

# OPTIMIZING ARTIFICIAL NEURAL NETWORKS FOR THE EVALUATION OF ASPHALT PAVEMENT STRUCTURAL PERFORMANCE

---

GAETANO BOSURGI<sup>1</sup>, ORAZIO PELLEGRINO<sup>2</sup>,  
GIUSEPPE SOLLAZZO<sup>3\*</sup>

<sup>1,2,3</sup>*Dept of Engineering, University of Messina, Messina, Italy*

Received 18 July 2018; accepted 11 December 2018

**Abstract.** Artificial Neural Networks represent useful tools for several engineering issues. Although they were adopted in several pavement-engineering problems for performance evaluation, their application on pavement structural performance evaluation appears to be remarkable. It is conceivable that defining a proper Artificial Neural Network for estimating structural performance in asphalt pavements from measurements performed through quick and economic surveys produces significant savings for road agencies and improves maintenance planning. However, the architecture of such an Artificial Neural Network must be optimised, to improve the final accuracy and provide a reliable technique for enriching decision-making tools. In this paper, the influence on the final quality of different features conditioning the network architecture has been examined, for maximising the resulting quality and, consequently, the final benefits of the methodology. In particular, input factor quality (structural, traffic, climatic), “homogeneity” of training data records and the actual net topology have been investigated. Finally, these results further prove the approach efficiency, for improving Pavement Management Systems and reducing deflection survey frequency, with remarkable savings for road agencies.

**Keywords:** Artificial Neural Network (ANN), asphalt pavement, Long Term Pavement Performance (LTPP), neural network optimisation, Pavement Management System (PMS), structural performance.

---

\* Corresponding author. E-mail: [gsollazzo@unime.it](mailto:gsollazzo@unime.it)

## Introduction

Pavement Management Systems (PMSs) are practical tools for road agencies, as they simplify road network administration and represent potential decision supporting tools for budget allocation and maintenance planning. However, their reliability and effectiveness are strictly related to the level of knowledge of the infrastructure asset characteristics. It is essential for road agencies, indeed, to rely their PMSs on exhaustive databases including information on the main pavement performance parameters (structural and functional) and the performed maintenance and rehabilitation activities. However, these datasets must be updated and fed continuously, allowing operators to know the deterioration trend of the performance parameters. This practice improves maintenance effectiveness and reduces diseconomies in budget allocation, increasing user satisfaction.

Modern survey methodologies guarantee economic and reliable high-performance data collection for roughness, texture, friction, or surface distresses – using high-speed profilers or laser-lightning detection systems (Sollazzo, Wang, Bosurgi, & Li, 2016; Wang, 2011). However, the evaluation of pavement structural performance is generally based on deflection measurements performed using the Falling Weight Deflectometer (FWD). These test procedures are slow, expensive, and cause adverse effects on traffic, due to the stop-and-go procedure (Elseifi, Abdel-Khalek, & Dasari, 2012; Rada, Perera, & Prabhakar, 2012). Despite the introduction of the Rolling Weight Deflectometer (RWD) equipment, the FWD methodology is still more widespread and reliable.

Then, road agencies likely need novel techniques for reducing deflection survey frequency and, thus, costs and interferences on traffic, assuring an adequate knowledge of the pavement structural performance. Although some authors tried to correlate functional to structural parameters for simplifying the problem, some direct mechanical correlations among specific parameters (such as roughness and structural capacity) are still undefined and unclear (Bianchini & Bandini, 2010; Rada, Perera, & Prabhakar, 2012; Zhang, Manuel, Damnjanovic, & Li, 2003), increasing analytical issues. Moreover, the problem is remarkably complex, and it is impractical to rely on mechanical-based evolution models requiring consistent computational and economic efforts. Because of this, various researchers proposed the introduction of structural performance indices in pavement management systems and developed numerical estimation models (Agarwal, Das, & Chakroborty, 2006; Tosti, Ciampoli, D'Amico, Alani, & Benedetto, 2018; Wu, Zhang, & Abadie, 2012).

However, the problem complexity and the lack of knowledge in some of the physical and mechanical relationships among the involved processes suggest the adoption of numerical and advanced “machine learning” approaches – and, in particular, Artificial Neural Networks (ANNs) – for solving the issue (Adeli, 2001). As it is known, ANNs have been widely applied in different areas of civil engineering (for example, structural, construction, environmental, geotechnical and infrastructure engineering) with positive results (Bosurgi & Trifirò, 2005; Bosurgi, D’Andrea, & Pellegrino, 2013; Ceylan, Bayrak, & Gopalakrishnan, 2014; Fwa & Chan, 1993; He, Qi, Hang, Zhao, & King, 2014; Pozarycki, 2015; Roberts & Attoh-Okine, 1998; Terzi, 2007).

Considering pavement engineering, numerous efforts have been spent on evaluating pavement functional performance using ANN (Attoh-Okine, 1994; Eldin & Senouci, 1995; Shekharan, 1999; Terzi, 2007; Ziari, Sobhani, Ayoubinejad, & Hartmann, 2016). However, in terms of structural performance evaluation, ANNs were adopted only for layer moduli estimation, back-calculation procedures or deflection measure processing (Plati, Georgiou, & Papavasiliou, 2016; Rakesh, Jain, Reddy, M., & Reddy, A., 2006; Yi, Kim, Y., Mun, & Kim, J., 2010). Consequently, by applying ANN potentiality to this issue also, it is possible to define useful models for assuring high levels of quality and accuracy to the final estimation of the pavement structural performance. In particular, ANNs seem to be useful for estimating asphalt pavement structural performance, considering specific influencing external factors (climatic, traffic, layers) and functional indices of the pavement, for reducing expensive and time-consuming deflection tests. However, the method is complicated and, thus, different aspects regarding the involved factors, the database characteristics and the net architecture obviously affect the final accuracy of the methodology. Then, it is important to optimise the architecture of such a network, investigating the possible advantages and drawbacks of different possible network configurations, for increasing the reliability of the final evaluation. At this regard, performing various tests and experimentations helps to deeply clarify the methodology potential and determine the choices increasing the approach accuracy and reliability.

Then, in this paper, the authors analyse the numerical reaction to different network parameters of an ANN defined for the estimation of asphalt pavement structural performance (considered regarding effective Structural Number  $SN_{eff}$  (Transportation Officials, 1993)). The improved methodology, applied in Pavement Management Systems, simplifies pavement maintenance planning and analysis, due to the possible consequent reduction in the frequency of FWD deflection tests. In periods among two performed deflection tests, in fact, by using the proposed approach, the structural performance derived directly from

available data (functional, geometrical, climatic, and traffic information), more economical, faster, and more efficient. In detail, for an appropriate methodology optimization, different aspects are examined for assessing their influence on the model accuracy such as:

- the selection of the most appropriate factors;
- the ANN architecture;
- the dataset homogeneity;
- the combination of different ANN in “committees of ANNs”, for better coverage of the entire solution space.

