An ANN-based approach for non-destructive asphalt road density measurement

Muyang Li1, Loulin Huang2*, and Bryan Pidwerbesky3

1School of Engineering, Computer and Mathematical Sciences, Auckland University of Technology, 6 St Paul Street, Auckland 1010, New Zealand. Email: muyang.li@autuni.ac.nz

2School of Engineering, Computer and Mathematical Sciences, Auckland University of Technology, 6 St Paul Street, Auckland 1010, New Zealand. Email: loulin.huang@aut.ac.nz

3Fulton Hogan Ltd, 15 Sir William Pickering Drive, Christchurch 8053, New Zealand. Email:Bryan.Pidwerbesky@fultonhogan.com

* *Corresponding author*

ABSTRACT

Asphalt pavement's density measurement is an important step in the quality control of asphalt road construction. It is usually achieved by applying the coring method (CM), nuclear density gauge (NDG), and electromagnetic density gauge (EDG). CM is the most accurate method, but it is a destructive method since the pavement is damaged when the cores are taken. NDG and EDG are non-destructive methods with high efficiency, but their measurement accuracy is poorer than that of CM. An EDG commonly used in density measurement is named pavement quality indicator (PQI). A novel method named density profiling system (DPS) is also based on the potential EDG. However, it is not applied to this research since more tests are required to verify its accuracy. This paper presents an approach to improve the accuracy of the non-destructive methods with NDG and PQI. It is based on the artificial neural network (ANN), which processes the raw data got from NDG and PQI and produces the predicted asphalt density as the output. The density measured in CM method is used as the target density and the error between ANN-predicted density and target density

is computed. To minimize this error, various ANN architectures and learning algorithms are tried in the ANN training process. Each established ANN model makes a substantial improvement in the performance of NDG or PQI in asphalt density measurement.

PRACTICAL APPLICATIONS

This research was initiated by Fulton Hogan (FH) Limited, a large road construction and maintenance company in New Zealand. FH lab teams are responsible for asphalt road density measurement in FH's road projects. A main method they use is to measure the densities of the cores taken from asphalt pavements (coring method). It is quite accurate but destructive and very time-consuming. They also use a nuclear density gauge (NDG) or pavement quality indicator (PQI), which are highly efficient non-destructive measurement devices. However, their measurement accuracy is poorer than that of the coring method. FH lab teams wanted to have a new density measurement method which is both accurate and efficient.

An ANN-based approach is presented in this paper to address the above issues faced by the FH lab teams. Densities collected with coring methods, NDG and PQI are used to train and validate the ANN models. The results from the ANNs show substantial improvements of the measurement accuracy and efficiency. The proposed approach has been presented to the FH lab teams who are impressed with its performance and plan to implement it in their projects.

INTRODUCTION

In asphalt pavement construction, it is necessary to monitor its density as an important step of quality control. A straightforward and accurate method for asphalt density measurement is the coring method (CM) (Ameri et al. 2014). In this method, asphalt core samples are extracted from the asphalt pavement by coring equipment (AS/NZS 2008). They are then transferred to a lab to be tested (AS/NZS 2014). The test procedures include the measurement of a sample's dry mass, the mass of the sample coated with a thin wax film, and the mass of the coated sample immersed into water. The so-called core density can be calculated from the mentioned three masses. The single-operator standard deviation (SD) of CM is 7.98 kg/m^3 while the multi-laboratory SD is

14.96 kg/m^3 (ASTM 2019). A problem with this method is that only a limited number of cores can be taken from a pavement. It is to avoid making the Swiss-cheese-like surface which is weak and unattractive. This method is hence called a destructive method. Apart from that, taking and measuring core samples and recovering the pavement are quite time-consuming and expensive.

More efficient measurement methods are the so-called non-destructive methods using nuclear density gauge (NDG) or electromagnetic density gauge (EDG) (Ziari et al. 2010). The working principle of the NDG is based on Compton scattering (Brisset et al. 2005). When a γ ray penetrates and collides with the electrons in the asphalt mixture, parts of the γ photons may scatter. The density is estimated from the number of scattered photons (Brisset et al. 2005; Malpass and Khosla 2002). The precision of NDG is high. The single-operator SD is 25.15 kg/m^3 while the multi-laboratory SD is 28.03 kg/m^3 (ASTM 2022). However, it still cannot match that of the CM method mentioned above and has much room for improvement. In addition, nuclear radiation hazard is a big concern for its application.

As for EDG, a commonly used device is the pavement quality indicator (PQI) from TransTech Systems.(Ziari et al. 2010; Allen and Schultz 2003). The operational theory of PQI is based on the effects of density on the electromagnetic parameters, such as relative permittivity (dielectric constant) (Allen and Schultz 2003). A primary function of the PQI is to generate an electrical field in the asphalt pavement to manifest and measure those effects, which can be used to estimate the density through built-in algorithms (Ziari et al. 2010). According to the ASTM standard, the single-operator SD of EDG is 20.50 kg/m^3 while the multi-laboratory SD is 23.55 kg/m^3 (ASTM 2016). However, due to the numerous factors (e.g., compositions, air voids, temperature) affecting the electromagnetic parameters of asphalt pavement, the actual accuracy of PQI is poorer than NDG or CM in the field measurement.

