Technical Paper

ISSN 1997-1400 Int. J. Pavement Res. Technol. 6(2):130-135

Application of Factor Analysis in Maintenance Management of Low Volume Roads

V. Sunitha[1+], A. Veeraragavan[2], Karthik K. Srinivasan[2], and Samson Mathew[1]

Abstract: The maintenance management of low volume rural roads in developing countries presents a variety of challenges to road designers and managers. Rural roads comprise over 85% of the road network in India, and their serviceable condition is crucial to the rural economy. The present study aims to develop an index for rating the condition of pavements of low volume rural roads under the national rural road program, namely Pradhan Mantri Gram Sadak Yojana (PMGSY) in a typical district in India, so that the index can be used as a decision support tool for the maintenance management of low volume rural roads. A visual condition survey of the selected roads included evaluation of the conditions for shoulders, drains, cross drainage structures, camber and pavement distresses viz., potholes, crack areas, and edge breaks at every 200 m section. The data were collected for a period of three years. A multivariate technique, factor analysis, was used to calculate the index number for every 200 m pavement section in the study area. Data on 12 indicators of pavement conditions were factor analyzed (principal components and varimax rotation) and four factors were extracted, namely distress factor, side drainage factor, shoulder vegetation factor, and cross drainage factor. A total variance of 77.8% was explained by these factors. The index numbers were standardized as percentages for easy comparison across all sections and named as Visual Condition Index (VCI).

DOI: 10.6135/ijprt.org.tw/2013.6(2).130

Key words: *Factor analysis; Pavement management system; Rural roads; Visual condition index.*

Introduction

Over 2.65 million km of the Indian road network consists of rural roads. The fund recommended by the 12th Finance Commission for maintenance of roads and bridges amounts to over 2.262 billion USD for the 2009-10 year, and is expected to increase in the coming years. An intensive rural road development program, Pradhan Mantri Gram Sadak Yojana (PMGSY) (www.pmgsy.nic.in), was launched in 2000 with the objective of providing all-weather roads to all villages with populations over 500 by the year 2007. The scientific approach was introduced in the design of rural roads in India. The program envisages connectivity to 66,802 habitations with all-weather roads, construction of 146,185 km of a new rural road network, an improvement on the existing 194,132 km rural road network with an investment of 9.403 billion USD over four years, while also ensuring quality and transparency in the program implementation. As the building of road network is being carried out at a fast pace, focus is now on the maintenance and preservation of these roads.

Quantification of Pavement Distresses

To develop an efficient pavement management system, the first step is to build a periodic database of pavement condition. Several types of condition data, such as roughness, skid resistance, and distress are routinely collected during network level pavement condition evaluation [1]. The different variables indicate various aspects of pavement condition. For example, the longitudinal and transverse surface profiles are measured using the road surface monitoring vehicle, based on laser technology. The different surface distresses are monitored visually, or by using automated image collection and interpretation techniques. The properties of the pavement layers and the subgrade can be deduced from surface deflections measured using Falling Weight Deflectometer (FWD) and the Ground Penetrating Radar (GPR) [2]. The degree of pavement damage is dictated by the particular distress type as well as its severity and extent. Accordingly, most affluent highway agencies use specific deduct value curves to quantitatively rate the manifestation of each distress type based on the accurate assessment of its severity and extent. Although detailed distress evaluation procedures facilitate appropriate project ranking for maintenance and rehabilitation and future condition prediction, collection of distress data itself would be painstaking and costly. Therefore, a vital need exists for rapid, cost effective, and reliable methods of pavement condition evaluation. This is especially applicable for agencies possessing a limited budget allocation for pavement management. For strategic-level decision making in pavement management, the condition of the network has to be described using indicators that summarize the vast information obtained from the measurements.

