Using support vector machines to predict the probability of pavement failure

Schlotjes, Megan; Burrow, Michael; Evdorides, Haralampos; Henning, Theuns

DOI:
10.1680/tran.12.00084

Using SVM to Predict the Probability of Pavement Failure

Megan R. Schlotjes, Michael P.N. Burrow, Harry T. Evdorides, and Theunis F.P. Henning

Corresponding Author:
Miss Megan R. Schlotjes
Joint Ph.D. Student – The University of Auckland and University of Birmingham
Department of Civil and Environmental Engineering, Faculty of Engineering, The University of Auckland
Private Bag 92019, Auckland Mail Centre
Auckland 1142
New Zealand

Telephone: +64 274280627
Email: msch102@aucklanduni.ac.nz

Other Authors:
Dr. Michael P.N. Burrow
University of Birmingham, Birmingham, United Kingdom
School of Civil Engineering, University of Birmingham
Edgbaston
Birmingham B15 2TT
United Kingdom
Email: m.p.n.burrow@bham.ac.uk

Dr. Harry T. Evdorides
University of Birmingham, Birmingham, United Kingdom
School of Civil Engineering, University of Birmingham
Edgbaston
Birmingham B15 2TT
United Kingdom
Email: h.evdorides@bham.ac.uk

Dr. Theunis F.P. Henning
The University of Auckland, Auckland, New Zealand
Department of Civil and Environmental Engineering, Faculty of Engineering, The University of Auckland
Private Bag 92019, Auckland Mail Centre
Auckland 1142
New Zealand
Email: t.henning@auckland.ac.nz

Word Count (incl. Tables): 5004
No. of Figures: 6
Date of Submission: 13th November 2012

ABSTRACT

This paper presents a method to predict the probability of structural failure of road pavements using information contained in road datasets. Expert knowledge was used to develop failure charts which identify the potential factors that may contribute towards pavement failure. A computational technique, known as a support vector machines, was built to use this information to determine from the datasets the probability of failure of individual road sections. With this prediction comes an indication of the predominant failure types, the causes of structural failure and the risk profile of a road network.

The usefulness of the approach was demonstrated on a dataset taken from the New Zealand Long Term Pavement Performance study of State Highways. The analysis of the dataset showed that the network was in a good condition, yet a small number of pavement sections with a high likelihood of failure were identified. Furthermore, the application of the failure paths examined the three predominant failure types occurring on the network and identified their possible causes. Rutting appears to be significantly influenced by the road pavement strength, fatigue cracking seems to be affected notably by the environment (i.e. water ingress) and shear failure is caused primarily by the combination of traffic, pavement composition and strength. In addition, it was confirmed that measured functional pavement condition alone is not a good identifier of failure, and that the inclusion of a parameter related to strength, such as pavement deflection, is essential.

KEYWORDS

Risk and Probability Analysis, Pavement Design, Maintenance and Inspection

NOTATIONS

SVM	Support vector machines
FMEA	Failure mode and effect analysis
FTA	Fault tree analysis
LTPP	Long term pavement performance programme
SH	State Highway, namely used as the research road network in this paper
2γ	Support vector machine margin calculated in the machine learning task, where γ = margin from the SVM decision boundary to the closest point, namely the support vectors
x_1, x_2	Closest points (vectors) to the decision boundary in the SVM model
w	Weight vector for the SVM model, defined by the research data (no units)
b	Bias for the SVM model, defined by the research data (no units)
$P_{FAILURE}$	Overall failure probability, predicted by the trained SVM model
$P(A), P(B), P(N)$	Predicted probability of failure for failure type A, failure type B, and failure type N, respectively
T_P	Predicted **T**rue **P**ositives (number)
T_N	Predicted **T**rue **N**egatives (number)
F_P	Predicted **F**alse Positives (number)
F_N	Predicted **F**alse **N**egatives (number)
N_{Total}	Total **N**umber of predictions
$P_{Predicted}$	Predicted probabilities
P_{Actual}	Actual failure probabilities, from the research dataset (binary output)
N_1	Total number of predicted non-failures

N_2 Total number of predicted failures

N_3 Total number of actual non-failures

N_4 Total number of actual failures

INTRODUCTION

Road asset managers with limited road maintenance budgets are faced with the challenging task of prioritising maintenance expenditure on road networks and thereby ensuring that the structural integrity of the network is preserved over time (Robinson *et al*., 1998). Once a failed road pavement has been identified, asset managers need to select the most appropriate maintenance treatment. However, without a comprehensive understanding of pavement failure, inappropriate maintenance is often carried out. At present, a combination of available data, such as traffic, road inventory and condition are used together with pavement deterioration models to estimate future network condition and to evaluate the maintenance requirements on a road network.

Predicting structural road pavement failure is a challenging task because of the complex interaction between the factors that contribute towards failure, the different modes or mechanisms by which a road may fail, the availability, quality and variability of data, and the inherent uncertainty of the behaviour of road pavements (Reigle, 2000). There are models that focus on singular or multiple types of failure (e.g. cracking or rutting) and there have been systems with diagnostic capabilities (Henning, 2008). The formulation of such models requires a thorough understanding of the complexities of pavement failure, which can in turn assist in selecting appropriate model variables (Isa *et al*., 2005). Whilst a number of researchers have developed approaches for infrastructure systems which utilise an understanding of failure types (Evdorides, 1994; Xiao *et al*., 2011), this practice is not widely used in the road sector arguably because of unavailability of both data of appropriate quantity and quality and computational techniques which are accessible for the practicing engineer.

To address this, this paper describes a computational methodology which quantifies the probability of structural failure of road pavement sections and identifies the most likely contributing factors. This is achieved using fault trees, developed using expert opinion, to identify the combination of factors which could contribute to failures. A computational technique, known as a support vector machines (SVM), automates the process by examining the possible failure paths in a given set of data associated with a road pavement to classify whether the scrutinised road pavement is sound or has failed, and to assign a probability of failure according to the potential failure paths identified from the fault tree analysis.

Accordingly, this paper presents the following:

1. The theoretical framework used for the diagnosis of the cause of failure and the probability of failure;
2. The development and testing of the methodology using data from New Zealand and
3. A discussion of the usefulness of the methodology developed.

PAVEMENT PERFORMANCE MODELLING

A number of approaches have been adopted to predict road pavement performance of which the probabilistic approach is becoming increasingly popular due to the stochastic nature of the variables measured on the road networks. This approach recognises that much of the data collected on road networks is highly variable (Martin, 2008). Methods used to this end include logistic regression, basic linear and non-linear models, Bayesian probabilities, genetic algorithms, and kernel-based learning methods (Henning, 2008; Martin, 2008; Park *et al*., 2008, Caruana and Niculescu-Mizil, 2006). In other fields such as medical diagnostics and other engineering disciplines, neural networks, SVM, fuzzy logic, and analytical hierarchy processes have been used successfully to calculate risk probabilities (Tu, 1996; Pal, 2006; Volinsky *et al*., 1997).