An up-and-coming method named density profiling system (DPS) is proposed based on ground penetrating radar (GPR), a potential EDG (Hoegh et al. 2020; Leiva et al. 2022). GPR is a prevalent device used to determine the thickness of the asphalt layer. It holds a similar operational theory to PQI, and hence it has the potential to measure the density of the asphalt pavement. In DPS, 3 GPRs are mounted onto a cart to measure the density continuously rather than the density of a fixed location by PQI or NDG. However, as a novel method, more tests are required to verify its accuracy. Thus, DPS is not applied in this research.

A common feature of non-destructive methods is that asphalt pavement's density is estimated indirectly through operational theories or empirical formulas in physics or engineering to describe the relations between the density and the nuclear or electromagnetic signals picked up directly by them. Measurement errors tend to occur from the inherent deficiency of those principles or formulas to reflect the actual relations mentioned above. Noises and factors other than the density affecting the sensors' readings also compromise the measurement accuracy.

Taking NDG and PQI as sensors, data processing analysis techniques for general sensors can be applied to improve their accuracy. One of the promising techniques is the artificial neural network (ANN) which is commonly used as an effective tool for processing data when the relations among the data are not easy to be established with traditional closed-form mathematical models (Bashiri and Farshbaf-Geranmayeh 2011; Tian and Shang 2006; Puig-Arnavat and Bruno 2015; Khadse et al. 2017). To mimic the biological neural system in the human body for decision-making, ANN consists of various artificial neurons organized in multiple layers (Negnevitsky 2011). Every two neurons in different layers have an internal link with a unique numerical weight, leading to the output layers where the outputs are generated. The performance of an ANN model is affected by those weights. The process to minimize the difference between the forecast output and the target is named the training process. After the training process, the optimized ANN model can then be applied to produce the output corresponding to the data collected in real time.

In the training process of an ANN, the weights in ANN are tuned with the learning algorithms. Three commonly used algorithms are Levenberg-Marquardt (LM), Bayesian regularization (BR), and scaled conjugate gradient (SCG) algorithms (Hagan and Menhaj 1994; Smith et al. 2019; Foresee and Hagan 1997). LM algorithm provides very efficient training for an ANN model since the second derivative computation is replaced by the approximated first derivatives computation (Hagan and Menhaj 1994). BR algorithm can improve the generation capability of an ANN model, which can keep a consistence performance on different data sets (Foresee and Hagan 1997). Similar to LM algorithm, SCG algorithm can also train an ANN model efficiently by employing the first derivative computation. In addition, it uses the least computational memory than other methods (Møller 1993). The architecture of the ANN also influences its performance. For example, too many or too few neurons in the hidden layer of ANN is not recommended (Negnevitsky 2011).

This paper presents an ANN-based approach to improve the accuracy of the non-destructive methods (NDG and PQI) for measuring asphalt density. The readings from NDG and PQI are used as the raw data for training the ANN models. The LM, BR, and SCG algorithms are selected as learning algorithms to train the ANN models with 2 to 15 neurons in their hidden layers. The performances of the ANN models are verified by comparing them to the performances of NDG and PQI.

ASPHALT DENSITY MEASUREMENT METHODS

Coring method (CM)

CM is a destructive method for measuring asphalt density. It consists of two parts: core sample extraction and lab tests. Generally, at least one core sample should be extracted in each 300 m^2's area of asphalt pavement, and its location should be chosen randomly within the area (AS/NZS 2008). After being extracted, it is then tested in a lab to determine the core densities using the waxing procedure (AS/NZS 2014). The dry mass of the sample is weighed primarily. The sample is then coated with a thin film of prepared wax, and the mass of the coated sample is weighed. Finally, the coated sample is immersed in water, and the mass of coated sample in water is weighed by a specific balance. The accurate core density can be calculated by using the following equation:

$$\rho_{\text{core}} = \frac{m_1 \rho_{\text{w}} \rho_{\text{wax}}}{\rho_{\text{wax}}(m_2 - m_3) - \rho_{\text{w}}(m_2 - m_1)} \tag{1}$$

Where ρ_{core} is the core density of sample (kg/m^3), ρ_{wax} is the density of the paraffin wax (kg/m^3), ρ_{w} is the density of water (kg/m^3), m_1 is the mass in air of the dry sample (kg), m_2 is the mass in air of the wax coated sample (kg), m_3 is the mass in water of the wax coated sample (kg).

This method is considered the most accurate method for determining the asphalt pavement's density (Ameri et al. 2014). However, the obvious problem is that both two parts of this method are not efficient enough, especially for a large pavement project. Furthermore, additional workload and costs are required to recover the destroyed pavement. An inappropriate recovery of the asphalt pavement even causes a local weakness in the road and highly affects the quality of the pavement (Ziari et al. 2010).

Nuclear density gauge (NDG)

NDG is a standard non-destructive device for measuring the asphalt density onsite (AS/NZS 2013a; AS/NZS 2013b). The field test with an NDG device is considered the most accurate non-destructive method, though its accuracy is still lower than that of CM method.