The most famous and widely used indices include the Pavement Condition Index (PCI) and the Present Serviceability Index (PSI). Pavement in excellent condition is considered to have a PCI value of 100; the severity and extent of various defects reduce the PCI to a minimum of 0 [3-5]. The PSI reflects the user rating of ride quality on a scale ranging from 5 (excellent) to 0 (poor), and it is calculated from the measured condition variables using regression equations [6]. An index, called Unified Pavement Distress Index (UPDI), is defined and has been used to measure the pavement distress condition [7]. The problem with many existing practices for summarizing condition index values is that they require identical

[1] Department of Civil Engineering, National Institute of Technology, Tiruchirappalli 620015, India.
[2] Department of Civil Engineering, Indian Institute of Technology Madras, Chennai 600036, India.
[+] Corresponding Author: E-mail sunitha@nitt.edu
Note: Submitted April 23, 2012; Revised September 20, 2012; Accepted October 3, 2011.

information from all roads in consideration. In Brazil, factorial programming is performed, and the variance analysis of pre-selected factors quantifies the effects of those factors and their interactions on the pavement condition (i.e., dependent variables: roughness and pavement deflection) [8].

Special Considerations in Rural Road Maintenance

The distresses and the condition of low volume roads are different from other roads, for example National Highway and State Highways. Usually in PCI and PSI calculations, only the distress pattern and the surface profile details are considered. However, the implication of a drainage factor is very important in any pavement condition. In this paper, apart from the distress factors in the evaluation of the performance of low volume roads, the drainage factor is also considered indirectly, viz., shoulder vegetation, condition of side, and cross drainage structures, etc. As these roads carry low traffic, sophisticated equipment is not required to perform a pavement condition assessment survey. The visual condition survey was conducted on all the selected roads for three years. The objective of this paper is to form an index summarizing the overall condition of the low volume rural road network so that appropriate maintenance decisions can be made. To achieve the above objective, a multivariate technique of factor analysis is used [9]. The scope of this work is limited to the network of low volume roads constructed under the PMGSY scheme in India.

Methodology Adopted

The methodology adopted is as follows:

Step 1: Identification of the network of low volume rural roads.

The road sections considered in the present study are those roads that were constructed under PMGSY.

Step 2: Data Collection.

A set of data depicting the pavement condition and other relevant information are collected for every 200 m of each road section for a period of three years. The detailed description of data collected during the visual condition survey is explained later.

Step 3: Factor analysis (principal components and varimax rotation) of the data.

Step 4: Identification of factors which represent the pavement condition.

Step 5: Computation of the weighted sum of each pavement section's factor scores.

Step 6: Calculation of overall index number for the pavement sections.

Since the objective was to summarize, rather than identify the underlying factors, the principal component analysis is used. Varimax rotation is used in the fine tuning process to develop pure factors. For clarification, factor is defined as the linear combination of the original variables. The other relevant definitions are as follows [10]:

1. Factor: In addition to the above definition, factors represent the underlying dimensions that summarize or account for the variances of the original set of observed variables.
2. Factor matrix: A table displaying the factor loadings of all the variables on each of the factors.
3. Factor loadings: The correlation between the original variables and the factors is the key to understand the nature of a particular factor. The squared factor loading indicates the percent of variance in an original variable and is explained by a factor.
4. Communality: The amount of variance an original variable shares with all other variables included in the analysis.
5. Factor scores: The score of a given factor is the linear combination of all the input variables, weighted by the corresponding factor loading.
6. Eigen values: The eigen value for a given factor measures the variance in all the variables which is accounted by that factor. The ratio of eigen values is the ratio of explanatory importance of the factors with respect to the variables. If a factor has a low eigen value, then its contribution to the explanation of variances in the variables is not significant and may be ignored as redundant with more important factors.

Method of Factoring

The Principal Component Analysis type of factoring is used in this work. This type is used in exploratory research, in which the researcher's goal is to reduce the number of variables to minimize the number of factors when the causal models are not known.

Deciding the Number of Factors

The number of factors is decided based on the eigen values. Only the factors having eigen values greater than one are considered.