In the following Sections, after brief theoretical notices on the ANNs, the results of several numerical tests are evaluated and discussed to assess efficiency and feasibility of the different net configurations.

## 1. Artificial Neural Networks and “committees of Artificial Neural Networks”

Artificial Neural Networks are computational models defined in analogy with the brain biological characteristics to simulate its decision process. They are used to approximate and estimate unknown functions depending on various and numerous input values, for solving complex and nonlinear problems using only elementary mathematical operations (Graupe, 2013; Priddy & Keller, 2005). In particular, ANNs represent a “black-box” approach, since the results are produced neglecting the input-output causal relationships (Rohani, Abbaspour-Fard, & Abdolahpour, 2011). In big-data analyses, the method potentiality is convenient and assures the development of generalised solutions to complicated problems using large sets of example data (Flood & Kartam, 1994).

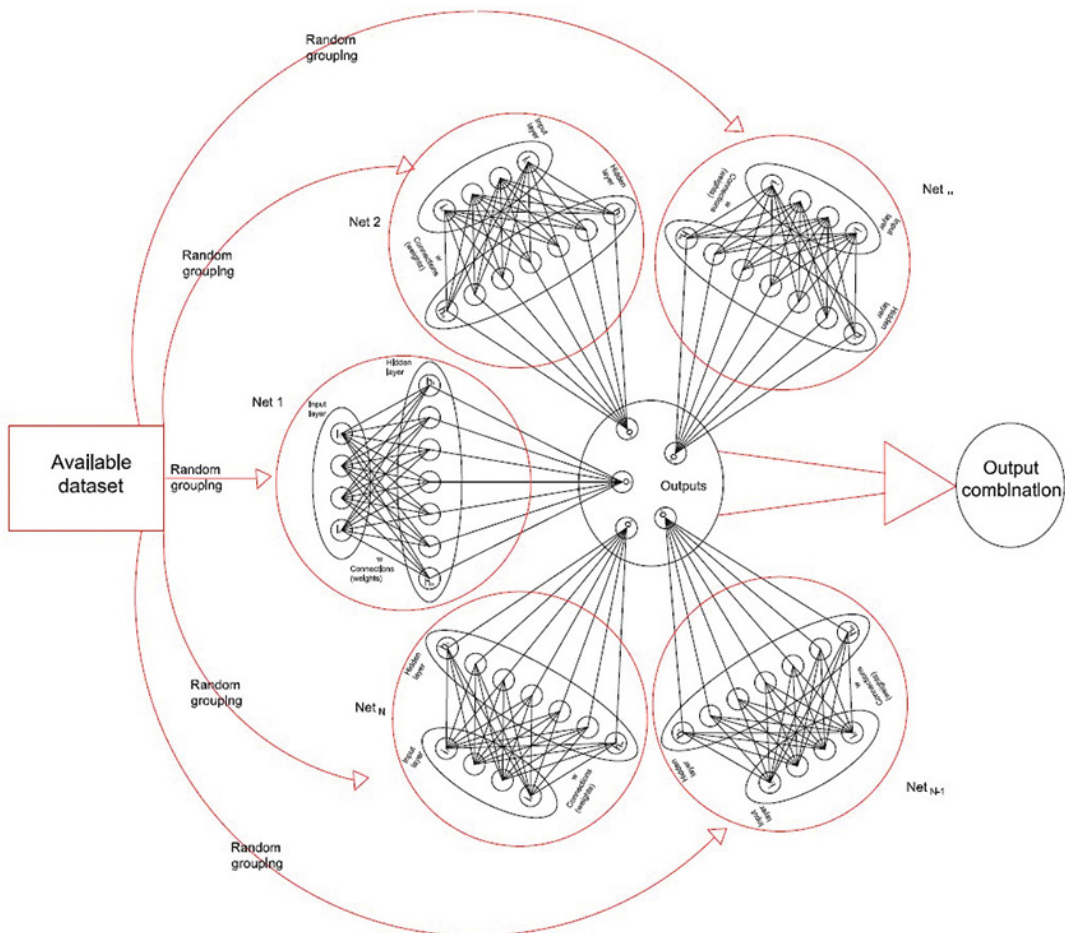
The Multilayer Feed-forward Neural Network (MFNN) is the most common type of ANN and includes three kinds of layers of interconnected neurons: input, hidden, and output layers. Each neuron processes the received inputs and, according to a properly defined activation function, produces an output. The neuron output is transmitted to the following neurons through specific connections defining the network topology. Each connection is associated with a specific weight ( $w_i$ ), which amplifies or reduces the input. For each neuron, the inputs ( $x_i$ )-output( $y_i$ ) relationship is defined using a specific transfer function, which usually has the logistic sigmoidal shape as follows, Eq. (1):

$$f(I) = \frac{1}{1+e^{-I}} \quad (1)$$

where  $I = \sum w_i x_i$ , is the sum of the weighted inputs  $x_i$ , produced by the previous neurons.

In a “supervised approach” – such as MFNN –, given a large set of input and output data, the training procedure consists in the modulation of the various weights to produce acceptable outputs. The results are expected to be very similar to the output provided for training. Usually, the training phase is performed using a back propagation model (Doughty, 1997) allowing the network to adjust the weights in a reverse direction, distributing the error among the different neurons and minimising it after each iteration.

To increase the efficiency of ANNs, different researchers have proposed novel approaches by combining the results of different networks (Alexandre, Campilho, & Kamel, 2000; Hashem, 1997; Hashem



**Figure 1.** Representation of an Artificial Neural Network Committee

& Schmeiser, 1995; Kuncheva, 2002). Among the various proposed methods, a simple solution consists of averaging the results of several ANNs, trained in parallel using different records, randomly extracted from the same dataset (Figure 1). In detail, each generated network is characterised by a different assignment of the available records to the training, validation, and test pools. This random procedure generates different networks and is expected to provide different estimates of the output variable. In this way, the outcome better represents the entire variable space, avoiding mistakes and inaccuracies possibly affecting a single net, even “the best”.

## 2. Training database

The training database represents the core element of the methodology. As discussed in Section 3, the quality of the model is strictly related to the available data characteristics. Three are the main aspects of the dataset definition:

- influencing input factors;
- database size;
- record “homogeneity”.

In general, owing to the relationship complexity and the specific features of the ANN, the dataset is expected to contain numerous significant variables affecting the structural performance. Apparently, as stated in previous studies, the higher the influence of the included variable on the output parameter, the more reliable the final result. Further, to produce a general model and guarantee an exhaustive ANN training in the entire solution space, it is essential to define an extensive and well-distributed dataset of road sections, to avoid model specification on reduced data sample. Finally, as a counterpart of the previous point, the higher the heterogeneity level of the adopted records – because of the characteristics of the analysed road network –, the lower the accuracy of the numerical correlation.