The operational theory of a NDG can be explained by the Compton scattering describing the interaction between a matter and the applied γ (gamma) rays (Brisset et al. 2005). When these rays penetrate the asphalt mixture, the γ photons collide with electrons presented in the measured asphalt mixture. The collisions lead to the scattering of these γ photons. The number of scattered photons is related to the density of measured material (Malpass and Khosla 2002). According to Beer-Lambert's law, the density of the measured material can be approximately expressed by (Laufenberg 1986):

$$\rho = \frac{\mathrm{lnB} - \ln(I_\mathrm{t}/\mathrm{I}_0)}{\mathrm{kx}} \tag{2}$$

Where I_t is the count of transmitted photons per second after attenuation (counts/s), I_0 is the count of transmitted photons per second before attenuation (counts/s), k is mass attenuation coefficient (m^2/kg), ρ is the density of measured material (kg/m^3), x is the thickness of the measured material (m) and B is a preset build-up factor.

However, the density measured by NDG is still not accurate enough compared with that measured by CM (Ameri et al. 2014; Ziari et al. 2010). In addition, there are radiation hazards in NDG is a concern.

Pavement quality indicator (PQI)

PQI is another type of device for the non-destructive measurement of asphalt density. It can also measure the air void content and the surface temperature of the asphalt mixture (Ziari et al. 2010; Allen and Schultz 2003).

A PQI contains an active region, an isolation ring, and a ground region. It measures the complex permittivity through the mentioned structure (Ziari et al. 2010; Allen and Schultz 2003; Kvasnak et al. 2007). After applying an electrical field to pure material, its complex permittivity can be gained by Debye's equation, as follows:

$$\varepsilon_{\mathrm{r}} = \varepsilon_{\mathrm{r}}' - j\varepsilon_{\mathrm{r}}'' \tag{3}$$

$$\varepsilon_{\mathrm{r}}' = \varepsilon_{\infty} + \frac{\varepsilon_{\mathrm{s}} - \varepsilon_{\infty}}{1 + (\mathrm{f}/\mathrm{f_r})^2} \tag{4}$$

$$\varepsilon_{\mathrm{r}}'' = \frac{(\varepsilon_{\mathrm{s}} - \varepsilon_{\infty})\mathrm{f}/\mathrm{f_r}}{1 + (\mathrm{f}/\mathrm{f_r})^2} \tag{5}$$

Where ε_s is the permittivity as f goes to 0 (F/m), ε_∞ is the permittivity as the frequency goes to infinity (F/m), $\mathrm{f_r}$ is the relaxation frequency (Hz) and f is the applied frequency (Hz). The first three terms are the properties of the material. The real part of complex permittivity is the dielectric constant which can be processed to estimate density through the built-in algorithm. The imaginary part can be converted into moisture (Kvasnak et al. 2007).

However, the asphalt mixture contains aggregate, asphalt binder, and little air. The relative permittivity of air is far lower than aggregate or binder (Porubiaková and Komačka 2015). Thus, the air void content of the asphalt mixture highly affects its complex permittivity. By the preset equations in the PQI, the measured complex permittivity is transferred to the air void content and then the density. However, asphalt mixture may also contain a little water in a wet condition. This may influence the PQI measurement since the permittivity of water is varied according to its temperature.

The models for non-destructive density measurement by NDG or PQI do not reflect all the parameters affecting the density. Their measurement accuracies are compromised and cannot match the destructive method such as CM. To improve their performance, data processing techniques like ANN can be applied.

ARTIFICIAL NEURAL NETWORK (ANN)

Artificial neural network (ANN) is an efficient technique for data processing and is widely used in various applications (Bashiri and Farshbaf-Geranmayeh 2011; Tian and Shang 2006; Puig-Arnavat and Bruno 2015; Khadse et al. 2017). It has been attempted to support monitoring the quality of the asphalt pavement construction, and all applied ANN models show excellent performances (Chandrakasu et al. 2021; Commuri and Zaman 2008; Commuri et al. 2011). In this research, ANN technique is employed to improve the performance of NDG and PQI devices. The NDG-measured and PQI-measured densities are used to train the ANN models, and the ANN models output the so-called ANN-predicted densities which are then compared with the core densities. Their differences, named errors, are minimized by the ANN training process, which is guided by learning algorithms. The architecture of ANN models, training process, and learning algorithms used in this approach will be explained respectively.

Architecture

An artificial neural network is a kind of computation model inspired by the human nervous system (Negnevitsky 2011; Haykin 2000). Like the biological neural network consisting of natural neurons, ANN consists of highly interconnected units named artificial neurons, as shown in Figure 1. They are connected by links passing signals from one to another. Each neuron receives input signals and the attached numerical weights through its input links. The neuron computes the weighted sum of the input signals and subsequently uses the sum to calculate the final output by an activation function. Finally, the output signal will be transmitted to other neurons as an input signal until the final output signals of the entire ANN are reached.

In an ANN model, artificial neurons are organized into layers. In this approach, the architecture of the ANN model is illustrated in Figure 2. It consists of an input layer, a hidden layer, and an output layer. The input layer contains two neurons which are used to pass the input data (NDG-measured and PQI-measured densities) to the neurons in the hidden layer. After the computation process in each neuron, the outputs of the neurons in the hidden layer are passed to the neuron in the output layer. After the same process, the neuron in the output layer outputs the results, which are so-called ANN-predicted densities.

