

Mapping of Global Road Systems Based on Statistical Discriminant Analysis

2010-01-0924

Published
04/12/2010

Ted T. Stawiarski and Joseph Wolkan
General Motors Company

Mohamed E. M. El-Sayed
Kettering University

Copyright © 2010 SAE International

ABSTRACT

Automotive manufacturers are facing continuously changing Global environment. Traditionally, these manufacturers relied on structural and general durability tests to validate vehicles. For these tests to remain representative of customer usage in a Global environment, the overall surface conditions of the Global road systems must be studied. Understanding and classifying these road systems conditions is an important step in dealing with vehicle durability in the Global environment.

In this paper, an approach to mapping the world road systems into Established Roads (ER) and Developing Roads (DR), utilizing Statistical Discriminant Analysis (SDA), is presented. The classification of Global regions as DR and ER road systems can be effectively used to recommend appropriate development and validation tests for each road system. A few examples are presented to demonstrate how the ER vs. DR mapping and consideration of country specific conditions can be used in selecting the appropriate vehicle content which would satisfy requirements and meet durability targets.

INTRODUCTION

In recent decade, a number of countries have undergone dramatic and accelerated economic development. Between other factors, the road surface quality has improved in many countries such as China, Mexico, Singapore, Poland, Check Republic and others. Since auto manufacturers produce vehicles that are designed for specific road conditions, the following key questions arise: is this development and improvements sufficient to ship vehicles built for established road systems? Or these vehicles need structures and

components capable of sustaining high loads and large cycles typical for developing road systems?

To answer these questions, in a practical approach, the world countries and their road systems are divided into two categories as countries with good or Established Roads (ER), and countries with Developing Roads (DR). This is done with the intent to recommend the appropriate ER or DR tests to be used as targets for vehicle design, development, and validation. This ER and DR systems classification illustrates the need for different general and structural durability tests representative and appropriate for various global customer usages. The detailed analysis supporting such a classification is being presented and discussed in the following paper.

ROAD SURFACE DATA

To decide if a particular region or state has ER or DR road system, it is essential to identify what kind of data and variables would be useful in conducting the analysis. Searches to collect the relevant data included public sources, and surveys.

PUBLIC DATA

Various sources of data are available in the public domain. Data of 210 countries were obtained and analyzed regarding number of vehicles and roads condition.

The search for data regarding road surface quality started with the State of Michigan Transportation Department where engineers suggested following data sources [1,2,3,4,5,6,7,8,9,10,11,12,13]:

- Michigan Department of Transportation (MDOT) Database: State highways information



Documents
Document Sharing Hub

- Research and Innovative Technology Administration (RITA): National Transportation Library
- Transportation Research Information Services (TRIS)
- International Transport Research Documentation (ITRD) database
- University of Michigan Library: Government Documents Center

Similar data sources have been searched on the national level:

- National Technical Information Service (NTIS) database
- U.S. Census Bureau : The 2009 Statistical Abstract
- US Department of Transportation: Highway Statistics 2001
- US Department of State: National Highway System Length Miles by Measured Pavement Roughness

The most effective data sources have been:

- World Bank: World Development Indicators
- International Road Federation (IRF)
- European Union Road Federation (ERF)
- Central Intelligence Agency: The World Fact Book

Since the late 1970's Tom Gillespie, Mike Sayers, Steve Karamihas, and other researchers at University of Michigan Transportation Research Institute (UMTRI) have investigated various methods of measuring and interpreting road roughness [14,15,16].

The international roughness index (IRI) was developed by the World Bank in the 1980s. IRI is used to define a characteristic of the longitudinal profile of a traveled wheel-track and constitutes a standardized roughness measurement. Profiling devices are used to provide accurate, scaled, and complete reproductions of the pavement profile within a certain range. These are available in several forms. Systems used today in the United States are frequently installed in vans which contain microcomputers and other data processing instrumentation. This method is very objective and therefore is very promising for future road systems assessments. The only disadvantage is the lack of comparisons between data measured in different countries worldwide.

SURVEY DATA

Some subjective evaluation has been obtained by conducting surveys of Global Regions regarding their road conditions. The evaluations were conducted by regional contacts, usually regional service engineers, whose work objective has been to address local requirements and service related issues. The evaluations were made using the same defined scale and the following criteria:

- Condition of roads connecting major cities (percentage of paved roads)
- Use of speed bumpers and frequency of potholes.
- City driving conditions and traffic congestion in the rush hours.
- Available technology and technical culture that helps to keep roads in good condition.

DISCRIMINANT ANALYSIS

Discriminate Analysis determines the multivariate distance of each member from the average value for each group, using all of the discriminating variables. The smaller of these 2 distances indicates to which group the “unknown” member belongs. The closer the unknown member is to the group's average (it's centroid), the higher the probability that it belongs to the group. The problem is multivariate when many variables are considered at the same time. The number of variables determines the dimensionality of the space of the problem. The multivariate distance (or Mahalanobis distance) is defined as [17-18]:

$$MD = (\bar{y} - \mu_0)' S^{-1} (\bar{y} - \mu_0) \quad (1)$$

Where:

$\bar{y}$ - Vector of the discriminating variables

μ_0 - Vector of the means of the discriminating variable for the ith group
 S - Common covariance matrix of the discriminating variables

To discriminate among 2 or more groups using a collection of discriminating variables, it is necessary to start with known members of the groups and to develop the discriminating functions. Once the critical set of discriminating variables becomes identified, the discriminating functions can be used to classify unknown observations. Therefore, it is important to identify with good confidence both known members of the classified groups of countries. For this problem, there were 19 countries which were considered to be ER and 23 countries which were considered to be DR. These known groups were used to develop a discriminating function that was used to classify all the remaining countries.

