Development of a Hybrid Machine Learning Model for Asphalt Pavement Temperature Prediction

ABDALRHMAN MILAD¹, IBRAHIM ADWAN², SAYF A. MAJEED³, ZUBAIR AHMED MEMON⁴, MUNDER BILEMA⁵, HEND ALI OMAR⁶, MAHER G. M. ABDOLRASOL⁷, ALIYU USMAN⁸ AND NUR IZZI MD YUSOFF²

¹Department of Civil and Environmental Engineering, College of Engineering and Architecture, University of Nizwa, Birkat-al-Mouz, 616, Nizwa, Sultanate of Oman; (e-mail: a.milad@unizwa.edu.om)
²Department of Civil Engineering, Universiti Kebangsaan Malaysia, 43600 UKM Bangi, Selangor, Malaysia; (e-mail: ibe598@gmail.com; izzi@ukm.edu.my)
³Technical Computer Engineering Department, Al-Hadba University College, Mosul, Iraq; (e-mail: sayf@hcu.edu.iq)
⁴Department of Engineering Management, Prince Sultan University (PSU), Riyadh, Saudi Arabia (e-mail: zamemon@psu.edu.sa)
⁵Faculty of Civil Engineering and Built Environment, Universiti Tun Hussein Onn Malaysia, Johor, Malaysia; (e-mail: mond0199131@gmail.com)
⁶Department of Civil Engineering, University of Tripoli, Tripoli, Libya; (e-mail: hend.omar@uot.edu.ly)
⁷Department of Electric, Electronics and System Engineering, Faculty of Engineering and Built Environment, Universiti Kebangsaan Malaysia, Bangi 43600, Malaysia; (e-mail: maher.abdolrasol@gmail.com)
⁸Department of Civil and Environmental Engineering, Universiti Teknologi PETRONAS, 32610, Bandar Seri-Iskandar, Perak, Malaysia; (e-mail: aliyu17005209@utp.edu.my)

Corresponding author: Abdalrhman Milad (e-mail: a.milad@unizwa.edu.om); Zubair Ahmed Memon (e-mail: zamemon@psu.edu.sa).

The authors would like to thank Prince Sultan University for their financial support.

ABSTRACT Machine learning (ML) models are excellent alternative solutions to model complex engineering issues with high reliability and accuracy. This paper presents two extensively explored ensemble models for predicting asphalt pavement temperature, the Markov chain Monte Carlo (MCMC) and random forest (RF). The RF and multiple MCMC (RF-MCMC) were used to hybridise the proposed algorithms for the optimal prediction of asphalt pavement temperature. This study used thermal instruments to measure the asphalt pavement temperature in Gaza Strip, Palestine. The temperature measurements were made at a two-hour interval from March 2012 to February 2013. The temperature data was used to model the pavement temperature. More than 7200 measured pavement temperatures were used to train and validate the proposed models. The validation showed that the ML models are satisfactory. The modelling results ensured the value of the proposed hybridisation models in predicting the asphalt pavement temperature levels. The developed hybrid algorithms regression model achieved acceptable and better prediction results with a coefficient of determination (R²) of 0.96. Generally, the results confirmed the significance of the proposed hybrid model as a reliable alternative computer-aided model for predicting asphalt pavement temperature.

INDEX TERMS Geophysical monitoring, Hybridisation Algorithms, Machine Learning, Measurement, Pavement Temperature Profile.

I. INTRODUCTION

A. RESEARCH BACKGROUND

Asphalt flexible pavements are the critical component of transport and highways and are more widely used than other materials because of their superior service performance [1-3]. The basis of flexible paving design is climatic conditions and axle load limit [4]. Ninety per cent of the pavements constructed throughout the world use asphalt binders [5]. The design of asphalt pavements is dependent on the structural capacity of the pavement layers, which is influenced by the applied stress and recoverable strain in the resilient modulus test, which involves different temperature ranges, loads, rest periods and loading axes [6, 7]. However, air temperature could cause severe damage to asphalt pavements and adversely affect pavement performance [8]. Therefore, it is essential to consider air temperature when designing flexible pavements. Hence, it leads to the asphalt
binder changes from Newtonian to a viscoelastic fluid to a viscoelastic solid [9-11]. The cumulative and consumption conditions of asphalt pavements lead to a wastage of structural and practical characteristics that reduce serviceability [12, 13].