The success of a particular modelling technique depends greatly on the appropriateness of the model for the situation at hand and its performance can be enhanced by understanding the underlying causes of failure (Isa *et al*., 2005). Two such widely employed techniques are that of failure mode and effect analysis (FMEA) and fault tree analysis (FTA) (Patev *et al*., 2005, Seyed-Hosseini *et al*., 2006). The former is an analytical tool for reliability analysis which can be used to identify possible failure causes to minimise, or eliminate, failure in systems. By using a weighted ranking system, each failure is assigned a risk number that represents the overall impact of failure. The causes of failure can be graphically represented using FTA, which further enables concurrently occurring failure factors to be included in the modelling process (Patev *et al*., 2005). With this approach, the failure paths can be established from the breakdown of the critical paths.

THEORETICAL FRAMEWORK

Conceptual Design

In order to determine the probability of road pavement failure from road datasets, the approach adopted used expert knowledge to identify the predominant types of failure on a road network and the associated foremost factors which contribute towards failure. Subsequently, a computational technique was identified and developed to analyse road pavement datasets containing these factors. The developed technique was capable of determining the probability of failure for each of the failure types and of identifying the most probable combination of factors which contribute to its failure. The probability of failure for each failure type were considered together to determine the overall failure probability of a pavement section. The overall approach consisted, therefore, of two main parts:

1. Fault charts to diagnose the cause of failure

2. A computational model to calculate the probability of failure.

Fault charts

Fault, or failure, charts were built by canvassing the views of a panel of experts, in conjunction with a preliminary analysis of road networks. Initially the predominant failure types, or mechanisms occurring on the road network, were identified and for each, the panel of experts identified fundamental groups of factors which contribute to failure. These were then broken down further into associated sub-factors and used by the panel of experts to develop a fault chart for each failure type, which can be used to identify the underlying causes of failure and the interactions between factors associated with failure and the failure modes. Three such charts which focus on the predominant mechanisms associated with structural failure occurring on New Zealand road networks are shown in Figures 2 to 4.

Computational model

In order to determine the probability of failure of road pavements it was necessary to select an appropriate computational technique which could make use of the data corresponding to the types identified by the panel of experts. A number of methods were examined for this purpose and they included logistic regression, neural networks, SVM, probability trees and random forests (Caruana and Niculescu-Mizil, 2006; Chandra *et al*., 2009). Following an extensive sensitivity analysis using road data from the New Zealand Long Term Pavement Performance (LTPP) State Highway (SH) network, a SVM technique was chosen for the task in hand (Schlotjes *et al*., 2012). A SVM is a supervised computational learning model with an associated training algorithm which can be used, for a given set of input data, to assign a probability to two possible categories to which the set of input data may belong (Van Looy *et al*., 2007). The SVM training algorithm uses input training data to build a model that can assign probabilities to new input datasets. Herein the input datasets consisted of road network information corresponding to the data types as identified by the panel of experts.

The SVM technique transforms typically non-linear data, or data difficult to separate with steadfast decision boundaries, using various kernel functions. Once transformed, the data is easily separated such that an unambiguous decision boundary is defined, as shown in Figure 1 (Van Looy *et al.*, 2007). By maximising the margin (2γ) between the separated data classes, the optimal solution is found to ensure confidence around the new predictions. To do so, using vector mathematics of the closest data points to the decision boundary (namely the support vectors), the following equation is maximised:

Equation 1

$$
\begin{aligned}
2\gamma &= \frac{1}{\|\boldsymbol{w}\|}\boldsymbol{w}^T(\boldsymbol{x}_1 - \boldsymbol{x}_2) \\
&= \frac{1}{\|\boldsymbol{w}\|}(\boldsymbol{w}^T\boldsymbol{x}_1 - \boldsymbol{w}^T\boldsymbol{x}_2) \\
&= \frac{1}{\|\boldsymbol{w}\|}((\boldsymbol{w}^T\boldsymbol{x}_1 + b) - (\boldsymbol{w}^T\boldsymbol{x}_2 + b)) = \frac{1}{\|\boldsymbol{w}\|}(1+1) \\
\gamma &= \frac{1}{\|\boldsymbol{w}\|}
\end{aligned}
$$

The model gives for each failure type, and for each failure path, a probability of a pavement section failing. For a pavement section the most probable failure path, for a particular failure type, is that which has the greatest failure probability.

The overall failure probability ($P_{FAILURE}$) for a pavement section with N failure types was calculated by assuming that each failure mechanism acts independently as follows:

Equation 2

$$P_{FAILURE} = max\,[P(A), P(B), \ldots, P(N)]$$

It should be noted that a similar approach is adopted in conventional pavement design.

Assessing the Performance of the Model

Assessing the performance of the model is an integral part of developing any machine learning tool, such as the SVM model developed herein. Four tests were considered: accuracy, misclassification, the f-score and phi coefficient (Parker, 2011):

The accuracy and misclassification tests were used to determine the number of incorrectly predicted road sections, and compare the predicted output with the actual failure data. The accuracy and misclassification percentage were calculated as follows (Parker, 2011):

Equation 3

$$Accuracy = \frac{\sum T_P + T_N}{N_{Total}} \times 100\ \%$$

Equation 4

$$Misclassification = \frac{\sum |P_{Predicted} - P_{Actual}|}{N_{Total}} \times 100\ \% = \frac{\sum F_P + F_N}{N_{Total}} \times 100\ \%$$

The f-score is a weighted average of, the fraction of the total number of correctly classified non-failed sections divided by the total number of predicted non-failed sections (precision), and of the fraction of correctly classified non-failed road sections divided by the total number of non-failed sections analysed (recall). It is calculated according to the following equation (Parker, 2011):

Equation 5

$$F\_Score = 2 \times \frac{precision \times recall}{precision + recall}$$

$$where: \qquad precision = \frac{T_P}{T_P + F_P}$$

$$recall = \frac{T_P}{T_P + F_N}$$

A f-score can have a value of between zero and one, the closer the value is to one the more accurate the method is regarded (Parker, 2011; Sokolova *et al.*, 2009).

The phi coefficient was used to measure how well the SVM technique predicted pavement failures and non-failures. As a measure of performance it is often favoured above the f-score because it takes into account all correctly predicted values, as opposed to the f-score where it's constituent precision and recall values only take account of the correctly predicted non-failures. The phi coefficient was determined using Equation 6 (Parker, 2011):

Equation 6

$$Phi\ coefficient = \frac{T_P \times T_N - F_P \times F_N}{\sqrt{(N_1 \times N_2 \times N_3 \times N_4)}}$$

A positive phi coefficient means that the majority of the results are correctly predicted, and vice versa. A value of zero indicates that there is no relationship between the prediction and input variables (Parker, 2011).

CASE STUDY

Dataset

Data was obtained from the LTPP programme which monitors 63 sites on the New Zealand SH network (Henning, 2008; Henning *et al.*, 2004). The large majority of the pavements in the network are thin, flexible, unbound granular pavements, carrying low volumes of traffic (i.e. <10,000 vehicles per day).