For this problem, a large set of possible predictors was used to develop the discriminating functions. These included variables associated with the numbers of vehicles operating in the countries as well as variables associated with the country's ability to maintain roads, the quantity of these roads, and various other associated variables. These variables were examined using a selection process in order to determine a small, yet powerful collection of variables classifying the countries as ER or DR.

Several models were tested in this analysis with different variables including: motor vehicles per thousand people (MVPTV), percent of roads that are paved, total roads network length per land area, total network per GDP, paved road length per thousand people, and others.

No assumptions have been made regarding the developing unclassified countries like Mexico, China, Middle East, Eastern Europe, and South-East Asia. No attempts were made to pre-judge the outcome. 210 countries have been identified, but there were much fewer countries with complete and reliable data. There were very few variables with good data for all 210 countries. Most of the variables were complete only for 112 countries.

The surveyed data were the most limited and suggested that a variable related to road maintenance might be beneficial if added to the examined set. In looking through the available variables from public data for a surrogate for road maintenance, gross domestic product per capita (GDPPC) seemed to be the best representative.

Consequently, a two-variable discriminate function, using MVPTP and GDPPC, was developed. This model has been found to be the most effective one in separating ER and DR road systems.

ANALYSIS RESULTS

The following [Tables 1,2,3,4,5](#) (please see the [Appendix](#)), present selected results of the Discriminant Analysis. In these tables, the first three columns list the countries and Discriminant Variables followed by the column indicating if the particular country belongs to one of the classified groups ER or DR or its classification was predicted (if it was left blank.) The following 4 columns describe ER and DR value of Discriminant Function and multivariate distance (MD.) The last two columns represent the predicted class and probability of ER.

[Table 1](#) lists all classified ER countries. Countries with strong economy like US, Canada, Great Britain, and Japan belong to this group. Please, note that Korea was found to be ER with probability 60% and South Africa has been predicted to be DR.

Similarly, [Table 2](#) presents results for classified DR countries. The results confirmed that all these countries were DR.

[Table 3](#) lists few countries that were predicted to be ER but initially were not classified. To this group belong countries like Iceland, Slovenia, Ireland, Finland, Kuwait, Greece, and Qatar.

The analysis has identified also a group of developing countries with strong ER characteristics. These countries are in Transition to ER road systems and their vehicles should be designed, developed, and validated to ER targets. [Table 4](#) presents countries predicted to be in transition to Established Road Systems. Note that probability of prediction for these countries is between 0 and 1. To this group belong countries like Czech Republic, Estonia, Poland, Hungary, Lebanon and Malaysia.

[Table 5](#) presents countries predicted with high confidence to be DR (Probability ER=0.) This group is led by Mexico with relatively high MVPTP of 211. China and Saudi Arabia also have been predicted to be DR. Other countries in this group are: India, Pakistan, Iran, Argentina, Russian Federation, Ukraine, and Turkey.

DISCUSSION

Using the ER/DR classifications, the variable of “Motor Vehicles per Thousand People” correctly classified 99.1% of the countries. Only South Africa was misclassified. It was classified as ER but found to be DR by calculation. The dividing point was about 250 MVPTP as shown In [Figure 1](#). None of the other models could outperform the MVPTP model. Even the model with all the variables didn't outperform the MVPTP.

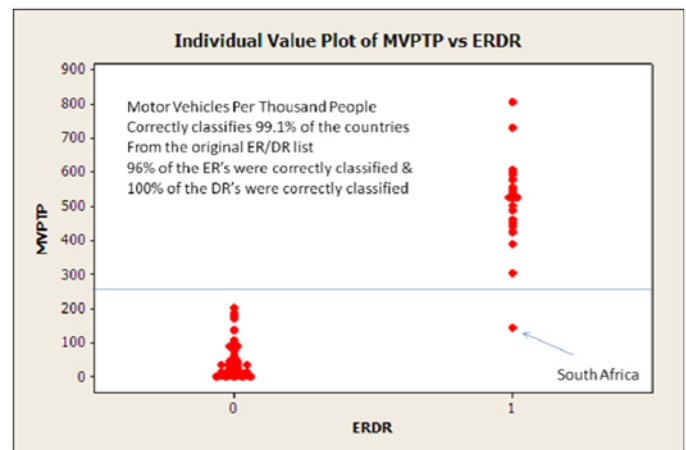


Fig 1. Analysis results with MVPTP model for classified groups of countries.

Since MVPTP ties to number of vehicles in a given country, the second variable was searched to reflect the road surface quality. Unfortunately, the road expenditure model did not improve the MVPTP model. Many other models were tested until it was found that the second critical variable that tied to road quality was gross domestic product per capita (GDPPC). [Figure 2](#), shows the scatter plot of classified countries using the MVPTP and GDPPC model.