The mechanistic–empirical design provides a more asphalt pavement approach to most of the materials used to measure the performance of the asphalt pavement, and it is used to determine the parameters affecting the asphalt pavement during different seasons [14-17]. An appropriate asphalt pavement temperature model is critical to ensure the safety and comfort of road users and extend the service life of road pavements [18,19]. Over the past few decades, researchers have developed many regression models to predict asphalt pavement temperatures, and some of these models have good accuracy. However, they require several input parameters for each prediction equation with their strengths and weaknesses [20,21]. The accuracy of a statistical method can only be established within the range of the original data used to develop the MLR models [22]. With technological advances, engineers now use computers and software to analyse data [22-24].

B. LITERATURE REVIEW
Barber was one of the first researchers to develop a thermal conductivity model for calculating the internal temperature profile of asphalt pavements [25]. In 1987, the Strategic Highway Research Program (SHRP) introduced the Long-Term Pavement Performance Program (LTTP), a 20-year research program to improve pavement characterisation at selected sites of the Seasonal Monitoring Program (SMP) [27-29]. In 1993, Lytton et al. developed an enhanced integrated climatic model (EICM) to predict the heat of pavement structures caused by climate parameter change [26]. And in the same year, researchers used the heat transfer theory to develop and validate the model for the summer using the highest pavement surface temperature [30]. Matic et al. developed a model to predict pavement temperature at particular depths using air temperature and other climate parameters [29]. In 2015, Islam et al. developed a series of statistical models for predicting the minimum, maximum and average pavement temperatures using the field measurement data from an instrumented road in New Mexico, USA [31]. Researchers installed two weather stations at a project site in Ashtabula, Ohio, USA, to collect climate data and used the data to develop regression models to predict the maximum and minimum daily temperatures [32]. Khan et al. developed and validated a regression profile to determine the average temperature in asphalt pavement layers. The temperature data is from a road segment near Albuquerque, New Mexico, USA, and the average depth for measuring the mean temperature is 2.54 inches [33].

In 2018, Li et al. developed and validated a statistical model for predicting the temperature at depths of more than 30 cm located in Alberta, Canada. The hourly pavement temperature measurement and meteorological data collection were made at three selected sites. The model was validated, and its applicability was verified using the available data from other locations [35]. In recent years, researchers have shown increasing interest in using artificial intelligence techniques, for example, machine learning (ML) and deep learning, to analyse and predict pavement temperature parameters [4, 36]. In 2014, a study in Serbia used artificial neural networks to develop models for predicting the minimum and maximum pavement temperatures based on the pavement’s surface temperature and depth [23]. The existing hybrid models for predicting the asphalt pavement’s temperature consider the impact of environmental factors on asphalt pavement. However, these models do not consider the reflect impact from the heated pavement that contribute to the heat accumulation in the surrounding environment, such as that in urban heat island (UHI) effects. Therefore, it is essential to conduct studies that take this factor into account.

C. MOTIVATION FOR THE RESEARCH
In the last decade, researchers used ML techniques to develop pavement temperature prediction models that use parameters such as solar radiation and air temperature. Their research was limited to the application of evolutionary algorithms such as gradient boosting regression tree, back-propagation (BP) neural network, Gaussian process regression (GPR) models, support vector machine (SVM) regression trees, ensembles of trees, linear regression models, and random forest [37]-[42]. There is an urgent need to develop a hybrid model to address these gaps. Specifically, the researchers proposed different ways to set up the RF parameters in the hybrid RF and MCMC models (too many trees can considerably slow down an algorithm). Even though these approaches could enhance the effectiveness of MCMC in various applications, the methods also particularly manifest in terms of RF parameter selection. The algorithm models have unique main aspects, strengths and weaknesses. The idea of merging these concepts holistically to combine their advantages has not been extensively explored in the literature.

D. RESEARCH OBJECTIVES
Based on the review of the reported literature, limited studies were conducted using machine learning to achieve better accuracy in predicting asphalt pavement temperature. The proposed model was built using ML and the field data from selected sites in the Gaza Strip, Palestine. The hybrid model was used to predict the asphalt temperature at varying depths. This study used the hybrid ML algorithms of random forest (RF) and Markov chain Monte Carlo (MCMC) regression to predict asphalt pavement temperature. Three input sets of correlation statistics were tested for this model to improve the influential parameters. At the same time, the validations were compared with other models, such as
artificial intelligence and empirical formulations. The remainder of this paper is organised as follows. Section 2 explains the observation location and data acquisition, Section 3 discusses the applied ML models, Section 4 presents the results and discussion, and Section 5 compares the result of this study with a previous study.

