving, 3 x 3 i.e. 9 rules (see Fig.4). As no definite membership functions were known to fit both these parameters, several trials were made with different membership functions. The membership functions were optimized by back propagation learning method and the checking error was estimated. Eventually, based on the estimated least error, Gaussian bell membership function was chosen as the ideal membership function. Besides, as no particular polynomial was known to fit the relationship between the input and the output parameters, therefore zero order Sugeno model i.e. Mamdani model was considered in the ANFIS training.

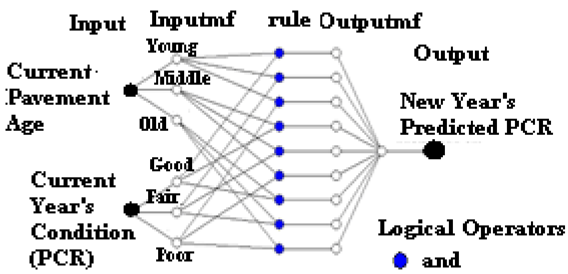


Fig.4 ANFIS Structure for MODEL-I

In MODEL-II, where subtractive clustering method of pattern recognition was followed, radius of influence was chosen as 0.304. This resulted in evolving 5 membership functions and 5 fuzzy rules (see Fig.5). This method uses first order (linear) Sugeno model to obtain the output. The model was trained for 100 epochs using back propagation learning method. The estimated training and checking RMSEs for both MODEL-I and MODEL-II reveals that the latter (checking RMSE = 4.95) was marginally better than the former one (checking RMSE = 5.25).

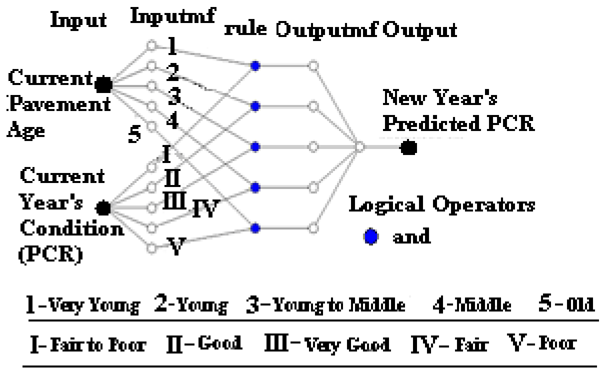
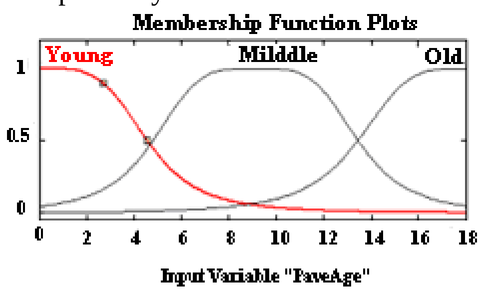
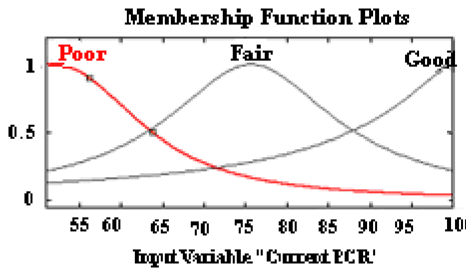


Fig.5 ANFIS Structure for MODEL-II

The optimized membership functions for the pavement age and current year's PCR generated for both MODEL-I and MODEL-II are presented in Fig.6 and Fig.7 respectively.