Failure Charts

Although other modes of failure are recognised for other pavement types and environments, the focus of this research paper is on only the three predominant failure types on New Zealand's low volume roads, namely rutting, load associated fatigue cracking, and shear (Schlotjes *et al.*, 2011). For each of these failure types a failure chart was developed by canvassing the opinion of a panel of experts. To achieve this they identified six groups of factors which contribute to failure as follows:

- Traffic (e.g. annual average daily traffic);
- Pavement Composition (e.g. number of layers, thicknesses, and ages);
- Pavement Strength (e.g. structural number);
- Environment (e.g. rainfall);
- Surface Condition (e.g. percentage of cracking, rutting depth, etc.), and
- Subgrade Sensitivity (e.g. low, medium and high).

These were used to group the data types available in the SH LTPP dataset as shown in Table 1.

The factors were then sub-divided according to the opinion of the panel of experts and used to develop a failure chart for each failure type. The charts so developed are presented in Figure 2 to Figure 4, where the notation ‘Trial “X”’ correlates with the combinations of factor listed in Table 3. It may be seen that some combinations of factors (failure paths) occur for more than one failure type because of the similar interactions between factors within the types of failure. For example, both rutting and fatigue cracking can be due to a combination of excessive strain and poor pavement support, as a result of composition issues (as shown by Trial 7 in Figures 2 and 3).

Computational model

As described above the SVM technique, was used to determine the probability of failure of the road pavements in the SH LTPP network dataset. To this end, the technique was used to compute, for each of the three failure types identified, the likelihood of failure of all pavement sections by each possible failure path in the failure charts. In developing the SVM model, a 10-

fold cross-validation approach was followed, where a random 90 % sample of the dataset was used for training (Rogers and Girolami, 2012). The performance of the SVM modelling technique was demonstrated using a number of measures as described below.

Results and Analysis

The results of the analysis are divided into two parts:

1. An assessment of the performance measures to show the applicability of the SVM modelling technique for the task in hand, and
2. An analysis of the SH LTPP road network to demonstrate the usefulness of the suggested methodology.

Assessment of the SVM technique

Table 2 presents the average results from cross-validation tests of the performance measures, from which it may be seen that the SVM model predicted accurately the three types of pavement failure, according to the accuracy, misclassification and f-score measures used. The relatively lower values of the phi coefficient however suggest weaker relationships between the road dataset and the predicted failure for each failure type.

The prediction of rutting and fatigue cracking may be seen to be slightly better than for the shear component, by all three measures. Shear failure can be strongly linked to the properties of pavement materials, and unfortunately, this information is lacking in network level datasets. Therefore, further work is required in the development of the shear failure prediction component of the model.

Factors Associated With Failure

Table 3 summarises, for each of the three failure types considered, the computed most probable causes of failure for the entire SH LTPP road network and the associated corresponding number of pavement sections. For all three failure types, strength is shown to be a major factor. As far as rutting is concerned, the road pavement strength is the only significant factor for 64 % of the road pavements analysed. The predominant factors associated with fatigue cracking are strength, traffic, composition, environment, and subgrade sensitivity. The environment factor occurs in 46 % of the pavement sections which have failed by fatigue cracking, and since environment is a measure of the cumulative amount of rainfall falling on an already cracked pavement (Table 1), it would suggest that water ingress is a major factor contributing to the deterioration of this network. For shear failure, traffic, pavement composition and strength together are the likely to be contributing factors towards failure of nearly half of the network.

The few occurrences of surface condition in Table 3 and the fact that it does not occur alone for any of the three failure mechanisms suggest that the functional pavement condition is not a good predictor of failure. This suggests that visual road condition assessment, on which pavement maintenance is often based, may not be sufficient on its own to determine appropriate maintenance.

Failure Probability

A simplified method was adopted, based on Equation 2, to calculate the overall failure probability, and is as follows:

Equation 7

$$P_{FAILURE} = max\ [P(Rutting), P(Fatigue\ Cracking), P(Shear)]$$

However, it is recognised that a more complex approach could improve the reliability of the failure probabilities, given the noticeable occurrence of combined or secondary failure mechanisms. The investigation of such an approach is beyond the scope of this paper.

Accordingly, three outputs were calculated:

1. Probable causes of failure
2. The probability of failure for each of the failure type
3. The overall failure probability of road sections

Figure 5 shows the frequency distribution of the probability of overall failure of the road pavement sections analysed and therefore the overall risk profile of the SH LTPP road network. The histogram shows that the majority (97 %) of the pavement sections on the network has a probability of failure of less than 0.2, and 79 % with a failure probability of less than 0.1, which suggests that the network is in good condition. However, a small number (2 %) of pavement sections are predicted to have failed ($P_{FAILURE} > 0.5$), with a high probability of failure.

Figure 6 presents the distribution of the most probable failure modes on the SH LTPP road network, from which it may be seen that shear failure is the most probable.

Practical Application

The methodology presented facilitates both project and network level analysis of a road network. At the project level, pavement sections that have a high probability of failure can be identified, further assessed if necessary, and appropriately treated. Such a predictive approach is likely to be more cost effective than a reactive one. For failed sections of the network, the methodology allows an insight into the causes of failure, enabling an appropriate remedial treatment to be applied, and can support, or replace, expensive site investigations. For example, pavement section #3804 was identified as having failed by fatigue cracking. Using the developed model, the computed overall failure probability and those of the individual failure types are as follows:

$$P_{FAILURE\ (\#3804)} = max[\ P_{RUT} = 0.0442,\ P_{FATIGUE\ CRACK} = 0.9874,\ P_{SHEAR} = 0.0767]$$

$$P_{FAILURE\ (\#3804)} = \ P_{FATIGUE\ CRACK} = 0.9874$$

The associated factor combination of the $P_{FAILURE}$ is Trial 23 (refer Table 3). From the failure charts, Figure 3 in particularly, it can be seen that pavement failed in fatigue cracking due to poor pavement composition (from an aged pavement or insufficient pavement thickness). This resulted in poor pavement support, which when combined with excessive traffic loadings caused failure.

Pavement section #4249 failed in both rutting and shear and its computed failure probabilities are:

$$P_{FAILURE\ (\#4249)} = max[\ P_{RUT} = 0.6935,\ P_{FATIGUE\ CRACK} = 0.0213,\ P_{SHEAR} = 0.9507\ ]$$

$$P_{FAILURE\ (\#4249)}\ = P_{SHEAR} = 0.9507$$

Trial 7 are the most probable failure factor combinations for both rutting and shear. According to the failure charts (Figures 2 and 4) the most likely failure paths for both of these mechanisms are traffic and composition, and although the same factor combinations are in the critical failure paths, the most probable causes are different. Shear failure is generally related to material performance, and rutting is a result of induced strains from excessive, traffic loadings and strain repetitions, so the cause of these failures (both rutting and shear) can be attributed to poor composition of the pavement combined with excessive traffic loading.