Rotation in Factor Analysis

The varimax method of rotation of factors has been adopted for the study. Varimax is indubitably the most popular rotation method. For varimax, a simple solution means that each factor has a small number of large loadings and a large number of zero (or small) loadings. This simplifies the interpretation because, after a varimax rotation, each original variable tends to be associated with one or a small number of the factors, and each factor represents only a small number of variables. In addition, the factors can often be interpreted from the opposition of a few variables with positive loadings to a few variables with negative loadings [11].

Study Area and Data Collection

The present study is concentrated on the Tiruchirappalli District in Tamil Nadu State in India, which extends over an area of 4,404 km^2 with a population density of 549 per km^2. This district is situated in the geographical co-ordinates of North Latitude between 10^o and 11^{o}30' and East Longitude between 77^{o}45' and 78^{o}50'. The study roads are low volume rural roads constructed under the PMGSY program. The age of the pavement varies from 4 to 10 years. A total of 124 study roads were considered, withlengths ranging from 0.5 to 4 km.

Pavement condition data is a pre-requisite for developing a pavement management system. An extensive visual condition survey of the study roads was conducted over three years (2009, 2010, and 2011). The pavement sections considered in the study were each of 200 m length. The different pavement distress data collected during the visual condition survey are shown in Table 1. The brief descriptions of the above parameters are given in the subsequent sections.

Pothole

Failure of wearing course observed may be due to the lack of proper mix design. Improper gradation of aggregates, inadequate binder contents, or an inferior binder results in a poor bituminous surfacing, which in turn leads to the development of potholes. A rating has been developed based on the number of potholes of 10 cm x 10 cm size.

Camber

Camber is the transverse slope provided to the road surface to drain off the rain water from the road surface. Water can cause the bitumen to lose its adhesive property, resulting in stripping and other moisture induced damages. Proper drainage and quick disposal of water from the pavement surface is important to ensure longevity of the pavement. The camber should be optimal.

Crack

Cracks can be of various types such as alligator cracks, longitudinal cracks, block cracking, etc. The rating provided was based on the percentage of cracked area on the total paved surface.

Edge break

Edge breaks are a serious issue in addressing pavement condition rating. There is a high probability of edge breaks progressing to the center of the pavement. The rating is based on the total cumulative length of the edge break in each pavement section of 200 m.

Shoulder Condition

Shoulders are provided to serve as an emergency lane for a vehicle compelled to leave the pavement or roadway. The transverse slope of the shoulder is an important factor since efficient drainage depends on the cross slope. Also, the presence of shrubs and vegetation affects the sight distance in these roads. Therefore, the rating for the shoulder is based on its condition, as well as the vegetation present over the section.

Table 1. Pavement Distresses and Their Indices Used.

Sl. No.	Indices	Description
1.	LDS	Left Drainage Shape
2.	LDI	Left Drainage - Silt and Debris
3.	LSV	Left Shoulder Condition - Vegetation and Slope
4.	POT	Pothole
5.	CAM	Camber
6.	CRA	Crack
7.	EDG	Edge Break
8.	RSV	Right Shoulder Condition - Vegetation and Slope
9.	RDS	Right Drainage Shape
10.	RDI	Right Drainage - Silt and Debris
11.	CDO	Cross Drainage Structure - Opening
12.	CDS	Cross Drainage Structure - Settlement

Table 2. Rating for Potholes, Camber, Cracked Area and Edge Break.

Rating	Potholes Number	Camber Provided	Cracked Area of Paved Surface	Length of Edge Break, m
5	Nil	Very Good	Nil	Nil
4	1 - 3	Good	1 - 5%	1 - 10
3	4, 5	Fair	6% - 10%	11 - 20
2	6, 7	Poor	11% - 15%	21 - 30
1	>7	Very Poor	>15%	> 30

Drainage - Shape, Silt and Debris

The side drains are an essential cross section element of any pavement. The drainage often consists of an open ditch parallel to the road carriageway with culverts at regular intervals to disperse the run-off to local water courses. The rating of the side drains consists of two components: one is the shape and side slope of the drain, and the other factor is the presence of silt and debris in the side drain.