(a) Membership function for Pavement Age



(b) Membership function for Current year's PCR

Fig.6 Gaussian bell membership functions evolved using grid partitioning method of pattern recognition (MODEL-I) and after 100 epochs of training

3.3 Markov Process of Prediction

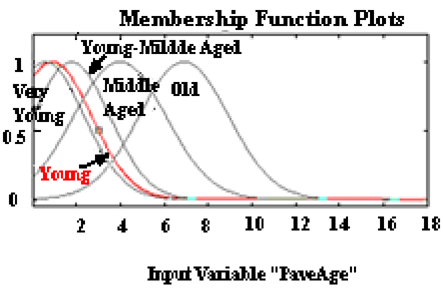
Following the PCR based classification of pavement as depicted in Fig. 1 and using the 1985-1986 PCR data Markov TPM (Table 1) was generated. These transition probabilities were then used to predict the condition of pavement for subsequent 4 years, i.e. 1987 – 1990, considering 1986 as the initial year.

Table 1 Markov Transition Probability Matrix generated from 1985-1986 PCR data

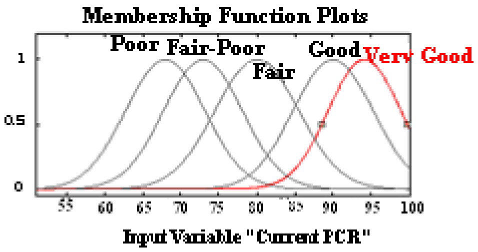
Condition	Good	Fair	Poor
Good	0.79	0.18	0.03
Fair	0	0.94	0.06
Poor	0	0	1

3.4 Model Efficiency

Although both the ANFIS based models were checked using randomly chosen 25% of the data from 1991-1996 data set, but the efficacy of these models and the Markov chain model were checked by predicting pavement condition for 1987-1990 considering 1986 as the initial year. The PCR output of the ANFIS based models were first translated to condition category following Fig.1 and then pavement condition distribution, i.e. percentage of "good", "fair" and "poor" for each year was estimated. The accuracy of this predicted condition distribution percentage is checked with the actual percentage distribution and overall RMSE was calculated.



(a) Membership function for Pavement Age



(b) Membership function for Current year's PCR

Fig.7 Gaussian bell membership functions evolved using subtractive clustering method of pattern recognition (MODEL-II) and after 100 epochs of training

Using the Markov TPM the pavement condition distribution percentage for the period 1987-1990 was estimated by multiplying each year's distribution with the TPM and its accuracy in prediction was also checked with the actual. The RMSE values obtained for MODEL-I, MODEL-II and the Markov chain model were compared to judge the efficiency of the pavement performance model.

All these analyses were carried out in MS Excel. A macro using Visual Basic was developed to invoke the FIS matrices for MODEL-I and MODEL-II from Excel. Using the 1986 data as the initial condition data, "succeeding PCR" was generated for the desired time period i.e.1987-1990. This "succeeding PCR" was then translated to pavement condition category following the adopted PCR classification (Fig.1).

3.5 Results

The actual pavement condition distribution and the distribution estimated from the PCR values predicted by MODEL-I, MODEL-II and Markov chain model are presented in Table 2 to Table 5.

Table 2 Actual Pavement Condition Distribution (in percentage)

Year / Condition	1986	1987	1988	1989	1990
Good	91.5	66.0	86.4	78.7	45.5
Fair	8.5	34.0	13.6	21.3	54.5
Poor	0.0	0.0	0.0	0.0	0.0

Table 3 Predicted Pavement Condition Distribution (in percentage) using Grid Partitioning based pattern recognition and BP learned based model (MODEL-I)

Year / Condition	1986	1987	1988	1989	1990
Good	91.5	83.0	83.0	55.3	21.3
Fair	8.5	17.0	17.0	44.7	78.7
Poor	0.0	0.0	0.0	0.0	0.0

Table 4 Predicted Pavement Condition Distribution (in percentage) using Subtractive Clustering based pattern recognition and BP learning based model (MODEL-II)

Year / Condition	1986	1987	1988	1989	1990
Good	91.5	87.2	72.3	59.6	10.6
Fair	8.5	12.8	27.7	40.4	76.6
Poor	0.0	0.0	0.0	0.0	12.8

Table 5 Predicted Pavement Condition Distribution (in percentage) using Markov chain model

Year / Condition	1986	1987	1988	1989	1990
Good	91.5	72.4	57.3	45.3	35.8
Fair	8.5	24.7	36.5	44.8	50.5
Poor	0.0	2.9	6.3	9.9	13.7

The RMSE values estimated using the data in Table 2 to Table 5 are presented in Table 6.