In this model, the numerical weights and the number of neurons in the hidden layer are the two essential parts required to be optimized. Their optimization processes will be introduced in the following sections.

Training Process and Learning Algorithms

The proposed ANN-based approach is illustrated schematically in Figure 3. In this research, the core densities are regarded as the target densities since they are the most accurate method to measure the density of the asphalt pavement. However, the limitation is obvious. The number of core samples that can be gained from a pavement is limited. The best way to solve this is to collect core densities from different projects containing pavements with similar conditions. The details of this part will be discussed in the next section, Data Collection.

As mentioned in the structure of ANN models, the NDG and PQI densities are input into the ANN models. After gaining the ANN-predicted densities, errors between them and the target densities are computed and are then used to tune the numerical weights of the ANN models. The process to search for the minimized errors and the optimized weights is named the training process.

Learning algorithms are employed to guide the search direction of this process. In this research, three learning algorithms are chosen, including the Levenberg-Marquardt algorithm, the Bayesian regularization algorithm, and the scaled conjugate gradient algorithm.

Based on the traditional Newton's method, the Levenberg-Marquardt (LM) algorithm is developed to reduce the time to train an ANN model (Hagan and Menhaj 1994). In Newton's method, the search direction is determined by calculating the Hessian matrix expressed by $\mathbf{H}_{ij} = \frac{\partial^2 \mathrm{F}(w)}{\partial w_i \partial w_j} (i = 1, ..., \mathrm{N})(i \neq j = 1, ..., \mathrm{N})$. w represents the weight, N represents the total number of weights and F represents the objective function to be minimized. The calculation of the second derivative requires a long time and a large storage space. These problems are solved by the LM algorithm where the objective function is considered as the sum of the squared error function, which can be described by $F(w) = \sum_{i=1}^{N} v_i^2(w)$. A Jacobian matrix is introduced and contains the elements $\mathbf{J}_{ij} = \frac{\partial v_i(w)}{\partial w_j}$. Accordingly, a Hessian matrix can be calculated approximately by $\mathbf{H} = 2\mathbf{J}^T\mathbf{J} + 2\sum_{i=1}^{N} v_i(w)\partial^2 v_i(w)$. The later part of $\mathbf{H}$ is commonly considered small enough and hence $\mathbf{H} \approx 2\mathbf{J}^T\mathbf{J}$. The Jacobian matrix only contains the first derivatives. Thus, compared with Newton's method, LM algorithm requires less time to train an ANN model.

The Bayesian regularization (BR) algorithm is proposed to improve the generalization capabilities of ANN models (Foresee and Hagan 1997). An ANN model with a good generalization capability means it performs nearly the same on both the data for training and other data. In other words, the ANN model is close to the true relationship between the input and output. With the assumption that the true relation is smooth, which corresponds to small weights in the ANN, the objective of the BR algorithm is to balance the minimization of the sum of the squared error and the sum of the weights in the ANN (Foresee and Hagan 1997; Hagan and Menhaj 1994). Hence, the BR algorithm requires more computation and is slower in training models than the LM algorithm.

Similar to the LM algorithm, the scaled conjugate gradient (SCG) algorithm is also proposed to train an ANN model efficiently (Møller 1993). In this algorithm, the so-called conjugate vectors, which are orthogonal to each other, are introduced. The search direction can be determined by calculating a set of conjugate vectors rather than the Hessian matrix and the Jacobin matrix. Although both a conjugate vector and the Jacobin matrix contain the first-order derivatives of the objective function, the former contains fewer elements than the latter. Thus, the SCG algorithm can train an ANN model more efficiently than LM algorithm does. In addition, less storage space is required during the training process.

The ANN models trained by the 3 learning algorithms are the so-called LM-ANN, BR-ANN, and SCG-ANN models, respectively. All models hold the same architecture mentioned above and will be used in the later optimization process.

DATA COLLECTION

To train the ANN model, a large number of data samples containing NDG, PQI, and core density are required. All data samples used in this research are collected from public data sources. The main advantage of using them is to save time, which may last up to several months. However, the disadvantage is that they may contain some factors influencing the accuracy of density measurement. According to the operation theory of NDG and PQI, the factors include the aggregate types and the layer thickness of the asphalt pavement. Thus, the ideal scenario for the data collection is that all the data samples are from one pavement or several pavements with similar or the same aggregate source and layer thickness. To meet the requirement, data samples from 2 public data sources are collected and then used to train the ANN models.

The same data collection process is used for them. Firstly, the number of test sections and the number of core samples that will be taken in each section are calculated. Secondly, locations for taking core samples are marked and numbered. Thirdly, the NDG is placed on each marked location to measure and record density. Fourthly, the PQI is used in the same way. Finally, after NDG and PQI measurements, all the core samples are taken and sent to a lab for measuring the core densities.