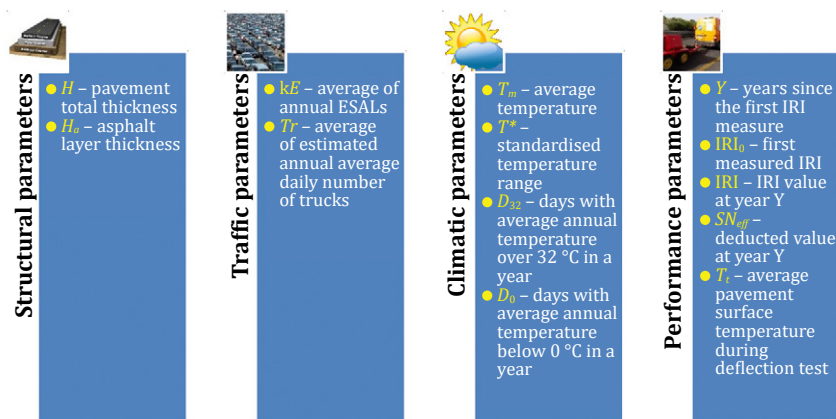
The LTPP database has been considered in the study to define an appropriate dataset in compliance with the previous requirements, as it contains several structural, traffic, climatic, and performance types of information related to over 2,500 North-American highway test sections for more than 25 years. The LTPP program was established as a part of the Strategic Highway Research Program (SHRP) and has been widely used as a reference database for analytical model definition and validation (Park, Thomas, and & Wayne Lee, 2007; Rada, Perera, & Prabhakar, 2012; Wang & Li, 2011). Maintenance and rehabilitation operations have been excluded from the dataset, i.e., for each section, the measurements following

the first maintenance operation have been excluded. The structural performance – representing the output variable – is evaluated in terms of effective Structural Number, according to the formulation provided by the AASHTO *Guide for Design of Pavement Structures* (Transportation Officials, 1993). Regarding the record size and features, different configurations have been adopted in the present work, in compliance with the specific research needs, as described in the following Section.

Concerning the input variables, the investigation starts from 12 “base factors” (Figure 2), previously selected in a similar ANN application (Sollazzo, Fwa, & Bosurgi, 2017).

These base factors, considered as a preliminary reference, were selected for representing significant variables related to traffic, climate, the structural configuration of the pavement section and functional performance and because they are expected to strongly condition the structural performance variations:

1. Structural parameters:
  - pavement total thickness ( $H$ ) in inches, including asphalt layers and eventual subbase;
  - asphalt layer thickness ( $H_a$ ) in inches;
2. Traffic parameters:
  - an average of annual ESALs (Equivalent Single Axle Load) in thousands in the LTPP lane ( $kE$ );
  - an average of the estimated annual average daily number of trucks in the LTPP lane ( $Tr$ );
3. Climatic parameters:
  - average temperature ( $T_m$ ), i.e., the mean of the annual average temperatures on the selected section in Celsius degrees;



**Figure 2.** Artificial Neural Network “base factors”

- standardised temperature range ( $T^*$ ), as a measurement of the temperature oscillation; this parameter has been evaluated using Eq. (2), where  $T_{max}$  and  $T_{min}$  are, respectively, the mean values of the annual maximum and the annual minimum temperature in Celsius degrees

$$T^* = \frac{T_{max} - T_{min}}{T_m}; \quad (2)$$

- the average number of days with average annual temperature over 32 °C in a year ( $D_{32}$ );
  - the average number of days with average annual temperature below 0 °C in a year ( $D_0$ );
4. Performance parameters:
- time passed ( $Y$ ) since the first profilometer survey in years for each section;
  - the first measured International Roughness Index (IRI) for each section ( $IRI_0$ ) – i.e., the IRI measured at year zero –, as starting reference value;
  - IRI value at a specific time for each section (average of left and right wheel path IRI);
  - $SN_{eff}$  at a specific time, according to the formulation provided by the AASHTO *Guide for Design of Pavement Structures* (Transportation Officials, 1993) as a function of deflection test measurements and pavement thickness;
  - the average pavement surface temperature ( $T_t$ ) of the section during the deflection test phase in Celsius degrees.

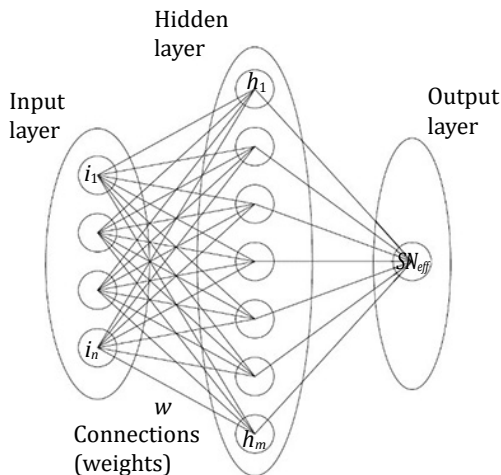
In addition to these base factors, two novel factors, with a considerable theoretical influence on the structural performance, have also been included in the dataset, for further analysis and investigation:

- subgrade resilient modulus ( $MR$ ); it is a structural parameter estimated through the procedure proposed in the AASHTO *Guide for Design of Pavement Structures* (Transportation Officials, 1993) for the  $SN_{eff}$  evaluation;
- annual average precipitation height ( $Pr$ ); it is a climatic parameter defined as the average of the precipitation heights (in mm) measured in the road section.

### 3. Numerical analysis

Several ANN architecture configurations have been considered to evaluate the influence of different aspects on the method accuracy. However, for simplifying the analysis, some general issues have been





**Figure 3.** The general architecture of the Artificial Neural Network

maintained invariant for all tests. The general architecture of the ANN is represented in Figure 3 (the target variable is always the  $SN_{eff}$ ). For all the ANNs, the considered records have been randomly divided among the training (70%), validation (15%), and test (15%) groups. Training was performed using the Levenberg-Marquardt algorithm and performance was measured concerning the Mean Square Error (MSE). The different aspects analysed in this paper, are discussed separately in the following Subsections.

### 3.1. Factor influence

By considering the variables introduced in Section 2, different architectures have been defined to further prove the influence of the various factors and verify the following assumptions:

- 1) adding other factors to the variable pool increases the correlation accuracy;
- 2) the more appropriate the factor selection (regarding influence on the output variable), the higher the resulting quality;
- 3) less strongly correlated factors are more efficient than many weakly correlated factors.

In detail, considering a starting dataset of 1021 records, four ANNs have been defined for exploring different scenarios and proving the previous assumptions:

- R – “reference net”: network trained considering all the base factors (12 factors);

- $F_1$ : network trained excluding climatic factors ( $T_m, T^*, D_0, D_{32}$ ) from the base factors (8 total factors);
- $F_2$ : network trained adding the two novel parameters introduced in Section 2 ( $MR, Pr$ ) to the base factors (14 total factors);
- $F_3$  – network trained considering the 12 factors (among the total 14 available) with the highest correlation with  $SN_{eff}$ ; compared to  $R, Y$  and  $T^*$  were replaced by the two novel parameters ( $MR, Pr$ ), according to the correlation rank (Pearson correlation coefficient) reported in Figure 4.