II. OBSERVATION LOCATION AND DATA ACQUISITION

The location for measuring pavement temperature is in Gaza, Palestine. An observation station was set up to measure pavement and air temperature at varying depths of 0 cm, 2 cm, 5.5 cm, and 7 cm at different time intervals during the four seasons of one year as shown in Figure 1. The location of the study area is shown in Figure 2, where manual measurements were performed to collect the data from March 2012 to February 2013 in the Gaza Strip, Palestine. The locations were pre-determined, and the condition at the data collection sites are well-known [35]. The data from the measurement station comprises three independent variables (air temperature, time, and depth) and one dependent variable (asphalt pavement temperature).

III. APPLIED MACHINE LEARNING MODELS

This study predicts asphalt temperature \( Y \) using air temperature \( x_1 \), time of the day \( x_2 \), and depth of measurement \( x_3 \). The data comprises the target data and input data in a machine learning problem, where \( Y \) is the target data and \( X = [x_1 \ x_2 \ x_3] \) is the input data. The linear regression formula for the data is given by Equation (1).

\[
Y_i = \beta_0 + x_1\beta_1 + x_2\beta_2 + x_3\beta_3
\]  

(1)

FIGURE 1. Measurement of asphalt pavement temperature in different depths.

Digital Thermometer

\[ \begin{align*}
20 \text{ cm} & \quad 20 \text{ cm} \\
2 \text{ cm} & \quad 5.5 \text{ cm} & \quad 7 \text{ cm} \\
5 \text{ mm Glycerin} & \\
\end{align*} \]

Asphalt Layer Surface

FIGURE 2. Location of the study area.
where $Y_i$ is the data point of $Y$, and $x_{1i}, x_{2i},$ and $x_{3i}$ are the $i^{th}$ data points of air temperature, time of the day, and depth. Therefore, $\beta_0, \beta_1, \beta_2,$ and $\beta_3$ are the unknowns in the equation, also known as regression coefficients. The regression coefficients are calculated using the MCMC method, which is explained in the following subsection.

A. MARKOV CHAIN MONTE CARLO

According to a survey, the Metropolis algorithm is among the ten algorithms that effectively influenced the progress of science and engineering in the last century [43], [44]. However, the MCMC algorithm is an example of a large sampling algorithm class with a significant part of computing science and statistics. This technique draws samples, and the following sample is based on the existing sample sequences of probabilistically related events to result in a solution within a reasonable period [45], [46].

This type of ML has the advantage of making reversible jumps because of its flexibility to include many possible moves based on the selection strategy compared to other schemes that rely on the indicator mix selections or diffusion processes [47]. The reversible jump is a tricky and time-consuming mechanism of this technique for reversible engineering actions. In addition, this technique is often used for integration or optimisation problems in bulky space dimensions. The MCMC is a class of algorithms used for sampling an unknown distribution when the set of observations is given [48]. The Metropolis-Hastings algorithm was chosen because of its suitability for multidimensional sampling problems [49], [50]. This application requires a four-dimensional sample, as shown in Figure 3.

![Graphical representation of MCMC algorithm.](image)

When using MCMC to solve regression problems, the regression coefficients are guess values (or random values). The algorithm is iterated using random processes and the given distribution, known as a transition function. Here, the Gaussian normal distribution is given by Equation (2) [49].

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma}exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \tag{2}$$

In each iteration, the transition function generates new regression coefficients with an optimal variance $\sigma$. However, a significant variance can cause a miss of maxima; a very small variance could increase the number of iterations, and the algorithm may be trapped at the local maxima. This study chose $\sigma = 0.0003$ based on the chain performance and $\mu = \beta_{old}$. This transition was achieved using a normal distribution random number generator. Hence, the transition is expressed as $\beta_{new} = N(\beta_{old}, \sigma)$. The next step in the Metropolis-Hastings algorithm is deciding whether to accept or reject the new regression coefficient values generated by the transition function. For this purpose, the ratio of posterior probabilities is calculated and compared with a random number. If the ratio of posteriors is greater than the random number, the values of $\beta$ are accepted. Otherwise, the values are rejected. This process is expressed using Equation (3) [49].

$$R = \frac{\pi(\beta_{new})}{\pi(\beta_{current})} \tag{3}$$
where $R$ is the acceptance ratio and function $\pi$ is the posterior. The probability of acceptance is defined as $P(\text{acc}) = \min(1, R)$; this constrains the value of the acceptance probability between 0 and 1. A random number from a uniform distribution, $U$, is generated. If the random number continue after storing the accepted values in an array and setting $\beta_{\text{current}} = \beta_{\text{new}}$. If $U \geq P(\text{acc})$, the new values of $\beta$ are rejected, and iterations are continued by setting $\beta_{\text{current}} = \beta_{\text{current}}$. Figure 3 shows the generalised working of MCMC. Figure 4 shows the flowchart for implementing the