The first data source is the field investigation taken by engineers and technicians from the North Carolina Department of Transportation (Sawyer et al. 2005). The investigation field is located on an interstate highway (I-95) in Wilson County, North Carolina. The type of asphalt mix is the Superpave Mix S-12.5D, and hence the nominal size of the aggregate is 12.5 mm. The binder grade is the PG 76-22 binder. The layer thickness is around 38 mm. 16 test sections (8 lots and 2 sub-lots per lot) on the asphalt pavement are selected and the area of each one is around 110 m^2. 6 test locations per section are marked for conducting NDG and PQI measurements and then taking core samples. Thus, 96 data samples are collected from this data source.

The second one is the evaluation of the non-destructive method taken by the engineers and technicians from Kentucky Transportation Center (Allen and Schultz 2003). The evaluation field is located on another interstate highway (I-75) in Scott County, Kentucky. The same Superpave mix type (S-12.5D) and binder grade (PG 76-22) are used. 44 test sections (11 lots and 4 sub-lots per lot) on the asphalt pavements are selected. 3 or 4 locations per section are marked for conducting the same process outlined in the last data source. During the PQI measurement, 2 PQI devices are employed. Their average readings are calculated and recorded as the PQI data. In this data source, 143 data samples are collected.

Totally, 239 data samples are collected. Each data sample contains a core density, a PQI density, and an NDG density. The core density ranges from 2037 to 2485 kg/m^3. The NDG density ranges from 1892 to 2465 kg/m^3 while the PQI density ranges from 1980 to 2502 kg/m^3. The averages of the three densities are around 2311, 2271, and 2292 kg/m^3, respectively. Since the core density is the target data, an NDG or PQI density is compared with the corresponding core density. The root mean squared errors (RMSEs) of NDG and PQI densities are calculated to describe their accuracies. They are 65.53 and 68.00 kg/m^3, respectively. The correlations between them are described by the correlation coefficients (Rs). Figure 4 shows the data distributions and the Rs of NDG and PQI densities. The horizontal axis represents the core density. The vertical axis represents the NDG or PQI density. The Rs of NDG and PQI densities are around 0.83 and 0.87, respectively. Some outliers are clumped off the fit line, obviously. This is because the PQI and NDG measurements are inaccurate under high temperature or moisture. Data drifts may occur and lead to outliers under that condition. When many data are measured under the same condition, the outliers are then clumped off the fit line.

All the data samples are combined and become an original dataset for training the ANN models.

ANN MODEL'S OPTIMIZATION

An experiment is implemented to optimize the architecture of 3 ANN models. The numerical weights are optimized directly in the training process. The number of neurons can be optimized indirectly. In the experiment, it is preset from 2 to 15. (Negnevitsky 2011). According to the architecture of the ANN models shown in Figure 2, one neuron is not enough to handle the relations with two inputs. Hence the minimum number of neurons is 2. The maximum number of neurons is commonly set based on the performance of the models. This will be discussed later.

A model with each number of neurons is trained 100 times to get the average performance. The architecture of the ANN model with the best performance is the optimized one.

In this experiment, ANN models are created, trained, and tested by employing MATLAB, which is a computational software and contains a powerful ANN toolbox. The original dataset is divided into 3 data groups, which are training, validation, and testing data groups. The training data group is directly applied for training ANN models. The validation data group is used to prevent ANN models from over-fitting and stop the training process in time. The test dataset is not used in the training process and can be used to test the accuracy and the generation capability of the ANN model. In this experiment, 60% data belongs to the training group, 20% data belongs to the validation group, and the rest 20% data belongs to the test group.

As shown in Table 1-3, the experimental results illustrate the performances of the 3 ANN models with different numbers of neurons in their hidden layers. In this research, the performance of an ANN contains its accuracy, generation capability, and training efficiency, which will be analyzed respectively.

Accuracy

The accuracy of each model is measured by the root mean squared error (RMSE) of its testing data group. RMSE is calculated by the following equation:

$$RMSE = \sqrt{\frac{\Sigma_{i=1}^{\mathrm{n}}(X_i - Y_i)^2}{\mathrm{n}}} \tag{6}$$

where X and Y are the core density and ANN-predicted density, respectively. n is the total number of densities.

The maximum neuron of the hidden layer is discussed. Commonly, the more neurons an ANN model holds, the more computation process it needs. It also means the more training time and space for computational storage. Thus, it is not worth having too many neurons unless the error goes down obviously with the increasing number of neurons. As shown in Figure 5, when the number of neurons increases from 10 to 15, the RMSEs of BR-ANN and SCG-ANN model keep constantly and that of LM-ANN model even slightly goes up. Thus, it is not necessary to test a model with more than 15 neurons.

In addition, as shown in Figure 5, the BR-ANN model holds the least RMSE among all three models. Obviously, the BR-ANN model holds a low RMSE when there are 4 or 10 neurons in its hidden layer. The two RMSEs are 27.29 and 27.67 kg/m^3, respectively. They are very close. The LM-ANN model holds the second least RMSE. As the number of neurons goes up from 2 to 4, the RMSE of the LM-ANN model goes down. On the other hand, as the number of neurons goes up from 4 to 15, its RMSE goes up. Thus, an optimized architecture of the LM-ANN model is that it holds 4 neurons in its hidden layer. The SCG-ANN model holds the largest RMSE, and the SCG-ANN model with 8 neurons is an optimized architecture.