The results of the four ANNs are compared in Figures 5 and 6, in which the regression charts of the four nets (comparing target vs output values) and the related MSE values are shown respectively. Both Figures provide a clear validation of all the hypotheses: the  $R^2$  values evidence a better performance of  $F_2$  and  $F_3$  than  $R$ , and, despite the exclusion of two factors,  $F_3$  assures higher accuracy than  $F_2$ .  $F_3$  overtakes  $F_2$  because probably the influence of the excluded factors is negligible. On the contrary, the exclusion of significant factors, as all the climatic ones, causes a sensible reduction of the model precision, as proved by the poor performance of net  $F_1$ .

In general, the results confirmed the input variable choice generally conditions the final quality of the model on the same dataset. In detail, the factor selection requires relevant accuracy, in compliance with the following suggestions:

- selecting factors likely conditioning the mechanical performance of the pavements increases the result accuracy significantly. This outcome has been proved in two ways. Considering  $F_1$ , it is possible to assess that excluding very relevant factors (such as

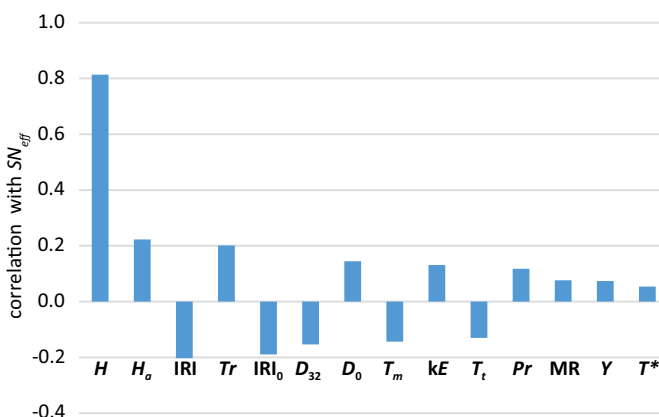


Figure 4. Correlation values among the various factors and  $SN_{eff}$

climatic ones) drastically reduces the numerical precision ( $R^2$  for  $F_1$  is 6% lower than for R). On the contrary, introducing strongly influencing parameters (MR and  $Pr$ ) in nets  $F_2$  and  $F_3$  caused a very significant improving (MSE for  $F_2$  and  $F_3$  is almost half than for R);

- a simple increase in the factor number is sometimes unproductive and disadvantageous. The comparison between  $F_2$  and  $F_3$  evidences that, although the simplistic introduction of the novel parameters improves the correlation accuracy (compared to R), a smarter definition of the input variable group is more reliable. Even if  $F_3$  results are only slightly better than  $F_2$ , the  $F_3$  architecture results optimized, involving fewer parameters, reducing computational and dataset management issues.

For more clarity, these factors simply represent a preliminary selection for investigating network behaviour and architecture. They are only some of the parameters influencing the pavement structural performance, and they can be further changed and improved for increasing quality. However, they have been selected as a starting point, but future researches are expected to improve the method including in the analysis other significant parameters, especially performance indicators (as cracking, rutting), with expected positive results.

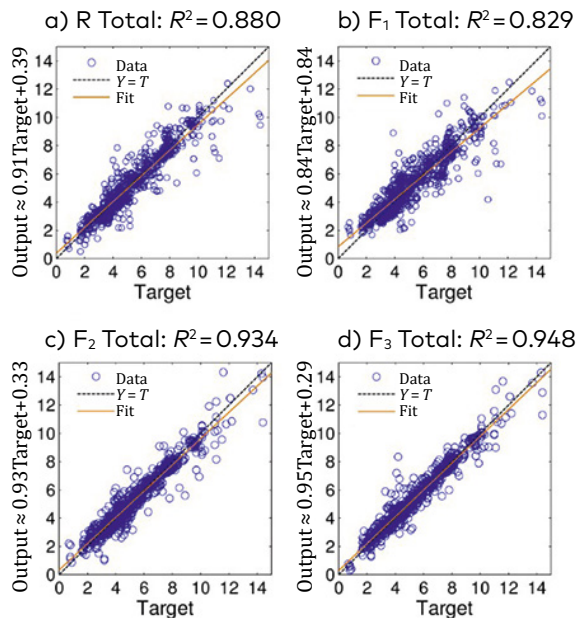
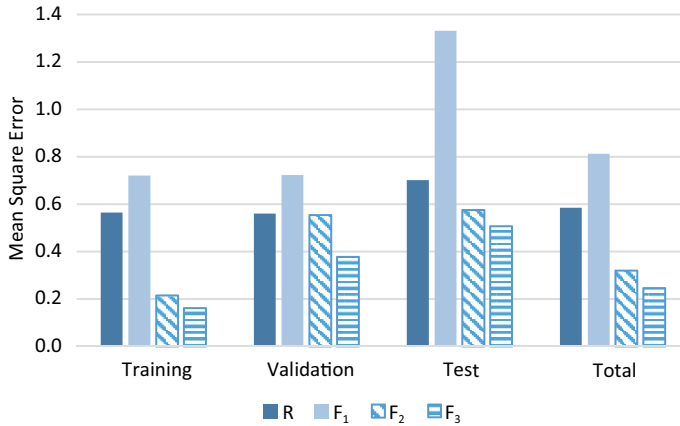


Figure 5. Regression charts for R,  $F_1$ ,  $F_2$ , and  $F_3$



**Figure 6.** Mean Square Error for R, F<sub>1</sub>, F<sub>2</sub>, F<sub>3</sub>

### 3.2. Record homogeneity influence

The ANN methodology represents a big-data approach and, thus, the resulting quality is strictly related to the quality of the initial dataset, mainly because of the lack of physical links with some of the involved phenomena (in particular, when numerous and different parameters are evaluated together). Obviously, the final accuracy depends on the data distribution in the solution space and on possible trends hidden in the records. It is possible to believe the higher the homogeneity in the database, the more reliable the quality of the correlation. In this context, homogeneity also means similarity (regarding external conditions influencing the structural performance) with the real road network on which the method is applied. To prove this assumption and check the effectiveness of different “homogenization” approaches, other specific networks, trained using only some of the available records, were compared to the reference net, R. These configurations were chosen in compliance with the following seven criterions:

1. R – the general database including all the 1021 records;
2.  $T_m$  between 15 °C and 25 °C; only 381 records of the R dataset;
3.  $T_m$  between 15 °C and 20 °C (246 records);
4.  $kE$  over 100 (386 records);
5.  $kE$  over 180 (153 records);
6.  $H$  between 15 cm and 25 cm (365 records);
7.  $H$  between 18 cm and 23 cm (202 records);