Cross Drainage Structure

Cross drainage structures such as box or pipe culverts are common elements in these road sections. The poor conditions of these structures have an adverse effect on the condition of road. These are rated based on the settlement and the erosion of the structures, along with the closure of the openings of these structures due to silts and other debris.

The rating adopted for the above pavement distresses and camber

Table 3. Rating for Shoulder Condition, Side Drain and Cross Drainage.

Rating	Shoulder Condition	Shape and Slope of Side Drain	Percentage Silts and Debris in Side Drain	Percentage of Opening in Cross Drainage	Percentage of Settlement in Cross Drainage
5	Very Good	Very Good	Nil	Fully open	Nil
4	Good	Good	1 - 25%	> 75%	1 - 25%
3	Fair	Fair	26% - 50%	51% - 75%	26% - 50%
2	Poor	Poor	51%- 75%	26% - 50%	51% - 75%
1	Very Poor	Very Poor	> 75%	< 26%	> 75%

Table 4. Rotated Factor Loadings.

S1. No	Index	Factors				Communality
		I	II	III	IV	
1	LDS	0.156	**0.65**	0.511	0.013	0.709
2	LDI	0.194	**0.902**	0.025	0.115	0.866
3	LSV	0.178	0.119	**0.896**	0.012	0.848
4	POT	**0.821**	0.128	0.229	0.07	0.748
5	CAM	**0.764**	0.082	0.115	0.034	0.605
6	CRA	**0.821**	0.284	0.012	0.07	0.76
7	EDG	**0.811**	0.157	0.208	0.062	0.729
8	RSV	0.197	0.135	**0.867**	0.043	0.811
9	RDS	0.179	**0.694**	0.457	0.038	0.724
10	RDI	0.189	**0.905**	0.032	0.117	0.869
11	CDO	0.071	0.105	0.015	**0.902**	0.83
12	CDS	0.082	0.08	0.038	**0.907**	0.838
Eigen Value		2.808	2.715	2.137	1.682	
Percentage of Variance		23.351	22.627	17.808	14.02	
Total Variance Explained				77.806		

Table 5. List of Indicators That Loaded High on Each Factor.

Factor	Index	Description
Factor I (Distress Factor)	POT	Pothole
	CAM	Camber
	CRA	Crack
	EDG	Edge Break
Factor II (Side Drainage Factor)	LDS	Left Drainage Shape
	LDI	Left Drainage - Silt and Debris
	RDS	Right Drainage Shape
	RDI	Right Drainage - Silt and Debris
Factor III (Shoulder Vegetation Factor)	LSV	Left Shoulder Condition – Vegetation and Slope
	RSV	Right Shoulder Condition – Vegetation and Slope
Factor IV (Cross Drainage Factor)	CDO	Cross Drainage Structure – Opening
	CDS	Cross Drainage Structure – Settlement

is given in Table 2. Table 3 depicts the rating for the shoulder condition, shape and slope of side drainage, percentage of silts and debris in side drain, and the percentage of opening and settlement in cross drainage structures.

Factor Analysis Results

From the above visual survey data, 550 sections of 200 m each were selected for the factor analysis. The data collected over three years was used in the analysis. Factor analysis was carried out using Statistical Package for the Social Sciences (SPSS) software. The factor analysis results identified four factors which associate twelve measures of pavement condition presented in Table 1. The varimax rotation option, which minimizes the number of variables that load highly on a factor, uses the orthogonal assumption. Orthogonal rotation assumes that the factors are at right angles to each other, i.e., the factors are not correlated. Only the factors with eigen values greater than one are considered. The factor loadings (correlation between the original variables and the factors) are shown in Table 4. For clarity, the highest factor loading for each indicator across all factors is shown in bold letters. Table 4 also contains eigen values, percentage of variance, communalities, and total variance. When the eigen values (column sum of squared factor loadings) are multiplied by the number of indicators, 12, the percentage of variance explained by that factor is obtained. In Table 4, if the eigen value for Factor I (Distress factor) 2.808 is multiplied by 12, the result is 23.351. This is the percentage of variance explained by Factor I. The percentage of variance explained ranges from as high as 23.351 for Factor I (distress factor) to as low as 14.02 for Factor IV (cross drainage factor). Communality, as given in last column of Table 4, is the proportion of each variable's variance that can be explained by the factors. It is defined as the sum of squared factor loadings for the variables.