Generalization capability

The generation capability of an ANN model is measured through the difference between the RMSEs of the training, validation, and testing data groups. As shown in Figure 6, as the number of neurons goes up, the generation capability of LM-ANN model becomes worse and worse. To gain a reasonable generation capability, the number of neurons in the hidden layer should be less than 5. As shown in Figure 7, the BR-ANN model illustrates an excellent generation capability since the RMSEs of all data groups are close to each other. The number of neurons in its hidden layer does not affect its generation capability. As shown in Figure 8, the SCG-ANN model also shows an excellent generation capability. When the number of neurons is more than 10, the difference between the RMSEs of the training group and the validation group becomes larger. Hence, in terms of the SCG-ANN model, the number of neurons in its hidden layer should be less than 10.

Training efficiency

The training efficiency of a learning algorithm is measured by the length of time for training an ANN model. As shown in Figure 9, The SCG algorithm is the most efficient for training the ANN model. The training time of SCG-ANN model is stable, which means it is not influenced by the number of neurons. The LM algorithm shows a similar training efficiency to the SCG algorithm when there are more than 3 neurons in the hidden layer of the trained model. The BR algorithm requires the longest time to train an ANN model. The training time becomes longer and longer when the number of neurons goes up. Thus, a BR-ANN model with 4 neurons in its hidden layer can be trained more efficiently than the one with 10 neurons in its hidden layer.

By comparing all three aspects of their performances, the optimized architectures of the 3 models are the LM-ANN model with 4 neurons, the BR-ANN model with 4 neurons, and the SCG-ANN model with 8 neurons in their hidden layers respectively.

ANN MODEL'S EVALUATION AND DISCUSSION

RMSE and the precision of measurement

There is no direct relationship between RMSE and the ASTM/AASHTO reported precision. According to the ASTM and AASHTO standards, the precision of NDG or PQI measurement is indicated by the standard deviation (SD) (ASTM 2022; ASTM 2016; AASHTO 2022). It measures the average variability of a kind of density. On the other hand, RMSE measures the difference between core and other densities. However, RMSE and SD have an indirect relation in this research. A lower RMSE means that the corresponding density is closer to the core density. It also means that the SD of the corresponding density becomes lower as the core density holds the lowest SD. Thus, if the ANN-predicted density holds a lower RMSE than that of NDG or PQI density, it means that the ANN model produces a density with a higher precision than the one measured by NDG or PQI.

As can be seen in Table 4, The RMSEs of the ANN-predicted densities are around 30 kg/m^3. On the contrary, the RMSEs of the NDG and PQI densities are both more than 60 kg/m^3, which is far more than the RMSEs of the ANN-predicted densities. Thus, the precision of the NDG or PQI density is improved greatly by employing the ANN models to predict the density.

Correlation coefficient (R)

Correlation coefficients (Rs) are employed to indicate the closeness of the ANN-predicted densities to the core densities. The following equation calculates R:

$$R = \frac{\Sigma_{i=1}^{\mathrm{n}}(X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\Sigma_{i=1}^{\mathrm{n}}(X_i - \bar{X})^2(Y_i - \bar{Y})^2}} \tag{7}$$

where X_i and Y_i are the core density and ANN-predicted density, respectively. $\bar{X}$ and $\bar{Y}$ are the average values of X_i and Y_i, respectively. n is the total number of densities.

As mentioned in the previous section, Figure 4 shows the original performance of NDG and PQI without the assistance of the ANN models. The Rs of NDG and PQI densities are around 0.83 and 0.87, respectively.

Figure 10 - 12 show the performance of the 3 ANN models with their optimized architectures. Each figure contains four graphs representing the performances of the training, validation, and testing data groups and the overall performance, respectively. The LM-ANN model holds the best overall performance ($R \approx 0.94$) among all the models. However, the BR-ANN model holds the best performance of its testing data group ($R \approx 0.94$). In addition, the BR-ANN model has the best generation capability since its 3 data groups hold a similar R (around 0.94, 0.95, and 0.94, respectively). The generation capability of the SCG-ANN model is also reasonable, although it holds the worst performance ($R \approx 0.92$).

As for the data distribution associated with the LM-ANN and BR-ANN models, there are not too many clumped outliers and most of the data points are close to the fit line, which leads to high R values. Thus, the models are close to the true relation. On the other hand, the SCG-ANN model produces some clumped outliers. The reason can be that the search direction of the SCG algorithm requires conjugate vectors, which contain less information than the Jacobian matrix required by that of the LM and BR algorithms. Hence, the relation searched by the SCG algorithm is commonly worse than that of the LM or BR algorithm.

Since each of the ANN models has a higher R-value than that of NDG or PQI, the ANN models are verified to improve the performance of NDG and PQI greatly. However, some outliers still exist. The reason may be that some input densities contain large drifts or errors and many of them are allocated to the training groups.