At a network level, risk profiles (Figure 5) can be produced to identify the overall serviceability of the network and the predominant failure mechanism(s) (Figure 6). This enables appropriate and timely maintenance to be carried out adjustments of any maintenance regimes,. Furthermore potential changes in the external environment can be quantified in terms of potential effects on network condition. For example, if traffic loading on the network is set to increase, the projected increase in loading can be included in the input parameters in the SVM and the effects modelled. Similarly the effects of the potential changes in precipitation due to climate change could be estimated.

CONCLUSIONS

A method has been developed which is capable of assessing the probability of the structural failure of road pavements. The method was based on using expert judgment to develop failure charts for the predominant types of failure on a road pavement which can be used to identify the contributing factors to pavement failure. A computational technique, known as SVM, was developed to analyse the probability of failure of pavement datasets and to determine the most probable failure paths for each failure type. The resulting probabilities for each failure type were used, in a simple approach, to determine an overall probability of pavement failure. Further work is being undertaken to calculate the overall failure probability within the SVM modelling process.

A case study using data from the New Zealand SH LTPP programme was used to demonstrate the performance of the proposed methodology. Four performance measures were used to assess the precision of the SVM technique in determining the probability of failure of pavements via three failure types, rutting, fatigue cracking and shear failure. Although the SVM performed satisfactorily in predicting failure, further development in the prediction of shear failure, and consideration of combined failure modes are both desirable and necessary. The analysis of the New Zealand dataset showed that the network may be regarded as being in a good condition, although a small number of pavement sections within the network have a high likelihood of failure. From this study it is evident that measured functional pavement condition alone is not a good identifier of failure, and that the inclusion of a parameter related to strength, such as pavement deflection, is essential.

REFERENCES

Caruana R and Niculescu-Mizil A (2006) An empirical comparison of supervised learning algorithms. Proceedings of the 23rd International Conference on Machine Learning, Pittsburgh, PA, USA.

Chandra DK, Ravi V and Bose I (2009) Failure prediction of dotcom companies using hybrid intelligent techniques. *Expert Systems with Applications* **36(3)**: 4830-4837 (Part 1).

Evdorides HT (1994) A Knowledge Based System for Pavement Analysis. Birmingham, UK: University of Birmingham [PhD Thesis].

Henning TFP (2008) The Development of Pavement Deterioration Models on the State Highway Network of New Zealand. Auckland, New Zealand: The University of Auckland [PhD Thesis].

Henning TFP, Costello SB, Dunn RCM, Parkman CC and Hart G (2004) The establishment of a long-term pavement performance study on the New Zealand State Highway network. *Road and Transport Research* **13(2)**: 17-32.

Isa AHM, Ma'soem DM and Hwa LT (2005) Pavement performance model for federal roads. Proceedings of the Eastern Asia Society for Transportation Studies **5**: 428-440.

Martin TC (2008) Predicting Sealed Granular Pavement Deterioration at a Road Network Level. Victoria, Australia: Monash University [PhD Thesis].

Pal M (2006) Support vector machines-based modelling of seismic liquefaction potential. International Journal for Numerical and Analytical Methods in Geomechanics **30(10)**: 983-996.

Park ES, Smith RE, Freeman TJ and Spiegelman CH (2008) A bayesian approach for improved pavement performance prediction. Journal of Applied Statistics **35(11)**: 1219-1238.

Parker C (2011) An analysis of performance measures for binary classifiers. *Proceedings of the 11th IEEE International Conference on Data Mining*, Vancouver, Canada.

Patev RC, Putcha C and Foltz SD (2005) Methodology for Risk Analysis of Dam Gates and Associated Operating Equipment Using Fault Tree Analysis. US Army Corps of Engineers, USA.

Reigle J (2000) Development of an Integrated Project-Level Pavement Management Model Using Risk Analysis. West Virginia, USA: West Virginia University [Ph.D. Thesis].

Robinson R, Danielson U and Snaith M (1998) Road Maintenance Management: Concepts and Systems. Palgrave, New York, USA.

Rogers S and Girolami M (2012) *A First Course in Machine Learning*. Taylor and Francis Group, Florida, USA.

Schlotjes MR, Henning TFP and Burrow MPN (2011) Improved network understanding – a diagnostic approach to the risk of pavement failure. *Proceedings of the 8th International Conference on Managing Pavement Assets*, Santiago, Chile.

Schlotjes MR, Henning TFP and Burrow MPN (2012) The Development of a Diagnostic Approach that Predicts Failure Risk of Flexible Road Pavements. *Proceedings of the 7th International Conference on Maintenance and Rehabilitation of Pavements and Technological Control (MAIREPAV)*, Auckland, New Zealand.

Seyed-Hosseini SM, Safaei N and Asgharpour MJ (2006) Reprioritization of failures in a system failure mode and effects analysis by decision making trial and evaluation laboratory technique. Reliability Engineering and System Safety **91(8)**: 872-881.

Sokolova M and Lapalme G (2009) A Systematic Analysis of Performance Measures for Classification Tasks. *Information Processing and Management* **45(4)**: 427-437.

Tu J (1996) Advantages and disadvantages of using artificial neural networks versus logistic regression for predicting medical outcomes. Journal of Clinical Epidemiology **49(11)**: 1225-1231.

Van Looy S, Verplancke T, Benoit D *et al.* (2007) A novel approach for prediction of tacrolimus blood concentration in liver transplantation patients in the intensive care unit through support vector regression. *Critical Care* **11(4)**.

Volinsky CT, Madigan D, Raftery AE and Kronmal RA (1997) Bayesian model averaging in proportional hazard models: Assessing the risk of a stroke. Journal of the Royal Statistical Society. Series C (Applied Statistics) **46(4)**: 433-448.

Wang L (ed.) (2005) *Support Vector Machines: Theory and Applications*. Springer, Heidelberg, Germany.

Xiao N, Huang HZ, Li Y, He L and Jin T (2011) Multiple failure modes analysis and weighted risk priority number evaluation in FMEA. Engineering Failure Analysis **18(4)**: 1162-1170.