$U < P(\text{acc})$, then new values of $\beta$ are accepted, and iterations continue after storing the accepted values in an array and setting $\beta_{\text{current}} = \beta_{\text{new}}$. If $U \geq P(\text{acc})$, the new values of $\beta$ are rejected, and iterations are continued by setting $\beta_{\text{current}} = \beta_{\text{current}}$. Figure 3 shows the generalised working of MCMC. Figure 4 shows the flowchart for implementing the

**FIGURE 4.** Metropolis-Hastings algorithm.

Furthermore, in order to decide on the acceptance or rejection of the proposed regression coefficients, there is a need for some sample distribution that roughly can be considered a posterior. If one follows the Bayes formula, the posterior can be defined using Equation (4) [50].

\[
\text{Posterior} = \frac{\text{Likelihood} \times \text{Prior}}{\text{Normalising Constant}} \quad (4)
\]

The denominator in Equation (4) is the normalising constant and cannot be evaluated analytically because the posterior is unknown. Therefore, the term is ignored when
calculating the approximate posterior for the acceptance ratio. Equation (4) can be written as Posterior × Likelihood + Prior; taking the ratio of two posteriors gives an approximate posterior ratio. The prior and likelihood can be taken as some simple distribution from which the samples can be obtained. This study took the normal distribution as the likelihood by representing ϕ as a precision parameter for simplicity, as given in Equation (5).

\[
L(\beta_0, \beta_1, \beta_2, \beta_3, \varphi) = \prod_{i=1}^{n} \left\{ \frac{\varphi}{2} \exp \left( -\frac{\varphi}{2} (Y_i - Y_{pr})^2 \right) \right\} \tag{5}
\]

where \( Y_{pr} = \beta_0 + x_1\beta_1 + x_2\beta_2 + x_3\beta_3 \) as in Equation (1). The prior is given by the gamma distribution expressed as Equation (6).

\[
\pi(\beta_0, \beta_1, \beta_2, \beta_3, \varphi) = \varphi^{a-1} \exp(-b\varphi) \prod_{k=0}^{\beta_3} \exp \left\{ \frac{-1}{2\sigma_k^2} \left( \beta_k - \mu_k \right) \right\} \tag{6}
\]

where \( a = 1 \) and \( b = 1 \) are the parameters for the gamma distribution; \( \sigma_k \) is the standard deviation for each value of \( \beta \), which is updated based on the \( \beta \) update in each iteration during the training. Hence, by using the substitutes for the likelihood and prior from Equations (5) and (6), the posterior can be written as Equation (7).

\[
\pi(\beta_k | y) = L(\beta_k, \varphi) \times \pi(\beta_k, \varphi) \tag{7}
\]

where \( \beta_k = \beta_0, \beta_1, \beta_2, \beta_3 \).

**B. RANDOM FOREST**

Ho introduced the random forest in 1995, and in 2001, Beriman further developed this algorithm [51]. RF is a computationally efficient technique for large data. Many research projects and real-world applications have successfully used RF in various fields [52]-[55]. The RF algorithm is excellent for learning methods using classification and regression trees (CART). This algorithm is dependent on the bootstrap resampling algorithm in a multiple sample extraction of the basic data; in short, it is based on the statistical learning theory [56], [57]. The predictions of each decision trees combined into multiple decision trees and averaged to obtain the best prediction output. However, increasing the diversity of the decision trees by randomly repeating the samples changes its predicted combination evolution for each tree decision. Because this study focused on predicting regression, the priority was introduced to a regression tree (RT). The leaf node implies that the samples were formulated using the mean square error (MSE) for each sample and regression tree branch. However, each minimum leaf node MSE was a branching condition executed till no more features were available or the optimal total MSE was satisfied, which was the end of the regression tree progress. It has two custom parameters, the number of RTs and random variables, such as the number of nodes known as Max depth [52]. Then, the data processing-related errors allowed only the minimisation to ensure the optimised parameters. The modelling of the RF algorithms used the bootstrap sampling-based extraction of N estimator training sets of the data. The training set is made up of about two-thirds of the original dataset. However, one-third of the dataset was not a part of each RF, and the bootstrap sample data were not plotted as a graph in the training progress [36]. This portion of the dataset is the untrained data. After that, the generating RT to every bootstrap training set, next by a “forest” of constructing a sum of N estimators’ RT which is not involved. The growth of each tree did not pick every optimal attribute, such as an internal node to each branching; in contrast, the selection of an optimal attribute was a random process based on the max depth attributes. The RF algorithm then created several training sets to enhance the differences between the regression models and improve the extrapolation prediction ability of the combined regression models. The regression model sequence \( t_1(x), t_2(x), \ldots, t_k(x) \) was derived using n-time model training; a sequence was used to build the multiple regression model system known as a forest. However, the prediction outcomes of the N estimator RTs were pooled and averaged to calculate the value of the updated sample in the search space. Figure 5 shows the RF model architecture. In this study, the task is a regression because the target data is asphalt temperature, which is continuous data. The term ‘Random Forest’ in this paper indicates ‘Regression Forest’ henceforth. RF can be defined as a ‘collection of tree structure classifiers/regressors \( h(x, \Theta_t), t = 1, \ldots \), where \( \Theta_t \) is independent identically distributed random vectors. Each tree casts a unit vote for the most popular class at input x in the classification task, and the mean of the output of all the trees is taken in the regression task, as shown in Figure 5 [58]. The RF implementation in this study was adapted from Criminisi et al. The output of the forest is the average of all the trees in Equation (8) [55].