CONCLUSION

This paper presents an ANN-based approach to improve the accuracy of asphalt density measurement with non-destructive methods such as NDG and PQI. The NDG-measured and PQI-measured densities are used as the inputs of the ANN models while the core densities are used as the target (desired) densities. Three learning algorithms, LM, BR, and SCG algorithms are applied to train the ANN models, and their performances are compared and the ANN architectures are optimized accordingly. It is shown that the proposed approach improves the accuracy of the non-destructive methods significantly. The next step is to collect additional data to train the proposed models continuously. The density from the pavements with a similar pavement thickness is helpful in further improving the accuracy of the data. The additional data can be directly combined with the current data samples since more data improves the ability of ANN models to filter noise and the outlier. Data containing temperature and moisture may also be helpful. The density from the pavement with a thinner or thicker layer may be helpful to improve the generation capabilities of the models.

DATA AVAILABILITY STATEMENT

All the data that support the findings of this study are from public data sources. All models, or code that support the findings of this study are available from the corresponding author upon reasonable request.

ACKNOWLEDGMENTS

The work presented in the paper was completed with the financial support of the Callaghan Innovation R&D Fellowship Grant of New Zealand.

REFERENCES

AASHTO (2022). “Standard method of test for in-place density of asphalt mixtures by nuclear methods.” *AASHTO T355*, American Association of State Highway and Transportation Officials (AASHTO).

Allen, D. and Schultz, D. (2003). “Evaluation of non-nuclear density gauges.” *Report No. KTC-03-24/FR115-01-1F*, College of Engineering, University of Kentucky, Lexington, Kentucky, US.

Ameri, M., Novin, M. K., and Yousefi, B. (2014). “Comparison of the field measurements of asphalt concrete densities obtained by ground-penetrating radar, pavement quality indicator and the borehole coring methods.” *Road Materials and Pavement Design*, 15(4), 759–773.

AS/NZS (2008). “Methods of sampling and testing asphalt - method 1.2: Sampling — coring method.” *AS/NZS 2891.1.2*, Standards New Zealand.

AS/NZS (2013a). “Methods of sampling and testing asphalt - method 14.1.2: Field density tests - determination of field density of compacted asphalt using a nuclear surface moisture-density gauge - backscatter mode.” *AS/NZS 2891.14.1.2*, Standards New Zealand.

AS/NZS (2013b). “Methods of sampling and testing asphalt - method 14.2: Field density tests — determination of field density of compacted asphalt using a nuclear thin-layer density gauge.” *AS/NZS 2891.14.2*, Standards New Zealand.

AS/NZS (2014). “Methods of sampling and testing asphalt - method 9.1: Determination of bulk density of compacted asphalt — waxing procedure.” *AS/NZS 2891.9.1*, Standards New Zealand.

ASTM (2016). “Standard test method for density of bituminous paving mixtures in place by the electromagnetic surface contact methods.” *ASTM D7113/D7113M*, American Society for Testing and Materials (ASTM).

ASTM (2019). “Standard test method for bulk specific gravity and density of non-absorptive compacted asphalt mixtures.” *ASTM D2726/D2726M*, American Society for Testing and Materials (ASTM).

ASTM (2022). “Standard test method for density of asphalt mixtures in place by nuclear methods.”

ASTM D2950/D2950M, American Society for Testing and Materials (ASTM).

Bashiri, M. and Farshbaf-Geranmayeh, A. (2011). "Tuning the parameters of an artificial neural network using central composite design and genetic algorithm." *Scientia Iranica*, 18(6), 1600–1608.

Brisset, P., Charbucinski, J., and Jin, J. (2005). "Technical data on nucleonic gauges." *Report No. IAEA-TECDOC-1459*, International Atomic Energy Agency (IAEA), Vienna, Austria.

Chandrakasu, M., Rajiah, M., Senathipathi, V., and Sekhar, C. R. (2021). "Comparative evaluation of distress prediction modeling of village roads in india using regression and ann techniques." *Journal of Transportation Engineering, Part B: Pavements*, 147(3).

Commuri, S., Mai, A. T., and Zaman, M. (2011). "Neural network based intelligent compaction analyzer for estimating compaction quality of hot asphalt mixes." *Journal of Construction Engineering and Management*, 137(9), 634–644.

Commuri, S. and Zaman, M. (2008). "A novel neural network-based asphalt compaction analyzer." *International Journal of Pavement Engineering*, 9(3), 177–188.

Foresee, F. and Hagan, M. (1997). "Gauss-newton approximation to bayesian learning." *Proceedings of International Conference on Neural Networks (ICNN'97)*, 1930–1935.

Hagan, M. and Menhaj, M. (1994). "Training feed-forward networks with the marquardt algorithm." *IEEE Transactions on Neural Networks*, 5(6), 989–993.

Haykin, S. (2000). "Neural networks: A guided tour." *Soft Computing and Intelligent Systems*, Academic Press, 71–80.

Hoegh, K., Steiner, T., Teshale, E. Z., and Dai, S. (2020). "Minnesota department of transportation case studies for coreless asphalt pavement compaction assessment." *Transportation Research Record*, 2674(2), 291–301.

Khadse, C. B., Chaudhari, M. A., and Borghate, V. B. (2017). "Electromagnetic compatibility estimator using scaled conjugate gradient backpropagation based artificial neural network." *IEEE Transactions on Industrial Informatics*, 13(3), 1036–1045.