TABLE 1 Factor Combinations for Modelling Using LTPP Data

Factor Group	Variables included in the Group
Traffic	- Average Annual Daily Traffic (AADT)[a,b,c] - Total percentage of Heavy Vehicles[a,b,c] - Cumulative number of Equivalent Standard Axles (ESA), given the base layer age[a,b,c] - Cumulative number of Equivalent Standard Axles (ESA), given the surface layer age[b]
Pavement Composition	- Base layer age[a,b,c] - Subbase layer age[a,b,c] - Surface age[b] - Total pavement thickness, excluding surface thickness[a,b,c] - Total pavement thickness, including surface thickness[b] - Pavement width[a,b,c] - Number of lanes[a,b,c]
Pavement Strength	- Strength of pavement (weak or strong)[a,b,c] - Structural number (SNP)[a,b,c] - Structural indices (SI) for rutting, flexure, shear and roughness[a,b,c] - Falling weight deflectometer (FWD) parameters[a,b,c]
Environment	- Cumulative rainfall once the pavement is cracked[a,b,c]
Surface Condition	- Rut depths for left-hand wheelpath, right-hand wheelpath, and lane[a,c] - Rut rate for left-hand wheelpath, right-hand wheelpath, and lane[a] - Total cracking (all cracking types)[b] - Crack rate[b] - Number of years of continual cracking[b] - Mechanical damage[c] - Structural patch[c] - Pothole diameter, depth and number[c] - Shoving[c]
Subgrade Sensitivity	- Sensitivity of pavement [a,b,c]

[a] – Rutting dataset; [b] – Fatigue cracking dataset; [c] – Shear dataset

TABLE 2 Summary of the Performance Measures

	Average Value over all Failure Paths			
	Accuracy (%)	**Misclassification (%)**	**F-score**	**Phi coefficient**
Rutting	97.70	2.30	0.99	0.22
Fatigue Cracking	98.21	1.79	0.99	0.31
Shear	94.52	5.48	0.97	0.16

TABLE 3 Factor Combinations of the SH LTPP Network per Failure Mechanism

Trial	Combinations of Factors	Rutting	Fatigue Cracking	Shear
3	Strength	3596	782	-
7	Traffic + Composition	45	547	177
8	Traffic + Strength	0	0	-
12	Composition + Strength	932	348	703
16	Strength + Environment	6	-	-
18	Strength + Subgrade Sensitivity	0	0	865
22	Traffic + Composition + Strength	120	313	1150
23	Traffic + Composition + Environment	-	199	-
24	Traffic + Composition + Surface Condition	-	45	-
25	Traffic + Composition + Subgrade Sensitivity	0	131	118
26	Traffic + Strength + Environment	0	396	-
28	Traffic + Strength + Subgrade Sensitivity	0	0	0
32	Composition + Strength + Environment	0	186	37
33	Composition + Strength + Condition	-	125	-
34	Composition + Strength + Subgrade Sensitivity	0	239	212
39	Strength + Environment + Subgrade Sensitivity	871	-	-
42	Traffic + Composition + Strength + Environment	0	151	182
43	Traffic + Composition + Strength + Surface Condition	-	253	-
44	Traffic + Composition + Strength + Subgrade Sensitivity	0	-	962
46	Traffic + Composition + Environment + Subgrade Sensitivity	3	103	-
47	Traffic + Composition + Surface Condition + Subgrade Sensitivity	-	18	-
49	Traffic + Strength + Environment + Subgrade Sensitivity	42	1077	614
50	Traffic + Strength + Surface Condition + Subgrade Sensitivity	-	229	-
52	Composition + Strength + Environment + Surface Condition	-	11	-
53	Composition + Strength + Environment + Subgrade Sensitivity	0	239	143
54	Composition + Strength + Surface Condition + Subgrade Sensitivity	-	4	-
57	Traffic + Composition + Strength + Environment + Surface Condition	-	3	-
58	Traffic + Composition + Strength + Environment + Subgrade Sensitivity	13	212	465
59	Traffic + Composition + Strength + Surface Condition + Subgrade Sensitivity	-	6	-
62	Composition + Strength + Environment + Surface Condition + Subgrade Sensitivity	-	8	-
63	Traffic + Composition + Strength + Environment + Surface Condition + Subgrade Sensitivity	-	3	-
	TOTAL	**5628**	**5628**	**5628**

- Not Applicable

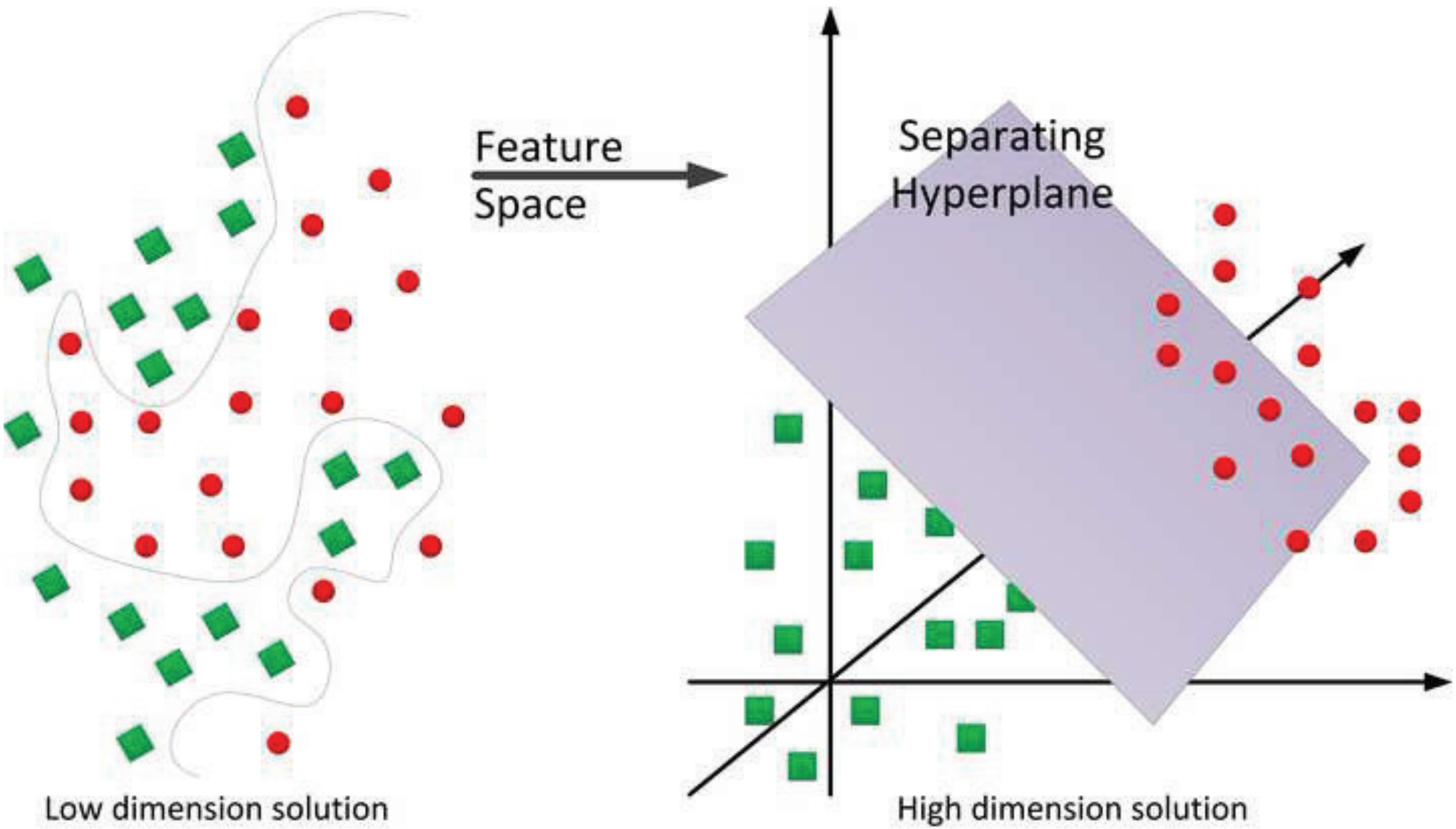

Feature
Space
Separating
Hyperplane
Low dimension solution
High dimension solution