Excluding the reference selection R, it is possible to classify and compare to the other selections according to the factor adopted for homogenization ( $T_m$ ,  $kE$ ,  $H$ ) and for the number of records (around

Table 1. Mean value of the parameters included in each network

Network	<i>H</i>	<i>H<sub>α</sub></i>	<i>Tr</i>	<i>kE</i>	<i>T<sub>m</sub></i>	<i>D<sub>32</sub></i>	<i>D<sub>0</sub></i>	<i>IRI<sub>0</sub></i>	<i>IRI</i>	<i>T<sub>t</sub></i>	<i>MR</i>	<i>Pr</i>
R	21.53	5.65	402.89	144.45	12.18	40.99	110.30	1.24	1.46	21.40	32449.18	876.85
<i>T<sub>m</sub></i> 15–25	17.99	4.37	448.69	150.22	18.53	80.36	41.78	1.19	1.42	26.07	34425.97	984.82
<i>T<sub>m</sub></i> 15–20	14.33	4.29	420.61	140.28	16.82	67.26	58.11	1.17	1.41	24.71	32506.70	1045.80
<i>kE</i> >100	23.79	6.48	732.26	301.52	11.92	39.31	112.30	1.33	1.55	20.97	33469.86	941.41
<i>kE</i> >180	19.63	5.95	1150.95	537.54	13.75	44.60	96.26	1.45	1.65	24.93	38973.49	1016.95
<i>H</i> 15–25	19.69	6.41	446.58	195.06	11.35	31.74	113.64	1.30	1.56	20.54	37793.26	944.82
<i>H</i> 18–23	20.15	5.81	400.61	138.00	11.03	32.14	112.71	1.33	1.61	19.31	35732.21	909.70

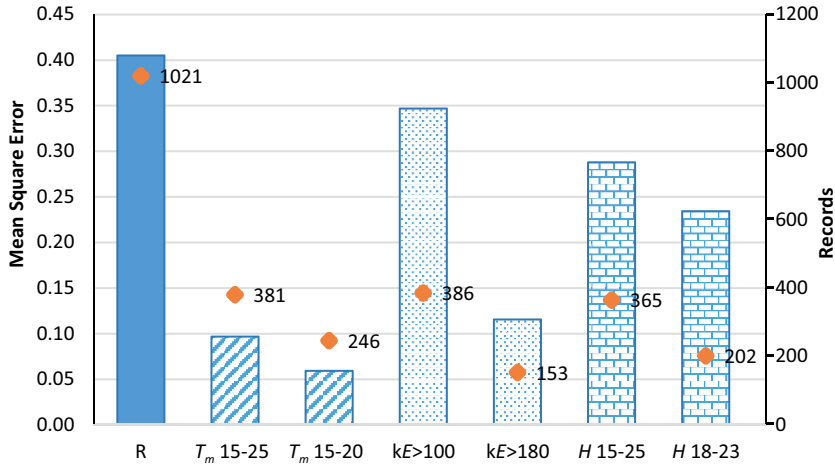
380 for selections 2, 4, 6 and around the half for selections 1, 3, 5). Details on the mean value and standard deviation for the different parameters for each selection are represented in Tables 1 and 2, respectively.

The results, provided in Figure 7 regarding MSE, represent the best performances of the different selections: similar patterns represent similar homogeneity criteria; moreover, for each network, the number of records is reported in the graph.

The MSE values show that selecting similar records in the datasets is generally productive and efficient for increasing the final accuracy: all the six selections show lower MSE values than R, despite the lower record number. If similar scenarios and sections are considered, the differences in the unknown behaviours are probably reduced and, thus, it becomes easier for the ANN to determine the numerical relationships between input and output variables. However, homogeneity effects vary as a function of the homogenization factor: according to the MSE values, the temperature seemed the most efficient homogenization factor (among the analysed ones). It is interesting to notice that, although the total thickness is stronger correlated to  $SN_{eff}$  (Figure 4) than  $T_m$  – and it

Table 2. The standard deviation of the parameters included in each network

Network	<i>H</i>	<i>H<sub>α</sub></i>	<i>Tr</i>	<i>kE</i>	<i>T<sub>m</sub></i>	<i>D<sub>32</sub></i>	<i>D<sub>0</sub></i>	<i>IRI<sub>0</sub></i>	<i>IRI</i>	<i>T<sub>t</sub></i>	<i>MR</i>	<i>Pr</i>
R	9.48	3.15	567.80	287.08	5.95	39.44	63.94	0.45	0.57	13.14	20183.53	404.30
<i>T<sub>m</sub></i> 15–25	7.59	2.82	744.59	343.00	2.64	31.53	29.13	0.43	0.53	11.58	25552.03	377.44
<i>T<sub>m</sub></i> 15–20	6.12	2.91	728.13	310.44	1.39	23.49	22.96	0.46	0.55	11.72	16165.03	415.13
<i>kE</i> >100	10.30	3.53	813.95	421.57	6.64	41.09	68.73	0.47	0.57	13.20	23448.27	359.03
<i>kE</i> >180	10.37	4.30	1159.21	597.30	5.31	31.91	54.69	0.56	0.66	12.31	32850.73	373.25
<i>H</i> 15–25	3.07	3.66	571.53	355.53	6.20	36.73	65.29	0.44	0.60	13.24	20980.08	428.34
<i>H</i> 18–23	1.79	3.28	584.94	263.54	6.32	39.61	66.98	0.49	0.63	12.60	17423.63	380.34



**Figure 7.** Mean Square Error for different homogeneous groups

is expected to have a higher influence in similarity selection-, reaction to total thickness is around 3 times worse than  $T_m$ . Then, the choice is difficult.

Further, beyond the homogeneity criterion, by comparing bars with the same pattern in Figure 7, it seems the higher the homogeneity level, the higher the final accuracy. Indeed, selections 3, 5, and 7 performed better than 2, 4, and 6, respectively. Although they were trained with a significantly smaller number of records (reduction varies from 35% to 60%). Naturally, attention should be paid on the record number, avoiding much-reduced databases likely causing improper overfitting and becoming useless.

However, even if these outcomes are remarkable, it is interesting to consider in the next studies a different and probably more performing homogenization approach, based on clustering techniques. Indeed, this solution generally represents a very efficient way for grouping different records according to their similarity in a multi-dimensional space (Amadore, Bosurgi, Pellegrino, & Sollazzo, 2016; Bosurgi, Carbone, Pellegrino, & Sollazzo, 2017) and, thus, it is probably beneficial in such a context.