It was seen earlier that subtractive clustering and BP learning based model, i.e. MODEL-I (checking RMSE – 4.95) was marginally better during training and checking than the grid partitioning and BP learning based model, i.e. MODEL-II (checking RMSE – 5.25). However, prediction of condition for 1987-1990 reveals that MODEL-I outperforms both MODEL-II and Markov chain prediction model.

Table 6 Comparison of RMSE calculated from the different predicted and actual condition distribution for the period 1987-1990

Model	RMSE
MODEL-I	16.1
MODEL-II	18.8
Markov chain model	18.1

4 Conclusion

The project was done with an objective to use soft computing technique in building pavement performance deterioration model and evaluate its utility in pavement management. Results indicate that ANFIS can be used as better approach in building pavement deterioration model as it captures many uncertainties that even probabilistic approach used in Markov chain process cannot identify and take care-of. It is important to note that the model should be build considering the homogeneity of data set, i.e. the data should be collected from the same population. In other words, different model should be developed and applied for different type of road network, i.e. Interstates/US routes and state routes and for different type of pavement sections. However, the rules evolved from both the grid partitioning and subtractive clustering based pattern recognition followed by BP learning was bit difficult to discern and evaluate them using our knowledge of pavement deterioration rules. Research studies may be carried on further to work on this demerit and make this neuro fuzzy technique of building a pavement performance deterioration model overwhelming.

References:

[1]. G. W. Flintsch and Chen Chen, Soft Computing Applications in Infrastructure Management, *Journal of Infrastructure Systems*, ASCE, Vol.10, No.4, 2004, pp 157-166.

[2]. Y. Kleiner, R. Sadiq, and B. Rajani, Modeling the deterioration of buried infrastructure as fuzzy Markov process, *Journal of* Water *Supply Research and Technology: Aqua*, v. 55, No. 2, March 2006, pp 67-80.

[3]. A.K. Sandra, V.R. Vinayaka Rao, K. S. Raju, and A. K. Sarkar, Prioritization of Pavement Stretches using Fuzzy MCDM Approach – A Case Study, *11th Online World Conference on Soft Computing in Industrial Applications*, September 18-October 6, 2006, http://www.cs.armstrong.edu/wsc11/accepted.html

[4]. W-D Yu, Hybrid Soft Computing Approach for Mining of Complex Construction Databases, *Journal of Computing in Civil Engineering*, ASCE, Vol.21, No.5, 2007, pp 343-352.

[5]. Ohio Department of Transportation, *Pavement Condition Rating Manual*, http://www.dot.state.oh.us/pavement/PavementManagement/Manual.htm

[6]. S. Sundin, and C.B-Ledoux, Artificial Intelligence – Based Decision Support Technologies in Pavement Management, *Computer-Aided Civil and Infrastructure Engineering*, 16, 2001, pp 143-157.

[7]. J.S.R. Jang, C.-T. Sun and E. Mizutani, *Neuro – Fuzzy and Soft Computing: A Computational Approach to Learning and Machine Intelligence*, Prentice Hall, Upper Saddle River, NJ 07458, 1997.

[8]. H. A. Taha, *Operation Research: An Introduction, Seventh Edition*, Prentice Hall, Upper Saddle River, NJ 07458, 2002.

[9]. J. J. Ortiz-Garcia, S. B. Costello and M.S. Snaith, Derivation of Transition Probability Matrices for Pavement Deterioration Modeling, *Journal of Transportation Engineering*, ASCE, Vol.132, No.2, 2006, pp 141-161.