Figure
Click here to download high resolution image

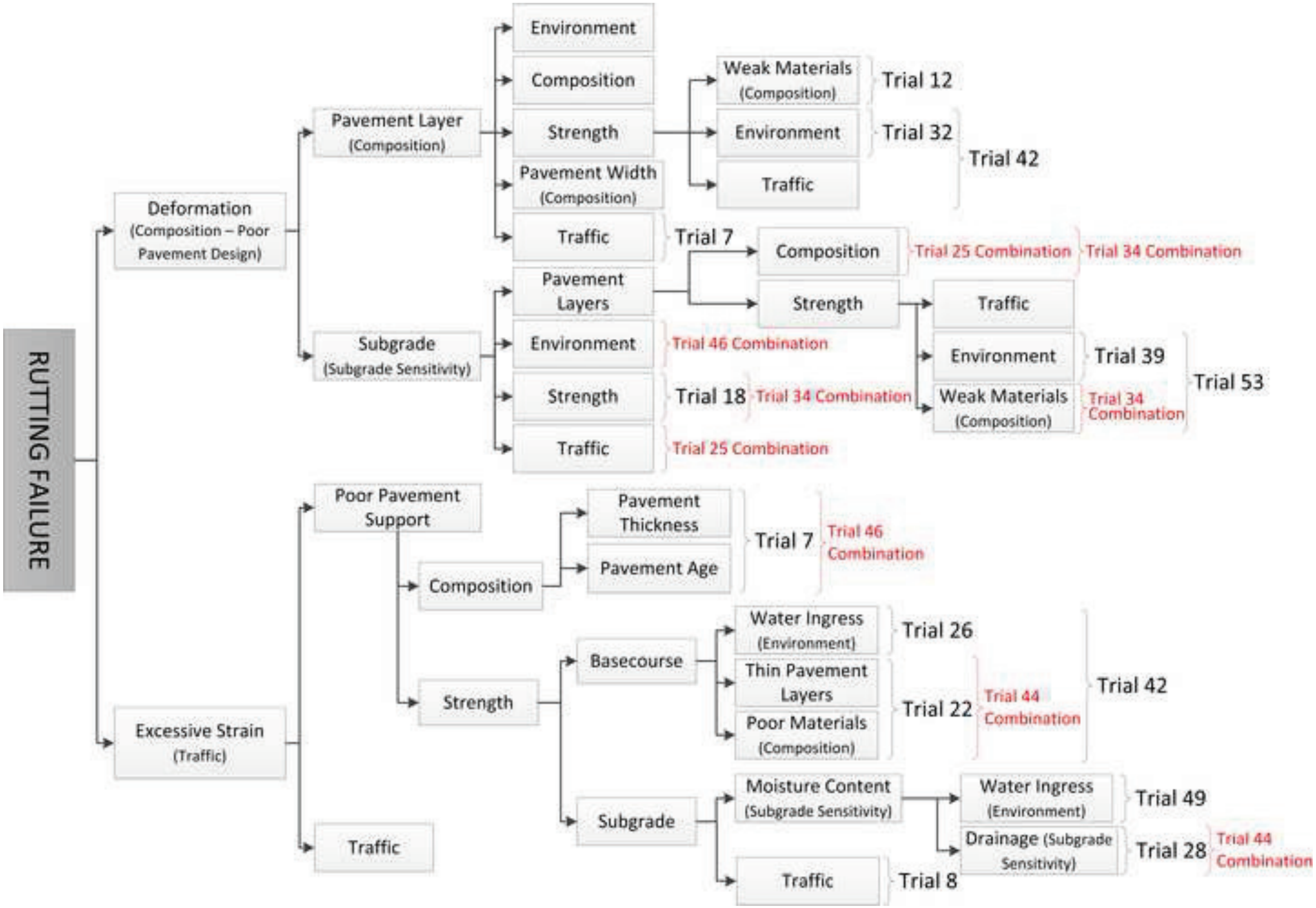


Figure
Click here to download high resolution image