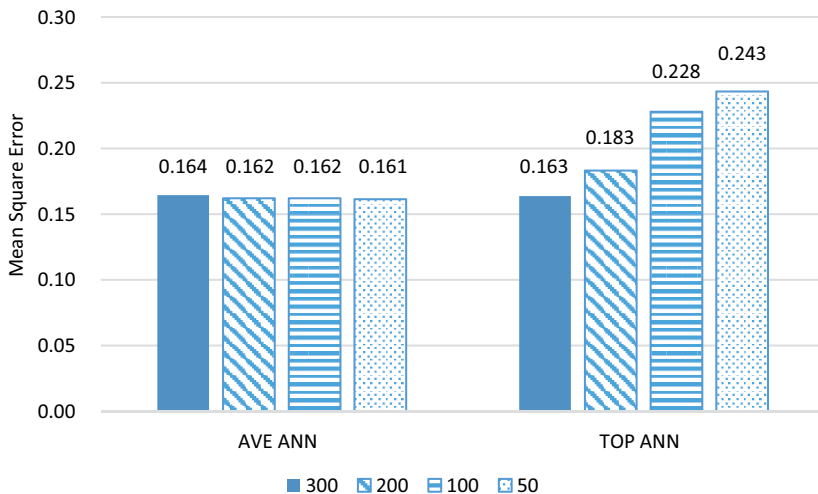
### 3.3. Committees of Artificial Neural Network

As introduced in Section 2, considering the combination of several ANNs is expected to significantly increase the accuracy of the correlation. According to the ANN nature and calculation, it is evident

that the result is strongly influenced by how the records are assigned to the three different pools. This distribution was performed randomly and, thus, different attempts produced different networks, characterised by different performances. In the previous experiments, for each architecture, the authors listed only the results of “the best ANN”, selected among several “randomly-identical” nets.

Then, it is interesting to focus on this aspect and verify whether the combination of multiple ANNs increases result accuracy. At this regard, the authors combined different “randomly-identical” nets by averaging their results. This operation is supposed to increase the correlation quality by averaging the specific inaccuracies of the different nets and reduce the final errors in the results.

For reducing the computational issues of these tests, a reduced dataset made up of 634 records has been adopted, and the reference architecture has been the  $F_3$  net described in Section 4.1, as it showed the best performance. Three hundred identical – regarding architecture – nets were trained in parallel, by randomly assigning the available records in the training, validation, and test pools. These nets were then combined in committees of different sizes, for evaluating the influence of the number of “averaged” nets. In detail, four specific committees were defined, by selecting 1/6, 1/3, 2/3 and all of the 300-trained nets (i.e., 50 nets, 100 nets, 200 nets, and 300 nets,

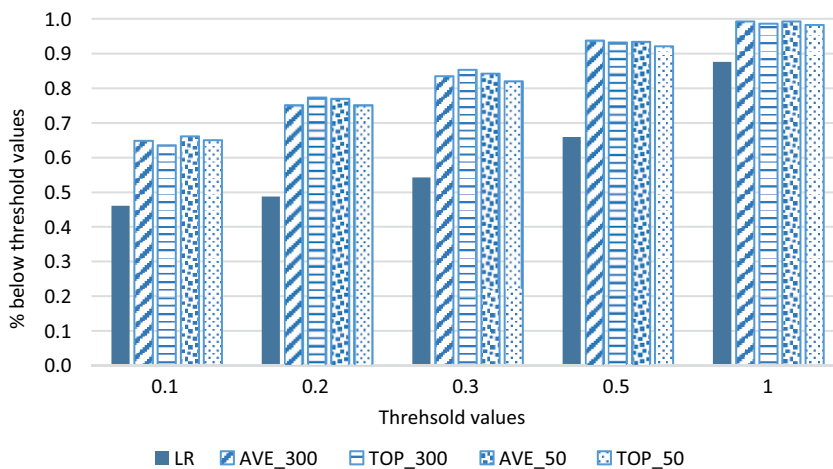


**Figure 8.** Mean Square Error for the average of all the net results and the best net for different committee size

respectively). For each group, the best net (TOP ANN, i.e., the network with the best performance in the first 50 nets, 100 nets, 200 nets, and 300 nets, respectively) and the average of all the net results (AVE ANN, i.e., the average of the first 50 nets, 100 nets, 200 nets, and 300 nets, respectively) have been considered. The related results, expressed in terms of MSE, are presented in Figure 8.

The outcomes are interesting: as expected, averaging the net results is always very effective. The average provides better results than the single best ANN for all the groups. This result proves the risks related to relying on a single net, even if very “promising”. Moreover, despite the relevant differences in committee sizes, the average results are very similar for the four groups. Then, an average of 50 nets is enough for increasing the correlation accuracy, avoiding the higher computational issues caused by bigger committee sizes. For more clarity, in Figure 9, the groups of 50 nets and 300 nets are also compared to Linear Regression (LR), regarding the percentage of errors below some threshold values (0.1, 0.2, 0.3, 0.5, and 1.0).

The chart further proves the higher accuracy of ANN compared to LR for all thresholds. The ANN approaches assure 95% of errors less than 0.5, while LR provides around 65% of errors at the same threshold. Furthermore, considering LR, over 10% of errors are more than 1. Finally, the average provides better results than the TOP nets in both cases, and this graph further confirms the quality invariance to the committee size.

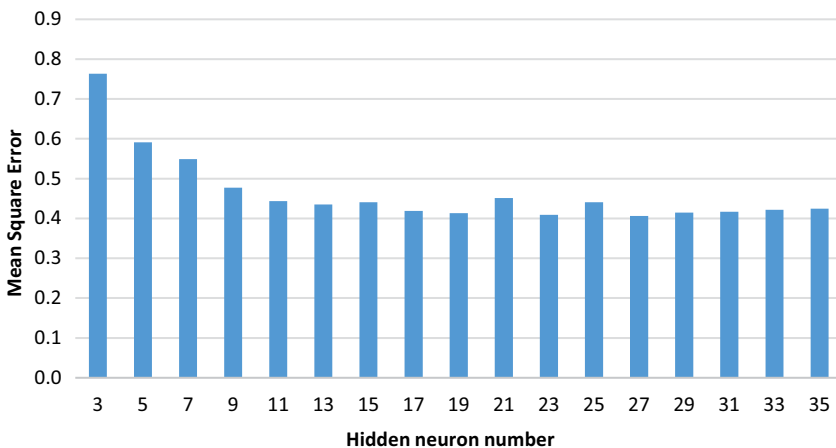


**Figure 9.** Committees of 50 nets and 300 nets compared to Linear Regression using error threshold values



### 3.4. Hidden neuron number

Finally, the number of hidden neurons is another parameter to consider for optimising the ANN architecture. In literature, there are suggestions for adopting a number of hidden neurons around 1 or 2 times the input number, and advice to limit them to twice the inputs (Swingler, 1996) for avoiding overfitting. The previous tests were performed considering 25 hidden neurons, according to quick preliminary “trial-and-error” experiments. However, since further investigating its influence in the correlation accuracy seems to be helpful, the authors tested systematically the  $F_3$  net presented in Section 4.1, considering different possible configurations for the hidden layer (varying from 3 neurons to 35 neurons). In detail, 50 nets have been trained for each number of hidden neurons. Figure 10 shows the MSE of the best net obtained for each hidden neuron number. The numerical results evidence it is preferable to adopt more than 10 hidden neurons in similar applications. Beyond this threshold, the net performance is only slightly conditioned by this variable, as differences in MSE are very small (around 0.04). Consequently, the adoption of 25 hidden neurons in the previous experiments appears as an acceptable preliminary choice.