The original objective was to explain as much of the total variance as possible and stay within an acceptable factor cut off point. It is important to note that principal components factor analysis considers both specific and common variance of indicators. Table 5 contains the list of indicators that loaded high on each factor. Although it is beyond the scope of this paper to explain and discuss each individual relationship, it is necessary to carefully study the associated groups or factors. After considering the list of indicators loading high on each factor, the authors made an effort to summarize their nature with a descriptive term. The selection of this term is an attempt to collectively name the group with the highest loading indicators having major influence on that decision. The following explanation adequately describes the general characteristics of the four factors extracted:

Factor I ($f = 1$): Distress factor.

Factor II ($f = 2$): Side drainage factor.

Factor III ($f = 3$): Shoulder vegetation factor.

Factor IV ($f = 4$): Cross drainage factor.

The factor score for all the pavement sections was computed using SPSS software based on the "regression" option. The score of each factor is a linear combination of all the measures, weighted by

Table 6. Factor Scores for Individual Pavement Sections.

Section No.	Factors				Initial Index	Standardized Index Visual Condition Index (VCI)
	I	II	III	IV		
1	-0.488	0.269	-0.796	0.497	-12.493	65
2	0.492	0.395	-1.205	0.407	4.674	73
3	0.703	-0.575	-2.754	0.497	-38.687	54
4	-0.374	-0.423	-2.576	0.566	-56.253	46
5	1.038	0.464	0.941	0.384	56.864	95
6	-0.761	0.754	-1.577	0.471	-22.169	61
7	-0.921	0.404	0.511	-1.008	-17.415	63
8	0.261	1.046	-0.48	0.379	26.528	82
9	-0.293	1.058	0.037	0.423	23.695	81
10	-0.172	1.297	-1.006	0.394	12.937	76
Percentage of Variance	23.351	22.627	17.808	14.02		

Table 7. Classification of VCI and No. of Sections in Each Category.

Classification	Visual Condition Index (VCI) Range	No. of Sections		
		2009	2010	2011
Very Good	80 - 100	268	210	62
Good	60 - 79	228	249	234
Fair	40 - 59	50	84	215
Poor	20 - 39	2	4	35
Very Poor	0 - 19	2	3	4

the corresponding factor loading. The authors followed the method used in formulating health index in the literature [10]. An index has been formulated for all sections based on the factor score of the sections and the variance of the factors as given in Eq. (1).

$$\mathrm{Index}_i = \sum_{f=1}^{4} S_{fi} V_f \tag{1}$$

where,

Index_i = index of the i^{th} pavement section

f = factors (1 to 4)

S_{fi} = factor score for the factor f for the i^{th} section

V_f = percentage variance of the factor f

As a sample, factor scores for ten sections are presented in Table 6. The percentage variance was used to weigh each factor score across all the factors. As an illustration, the index values for section 1 (-12.493) were obtained as given in Eq. (2):

$$-0.488*23.351+0.269*22.627+(-0.796)*17.808 \\ +0.497*14.02=-12.493 \tag{2}$$

These values are from Table 6, where 23.351 is the percentage of variance for Factor I and it is multiplied by the score for Factor I, which is -0.488. This product is added to the percent variance, 22.627 and multiplied by a factor score of 0.269 and so on. This technique of multiplication and summation is followed across all four factors for section 1. The initial index numbers were all determined by the same procedure. Since it would be difficult to interpret or compare such figures, they were standardized and scaled based on Eq. (3) to a percentage of 0 to 100 (0 being the worst and 100 the best among the sections in three years).

$$Standardized\ index = 100*(initial\ index - initial\ index_{min}) \\ /(initial\ index_{max} - initial\ index_{min}) \tag{3}$$

where *initial index*$_{max}$ and *initial index*$_{min}$ are the maximum and the minimum values of the initial index, respectively.