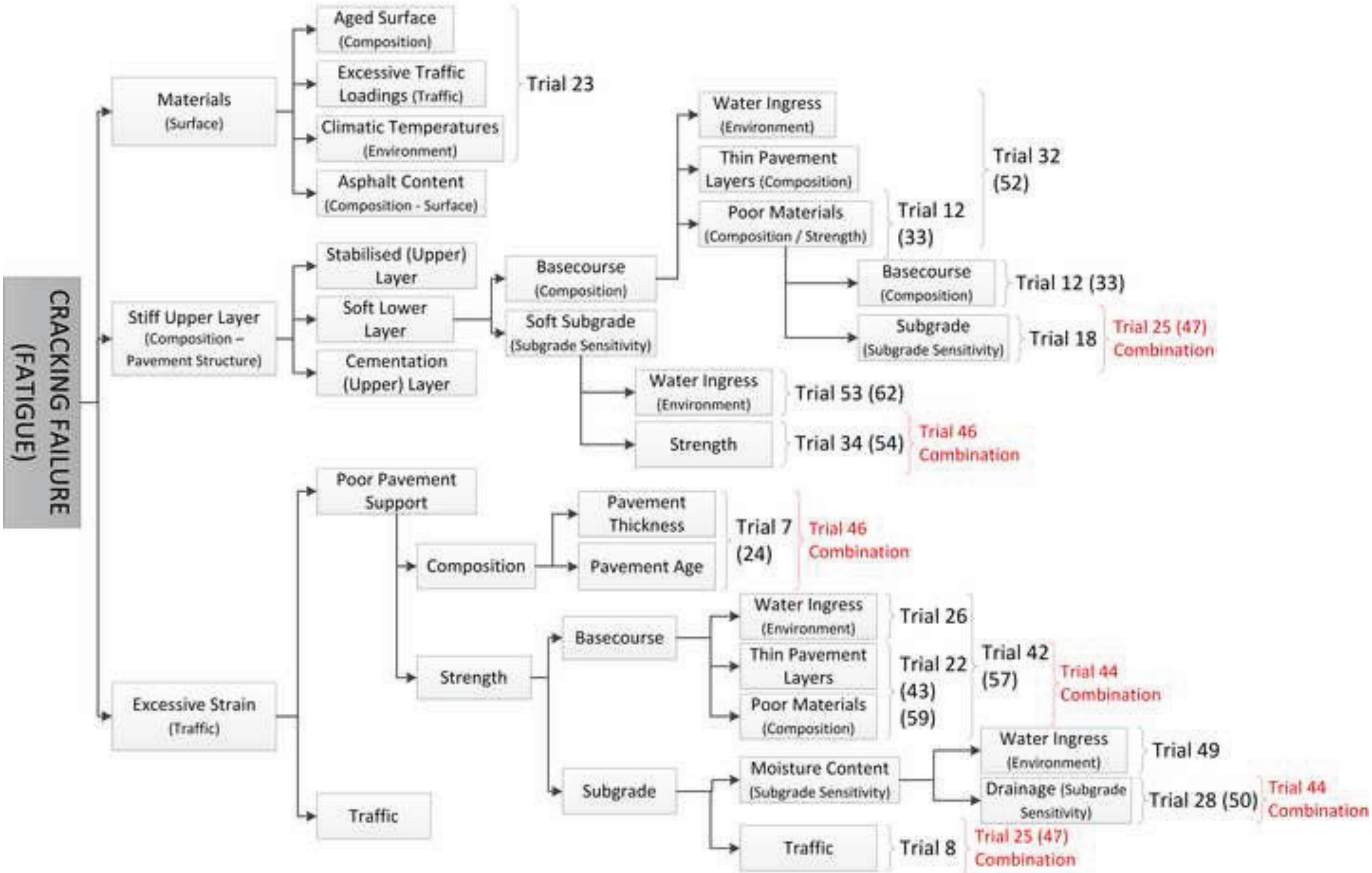
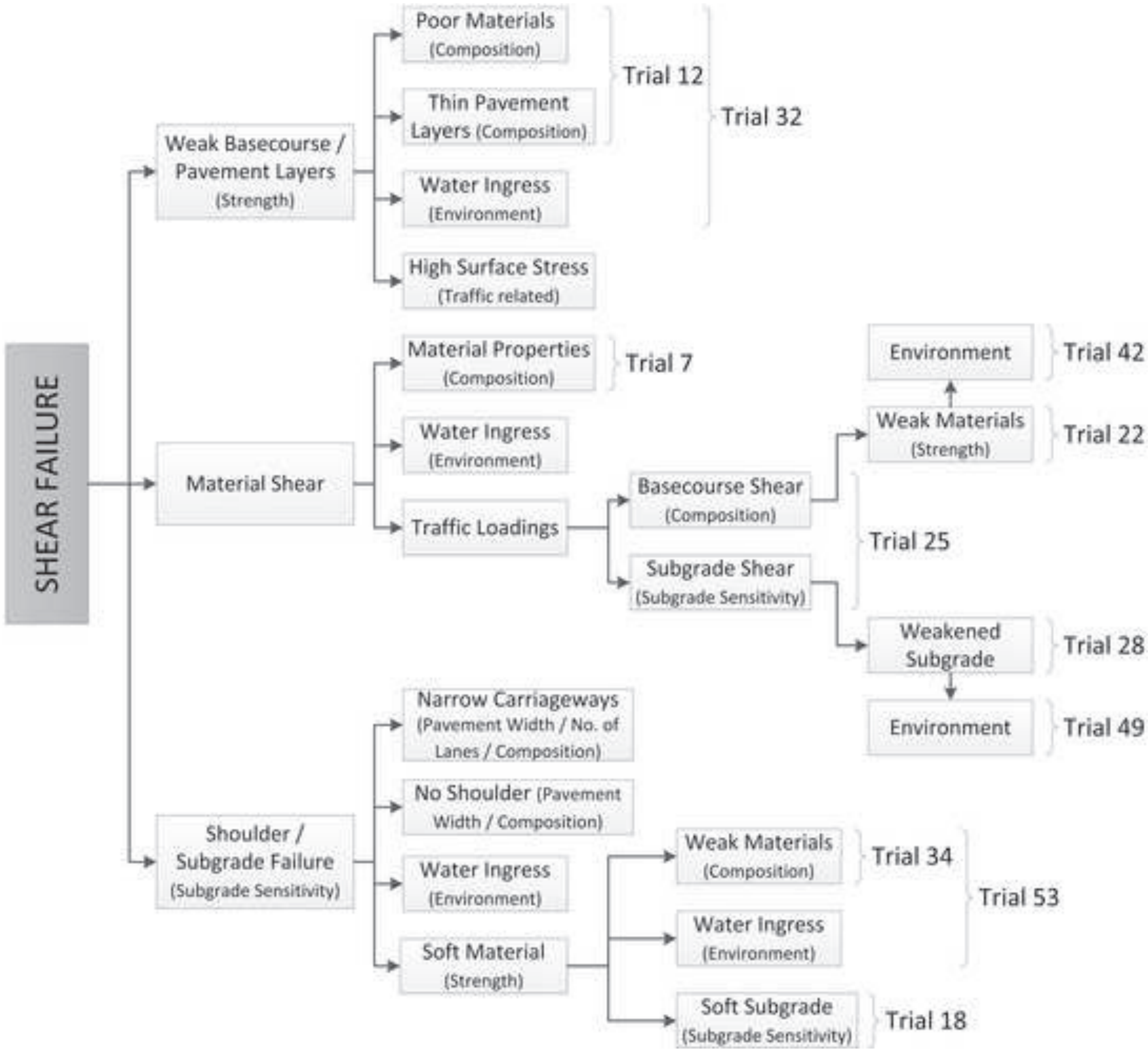
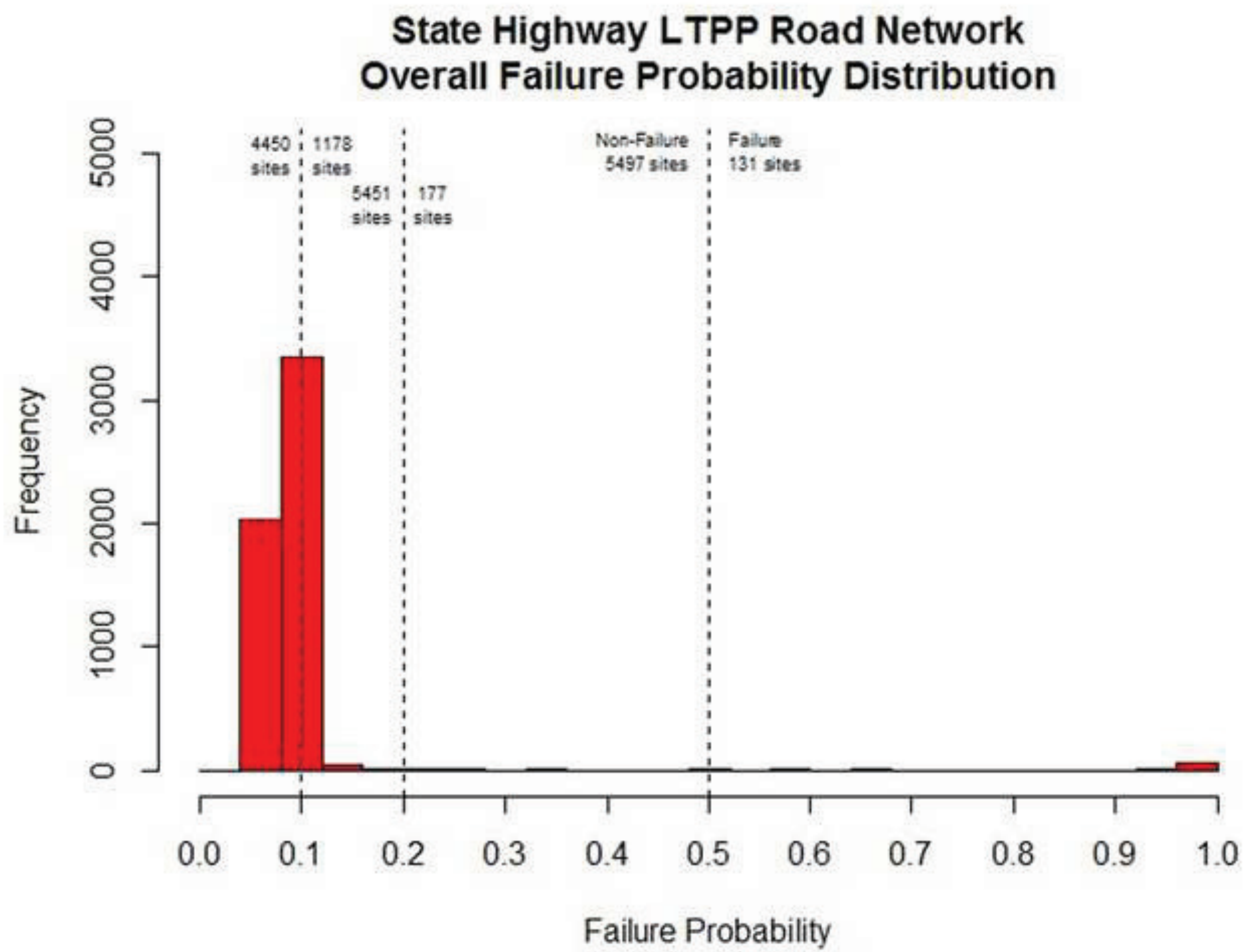