**Figure 10.** Influence of hidden neuron number on  $F_3$  net performance

## Conclusions

In this paper, the authors have examined and analysed an Artificial Neural Network for structural performance estimation in asphalt pavements, aiming to optimise its configuration and some influencing factors for maximising the resulting quality and the estimation reliability. In detail, several aspects conditioning the Artificial Neural Network accuracy have been considered, to derive some indications for proper network optimisation. The considered nets are based on a large data set from the Long Term Pavement Performance database, including some of the parameters influencing the pavement structural performance, such as structural, climatic, traffic, and performance variables. The numerical tests evidenced how the factor selection, the record homogeneity and the hidden neuron number modify the Artificial Neural Network final accuracy, evidencing possible advantages and drawbacks of different configurations. In detail, the outcomes proved the following issues:

- the input variables are strategic for the final accuracy; in detail, it appeared that an arbitrary increase of the factor numbers has a reduced direct influence, but the network is optimised when factors actually conditioning the mechanical performance of the pavement are selected;
- if the available dataset is homogeneous regarding variability and ranges of the selected factors, the performance significantly grows; however, some factors showed higher influence than others on the accuracy regarding homogeneity and, thus, the dataset selection for network training requires a proper design;
- the number of hidden neurons represents a strategic factor too: indeed, experiments showed above 10 neurons must be adopted in similar applications (better around 1–2 times the input number), but further tests are required exhaustively relate to this number to the dataset features.

Moreover, the numerical results proved that averaging the results of parallel Artificial Neural Networks assures better results since this solution assures a more accurate coverage of the solution space. Results and discussions appear useful for optimising the Artificial Neural Network architecture for structural performance evaluation in practical applications. It is suggested to spend further research efforts for including in the dataset rutting measures and distress ratings, collected through the innovative, high-performance survey devices, for further exploiting the methodology potential. Furthermore, clustering techniques are expected to improve the dataset definition, as they are helpful in simplifying the identification of similar records in this kind

of database. However, the numerical results confirm the effectiveness of such a methodology for actually improving Pavement Management Systems and reducing deflection survey frequency, producing remarkable savings for the road agencies.

## REFERENCES

- Adeli, H. (2001). Neural networks in civil engineering: 1989–2000. *Computer-Aided Civil and Infrastructure Engineering*, 16(2), 126-142.  
<https://doi.org/10.1111/0885-9507.00219>
- Agarwal, P. K., Das, A., & Chakroborty, P. (2006). Simple model for structural evaluation of asphalt concrete pavements at the network level. *Journal of infrastructure systems*, 12(1), 41-49.  
[https://doi.org/10.1061/\(ASCE\)1076-0342\(2006\)12:1\(41\)](https://doi.org/10.1061/(ASCE)1076-0342(2006)12:1(41))
- Alexandre, L. A., Campilho, A. C., & Kamel, M. (2000). Combining independent and unbiased classifiers using weighted average. In *Pattern Recognition, 2000. Proceedings. 15th International Conference on* (Vol. 2, pp. 495-498). IEEE. <https://doi.org/10.1109/ICPR.2000.906120>
- Amadore, A., Bosurgi, G., Pellegrino, O., & Sollazzo, G. (2016). Compaction of Open-Graded HMAs Evaluated by a Fuzzy Clustering Technique. In *8th RILEM International Symposium on Testing and Characterization of Sustainable and Innovative Bituminous Materials* (pp. 243-254). Springer, Dordrecht.  
[https://doi.org/10.1007/978-94-017-7342-3\\_20](https://doi.org/10.1007/978-94-017-7342-3_20)
- Attoh-Okine, N. O. (1994). Predicting roughness progression in flexible pavements using artificial neural networks. In *Transportation Research Board Conference Proceedings* (Vol. 1, No. 1).
- Bianchini, A., & Bandini, P. (2010). Prediction of pavement performance through neuro-fuzzy reasoning. *Computer-Aided Civil and Infrastructure Engineering*, 25(1), 39-54. <https://doi.org/10.1111/j.1467-8667.2009.00615.x>
- Bosurgi, G., & Trifirò, F. (2005). A model based on artificial neural networks and genetic algorithms for pavement maintenance management. *International Journal of Pavement Engineering*, 6(3), 201-209.  
<https://doi.org/10.1080/10298430500195432>
- Bosurgi, G., Carbone, F., Pellegrino, O., & Sollazzo, G. (2017). Time Reduction for Completion of a Civil Engineering Construction Using Fuzzy Clustering Techniques. *Periodica Polytechnica Transportation Engineering*, 45(1), 25-34.  
<https://doi.org/10.3311/PPtr.9810>
- Bosurgi, G., D'Andrea, A., & Pellegrino, O. (2013). What variables affect to a greater extent the driver's vision while driving?. *Transport*, 28(4), 331-340.  
<https://doi.org/10.3846/16484142.2013.864329>
- Ceylan, H., Bayrak, M. B., & Gopalakrishnan, K. (2014). Neural networks applications in pavement engineering: A recent survey. *International Journal of Pavement Research and Technology*, 7(6), 434-444.  
[https://doi.org/10.6135%2fijprt.org.tw%2f2014.7\(6\).434](https://doi.org/10.6135%2fijprt.org.tw%2f2014.7(6).434)