Kvasnak, A., Williams, C., and Ceylan, H. (2007). "Investigation of electromagnetic gauges for determining in-place hma density." *Report No. IHRB Project TR-547*, Iowa State University, Iowa, US.

Laufenberg, T. L. (1986). "Using gamma radiation to measure density gradients in reconstituted wood products." *Forest products journal*, 36(2), 59–62.

Leiva, F., Gatiganti, S., and Brenes, A. (2022). "Pavement thickness and base layer effects on density profile system measurement." *Journal of Testing and Evaluation*, 51(4).

Malpass, G. and Khosla, N. (2002). "Evaluation of gamma ray technology for the measurement of bulk specific gravity of compacted asphalt concrete specimens." *Journal of the Association of Asphalt Paving Technologists*, 70, 352–367.

Møller, M. F. (1993). "A scaled conjugate gradient algorithm for fast supervised learning." *Neural Networks*, 6(4), 525–533.

Negnevitsky, M. (2011). *Artificial intelligence: a guide to intelligent systems*. Pearson Education Limited, London, UK.

Porubiaková, A. and Komačka, J. (2015). "A comparison of dielectric constants of various asphalts calculated from time intervals and amplitudes." *Procedia Engineering*, 111, 660–665.

Puig-Arnavat, M. and Bruno, J. C. (2015). "Artificial neural networks for thermochemical conversion of biomass." *Recent Advances in Thermo-Chemical Conversion of Biomass*, 133–156.

Sawyer, J., Ledford, D., and Gilliam, J. (2005). "Field investigation: Evaluation of non-destructive density testing devices on hot mix asphalt." *Report No. 36691.3.1*, North Carolina Department of Transportation, North Carolina, US.

Smith, J. S., Wu, B., and Wilamowski, B. M. (2019). "Neural network training with levenberg–marquardt and adaptable weight compression." *IEEE Transactions on Neural Networks and Learning Systems*, 30(2), 580–587.

Tian, H. and Shang, Z. (2006). "Artificial neural network as a classification method of mice by their calls." *Ultrasonics*, 44, 275–278.

Ziari, H., Behbahani, H., Izadi, A., and Divandary, H. (2010). "Field evaluation of nondestructive tests in measuring the pavement layers density." *Journal of Applied Sciences*, 10(23), 3091–3097.

LIST OF FIGURE CAPTIONS

TABLE 1. The performance of the LM-ANN model with different numbers of neurons in its hidden layer

The number of neurons in the hidden layer	LM-ANN model			
	Training time (s)	RMSE (kg/m^3)		
		Training	Validation	Testing
2	0.15	29.53	29.48	30.61
3	0.08	27.79	28.89	29.01
4	0.08	26.50	27.42	28.63
5	0.08	25.52	27.01	28.90
6	0.07	25.44	27.38	30.39
7	0.07	25.43	27.85	30.82
8	0.07	24.93	28.74	31.57
9	0.07	25.11	28.33	30.14
10	0.08	24.66	28.92	31.37
11	0.07	24.61	28.09	30.01
12	0.07	24.65	28.33	30.78
13	0.07	24.33	28.71	31.94
14	0.08	24.24	29.21	31.21
15	0.07	25.46	31.29	32.03

TABLE 2. The performance of the BR-ANN model with different numbers of neurons in its hidden layer

The number of neurons in the hidden layer	BR-ANN model			
	Training time (s)	RMSE (kg/m^3)		
		Training	Validation	Testing
2	0.09	27.23	27.44	28.23
3	0.09	26.69	27.29	28.16
4	0.10	26.40	27.24	27.29
5	0.10	26.55	27.52	27.87
6	0.11	26.44	26.92	28.04
7	0.10	27.16	27.78	28.16
8	0.10	26.74	28.04	28.93
9	0.10	27.79	28.33	29.41
10	0.12	26.92	27.56	27.67
11	0.11	27.26	28.28	28.52
12	0.11	26.83	27.92	28.09
13	0.11	26.42	27.77	28.45
14	0.13	26.44	27.43	28.03
15	0.11	27.11	28.41	28.88

TABLE 3. The performance of the SCG-ANN model with different numbers of neurons in its hidden layer

The number of neurons in the hidden layer	SCG-ANN model			
	Training time (s)	RMSE (kg/m^3)		
		Training	Validation	Testing
2	0.08	33.47	33.54	35.35
3	0.07	32.94	32.73	33.74
4	0.07	32.34	32.19	33.40
5	0.07	32.26	31.91	33.45
6	0.07	31.82	32.13	33.05
7	0.07	31.74	31.18	33.39
8	0.07	30.77	30.55	31.80
9	0.07	30.76	30.92	33.03
10	0.07	30.97	30.49	32.27
11	0.07	30.61	29.71	33.37
12	0.07	30.21	30.98	32.51
13	0.06	29.52	31.58	32.62
14	0.07	29.70	31.31	32.42
15	0.07	30.12	30.46	33.09

TABLE 4. The RMSEs of the ANN-predicted densities and the measured densities.

ANN-Predicted Densities			NDG	PQI
LM-ANN model	BR-ANN model	SCG-ANN model	Density	Density
28.63	27.29	31.08	65.53	68.00