The standardized index was named Visual Condition Index (VCI) by the authors. As it is difficult to present the VCI of all the sections, the VCI of ten sample sections is presented in the last column of Table 6. Likewise, the standardized index was calculated for all 550 sections for the period from 2009 to 2011.

Classification of Pavement Sections Based on Visual Condition Index

For additional interpretation, the sections were ranked according to the VCI. The pavement section with a VCI of 100 ranks at the top, and the section with a VCI of 0 ranks last. As it is difficult to present the VCI of all the sections on a yearly basis, the sections are categorized according to the range of VCI as VERY GOOD, GOOD, FAIR, POOR, and VERY POOR, and the number of sections falling in each category in the years 2009, 2010 and 2011 are specified in Table 7.

Summary and Conclusions

Four factors were identified after factor analysis of the survey data with 12 indicators. The four factors that were identified are: factor I - distress factor, factor II - side drainage factor, factor III - shoulder vegetation factor, and Factor IV - Cross drainage factor. The percentage of variance explained by the different factors are 23.351% for distress factor, 22.627% for side drainage factor, 17.808% for shoulder vegetation factor, and 14.02% cross drainage factor. The percentage of variance explained by the factors is high for factor 1 (distress factor) and is low for factor IV (cross drainage factor). Thus, the first extracted factors are more strongly associated and account for the highest percentage of variance explained. The total variance of 77.8% was explained by the factor analysis. The pavement sections can be ranked according to the standardized index called VCI. The pavement section with a VCI of 100 ranks at the top, and the section with a VCI equal to 0 ranks last. The VCI can be used for further processes in the Pavement Management System, viz., deterioration modeling and optimization of the maintenance strategies.

References

1. Bandara, N. and Gunaratne, M. (2001). Current and Future Pavement Maintenance Prioritization Based on Rapid Visual Condition Evaluation, *Journal of Transportation Engineering*, 127(2), pp. 116 -123.
2. Ruotoistenmaki, A. (2005). Pavement Condition Rating Based on Factor Analysis of Pavement Condition Measurements. *Presented at 84th Annual Transportation Research Board (TRB) Meeting*, Washington DC, USA.
3. Haas, R., Hudson, W.R., and Zaniewski, J.P. (1994). *Modern Pavement Management.* Krieger Publishing Company, Malabar, Florida, USA.
4. Shahin, M.Y. and Walter, J.A. (1990). Pavement Maintenance Management for Roads and Streets Using the PAVER System, *USACERL Technical Report M-90/05*, Champaign, IL, USA.
5. Sun, L., Kan, S., and Ji, Y. (2004). Evaluation and Rehabilitation Models of Airport Pavement Management System: A Case Study of Shanghai Pudong International Airport, *6th International Conference on Managing Pavements,* Brisbane, Australia.
6. AASHO (1962). *The AASHO Road Test, Report 5 – Pavement Research, Special Report 61-E*, Highway Research Board, National Research Council, Washington, DC, USA.
7. Jung, C.H. and Amirkhanian, S.N. (1992). Unified Pavement Distress Index for Managing Flexible Pavements. *Journal of Transportation Engineering,* 118(5), pp. 686-699.
8. Fernandes, Jr. J.L. and Yshiba, J.K. (2006). Pavement Performance Prediction Models Based on a Factorial Experiment in the State of Parana – Brazil. *Presented at 85th Annual Meeting of the Transportation Research Board*, Washington, DC, USA.
9. Hair, J., Anderson, R., Tatham, R., and Black, W. (1998). *Multivariate Data Analysis.* Prentice-Hall, Englewood Cliffs, New Jersey, USA.
10. Hightower, W.L. (1978). Development of an Index of Health Utilizing Factor Analysis. *Medical Care*, 16(3), pp. 245 - 255.
11. Abdi, H. (2003). Factor Rotations in Factor Analyses, (unpublished manuscript), University of Texas at Dallas, Program in Cognition and Neurosciences, Dallas, TX, USA.