- Doughty, M. S. (1997). Applications of neural network in transportation. *Transportation Research Part C: Emerging Technologies*, 5(5), 255-257. [https://doi.org/10.1016/S0968-090X\(97\)00013-2](https://doi.org/10.1016/S0968-090X(97)00013-2)
- Eldin, N. N., & Senouci, A. B. (1995). A pavement condition-rating model using backpropagation neural networks. *Computer-Aided Civil and Infrastructure Engineering*, 10(6), 433-441. <https://doi.org/10.1111/j.1467-8667.1995.tb00303.x>
- Elseifi, M., Abdel-Khalek, A. M., & Dasari, K. (2012). Implementation of rolling wheel deflectometer (RWD) in PMS and pavement preservation. *Report FHWA/11.492, Louisiana Department of Transportation and Development*.
- Flood, I., & Kartam, N. (1994). Neural networks in civil engineering. I: Principles and understanding. *Journal of computing in civil engineering*, 8(2), 131-148. [https://doi.org/10.1061/\(ASCE\)0887-3801\(1994\)8:2\(131\)](https://doi.org/10.1061/(ASCE)0887-3801(1994)8:2(131))
- Fwa, T. F., & Chan, W. T. (1993). Priority rating of highway maintenance needs by neural networks. *Journal of Transportation Engineering*, 119(3), 419-432. [https://doi.org/10.1061/\(ASCE\)0733-947X\(1993\)119:3\(419\)](https://doi.org/10.1061/(ASCE)0733-947X(1993)119:3(419))
- Graupe, D. (2013). *Principles of artificial neural networks* (Vol. 7). World Scientific.
- Hashem, S. (1997). Optimal linear combinations of neural networks. *Neural networks*, 10(4), 599-614. [https://doi.org/10.1016/S0893-6080\(96\)00098-6](https://doi.org/10.1016/S0893-6080(96)00098-6)
- Hashem, S., & Schmeiser, B. (1995). Improving model accuracy using optimal linear combinations of trained neural networks. *IEEE Transactions on neural networks*, 6(3), 792-794. <https://doi.org/10.1109/72.377990>
- He, J., Qi, Z., Hang, W., Zhao, C., & King, M. (2014). *Predicting freeway pavement construction cost using a Back-Propagation neural network: A case study in Henan, China*. *Baltic Journal of Road and Bridge Engineering*, 9 (1), 66-76. <https://doi.org/10.3846/bjrbe.2014.09>
- Kuncheva, L. I. (2002). A theoretical study on six classifier fusion strategies. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, (2), 281-286. <https://doi.org/10.1109/34.982906>
- Park, K., Thomas, N. E., & Wayne Lee, K. (2007). Applicability of the international roughness index as a predictor of asphalt pavement condition. *Journal of Transportation Engineering*, 133(12), 706-709. [https://doi.org/10.1061/\(ASCE\)0733-947X\(2007\)133:12\(706\)](https://doi.org/10.1061/(ASCE)0733-947X(2007)133:12(706))
- Plati, C., Georgiou, P., & Papavasiliou, V. (2016). Simulating pavement structural condition using artificial neural networks. *Structure and Infrastructure Engineering*, 12(9), 1127-1136. <https://doi.org/10.1080/15732479.2015.1086384>
- Pozarycki, A. (2015). *Pavement diagnosis accuracy with controlled application of artificial neural network*. *Baltic Journal of Road and Bridge Engineering*, 10 (4), 355-364. <https://doi.org/10.3846/bjrbe.2015.45>
- Priddy, K. L., & Keller, P. E. (2005). *Artificial neural networks: an introduction* (Vol. 68). SPIE press.
- Rada, G. R., Perera, R., & Prabhakar, V. (2012). *Relating Ride Quality and Structural Adequacy for Pavement Rehabilitation/Design Decisions* (No. FHWA-HRT-12-035).

- Rakesh, N., Jain, A. K., Reddy, M. A., & Reddy, K. S. (2006). Artificial neural networks—genetic algorithm based model for backcalculation of pavement layer moduli. *International Journal of Pavement Engineering*, 7(3), 221-230. <https://doi.org/10.1080/10298430500495113>
- Roberts, C. A., & Attoh-Okine, N. O. (1998). A comparative analysis of two artificial neural networks using pavement performance prediction. *Computer-Aided Civil and Infrastructure Engineering*, 13(5), 339-348. <https://doi.org/10.1111/0885-9507.00112>
- Rohani, A., Abbaspour-Fard, M. H., & Abdolahpour, S. (2011). Prediction of tractor repair and maintenance costs using Artificial Neural Network. *Expert Systems with Applications*, 38(7), 8999-9007. <https://doi.org/10.1016/j.eswa.2011.01.118>
- Shekharan, A. (1999). Assessment of relative contribution of input variables to pavement performance prediction by artificial neural networks. *Transportation Research Record: Journal of the Transportation Research Board*, (1655), 35-41. <https://doi.org/10.3141/1655-06>
- Sollazzo, G., Fwa, T. F., & Bosurgi, G. (2017). An ANN model to correlate roughness and structural performance in asphalt pavements. *Construction and Building Materials*, 134, 684-693. <https://doi.org/10.1016/j.conbuildmat.2016.12.186>
- Sollazzo, G., Wang, K. C. P., Bosurgi, G., & Li, J. Q. (2016). Hybrid procedure for automated detection of cracking with 3D pavement data. *Journal of Computing in Civil Engineering*, 30(6), 04016032. [https://doi.org/10.1061/\(ASCE\)CP.1943-5487.0000597](https://doi.org/10.1061/(ASCE)CP.1943-5487.0000597)
- Swingler, K. (1996). *Applying neural networks: a practical guide*. Morgan Kaufmann.
- Terzi, S. (2007). Modeling the pavement serviceability ratio of flexible highway pavements by artificial neural networks. *Construction and Building Materials*, 21(3), 590-593. <https://doi.org/10.1016/j.conbuildmat.2005.11.001>
- Tosti, F., Ciampoli, L. B., D'Amico, F., Alani, A. M., & Benedetto, A. (2018). An experimental-based model for the assessment of the mechanical properties of road pavements using ground-penetrating radar. *Construction and Building Materials*, 165, 966-974. <https://doi.org/10.1016/j.conbuildmat.2018.01.179>
- Transportation Officials. (1993). *AASHTO guide for design of pavement structures, 1993* (Vol. 1). AASHTO.
- Wang, K. C. (2011). Elements of automated survey of pavements and a 3D methodology. *Journal of Modern Transportation*, 19(1), 51-57. <https://doi.org/10.1007/BF03325740>
- Wang, K. C., & Li, Q. (2011). Pavement smoothness prediction based on fuzzy and gray theories. *Computer-Aided Civil and Infrastructure Engineering*, 26(1), 69-76. <https://doi.org/10.1111/j.1467-8667.2009.00639.x>
- Wu, Z., Zhang, Z., & Abadie, C. (2013). Determining structural strength of existing asphalt layer using condition survey data. *International Journal of Pavement Engineering*, 14(7), 603-611. <https://doi.org/10.1080/10298436.2012.677845>

- Yi, J.H., Kim, Y.S., Mun, S.H., and Kim, J.M. (2010). *Evaluation of structural integrity of asphalt pavement system from FWD test data considering modeling errors*. *Baltic Journal of Road and Bridge Engineering*, 5 (1), 10-18.  
<https://doi.org/10.3846/bjrbe.2010.02>
- Zhang, Z., Manuel, L., Damnjanovic, I., & Li, Z. (2003). Development of a new methodology for characterizing pavement structural condition for network-level applications. *Texas Dept. of Transportation, Austin, TX*.
- Ziari, H., Sobhani, J., Ayoubinejad, J., & Hartmann, T. (2016). Prediction of IRI in short and long terms for flexible pavements: ANN and GMDH methods. *International journal of pavement engineering*, 17(9), 776-788.  
<https://doi.org/10.1080/10298436.2015.1019498>