Figure
Click here to download high resolution image

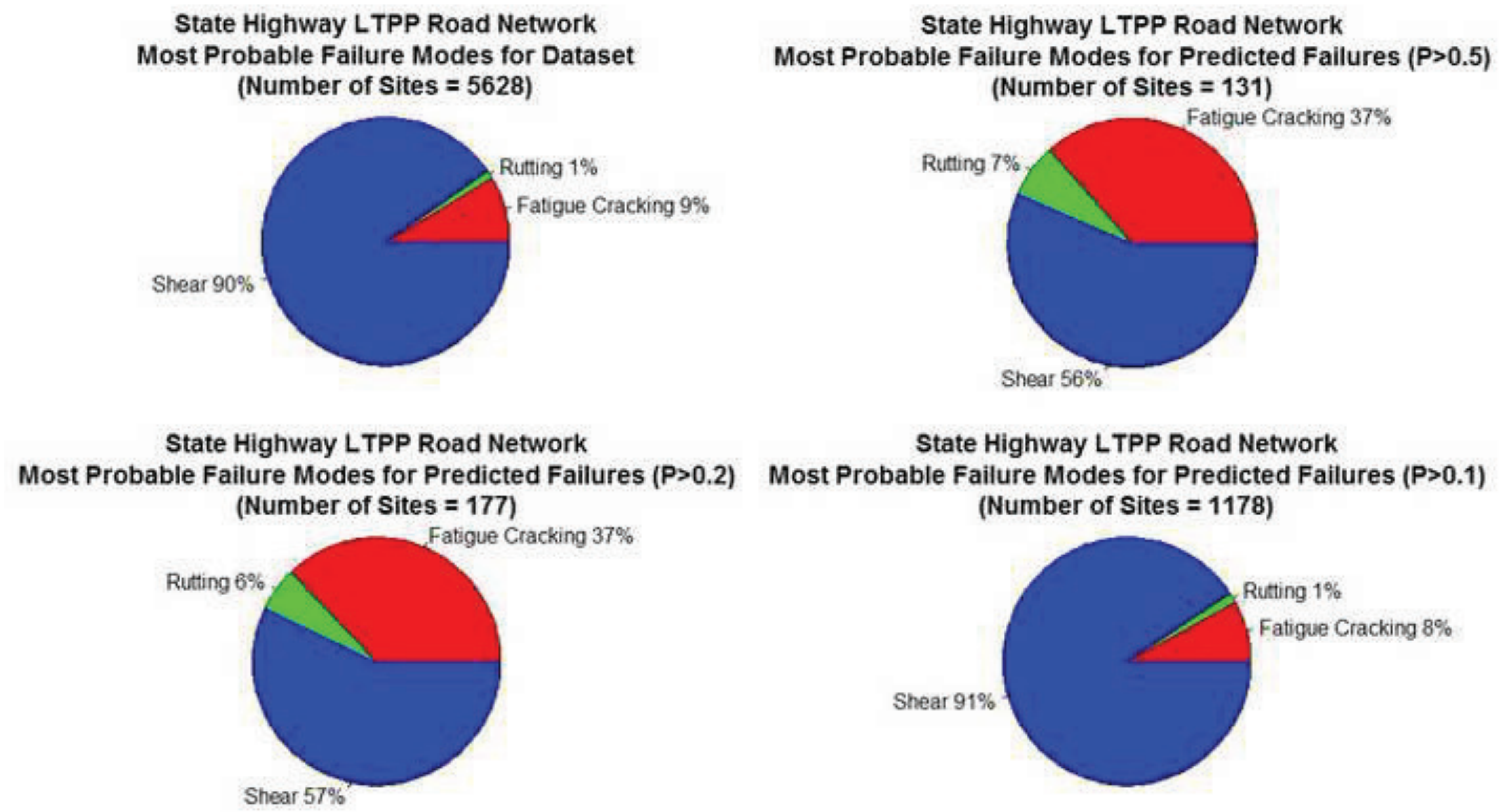


Figure

Click here to download high resolution image



Figure


State Highway LTPP Road Network
Most Probable Failure Modes for Dataset
(Number of Sites = 5628)
Rutting 1%
Fatigue Cracking 9%
Shear 90%
State Highway LTPP Road Network
Most Probable Failure Modes for Predicted Failures (P>0.5)
(Number of Sites = 131)
Fatigue Cracking 37%
Rutting 7%
Shear 56%
State Highway LTPP Road Network
Most Probable Failure Modes for Predicted Failures (P>0.2)
(Number of Sites = 177)
Fatigue Cracking 37%
Rutting 6%
Shear 57%
State Highway LTPP Road Network
Most Probable Failure Modes for Predicted Failures (P>0.1)
(Number of Sites = 1178)
Rutting 1%
Fatigue Cracking 8%
Shear 91%