\[
\pi(y|x) = \frac{1}{T} \sum_{t} \pi_t(y|x) \tag{8}
\]

where \( \pi(y|x) \) is the posterior average of all forests, and \( T \) is the number of trees. The posterior of each individual was obtained by training the tree for an objective function, which is to maximise the information gain, as given by Equation (9).

\[
l_f = \sum_{v \in S_j} \log(|\Lambda_v(u^v)|) - \sum_{i \in (L,R)} \left( \sum_{v \in S_j} \log(|\Lambda_v(u^v)|) \right) \tag{9}
\]

where \( \Lambda_v \) is a conditional covariance matrix, \( S_j \) is the training data available at node \( j \), and \( S_L^j \) and \( S_R^j \) are the left and right node splits of the data, respectively. However, the output in this dataset is a single variable array, and the information is expressed as Equation (10).
where $y$ is the estimated output, and $y_j$ is the mean of the training sample at the $j^{th}$ node. In this maximum information gain process, the data arriving at node $j$ were split into its right or left child. The data split was governed by the decision-making component of the tree called ‘weak learner’ given by Equation (11).

$$h(x, \theta_j) = [\tau_1 > \varphi(x), \psi > \tau_2]$$

where $\varphi(x)$ represents the feature space of the data, $\psi$ is the geometric shape representing the data separation entity, and $\tau_1$ and $\tau_2$ are the constraining variables trained in each iteration. In this study, the data separators are vertically aligned to the $y$-axis. These axis-aligned separators are called ‘Stumps’ [59]. Decision stumps separate the data showing the relationship between each input parameter and the asphalt temperature. The decision stumps for time and depth are simple and directly defined because of their discrete nature. However, for the relationship between the air temperature and asphalt temperature, the decision stumps have to be calculated separately in each training iteration depending upon the difference between the two consecutive values to be separated.

These figures show that the data did not fit well with the predicted data when using a single linear regression model consisting of four regression coefficients given by Equation (1). Figure 7 shows that the time consists of ten finite sets of time, and Figure 8 shows that the depth consists of four finite sets of depths. Therefore, the data has 40 labelled discrete sets, which increase the total number of correlation coefficients from 4 to 160 ($4 \times 40$). RF is used to learn the labels and the MCMC regression for each dataset combined with the labels.

$${\mathcal{I}}_j = \sum_{v \in S_j} (y - \bar{y}_j)^2 - \sum_{i \in \{LR\}} \left( \sum_{v \in S_j} (y - \bar{y}_j)^2 \right)$$

(10)
FIGURE 6. Scatter plot of asphalt temperature (°C) vs air temperature (°C).

FIGURE 7. Scatter plot of asphalt temperature (°C) vs time of the day.
The training of the algorithm was carried out as follows.

1) Clustering the Data Using the K-Means Algorithm
The k-means clustering algorithm can be explained in the following four steps [59].

i. Randomly choose the initial j clusters C = c1; : : : ; cj .
ii. Iterate over the j clusters, and select the closest set of points to each cluster.
iii. Set a centroid, ci, for each cluster.
iv. Repeat steps ii and iii until there is no cluster change.

The algorithm divided the training dataset into j = 40 clusters, X = [x1, x2, . . . , x40], and assigned each data point a class named after its cluster.

2) Train the Random Forest to Learn the Cluster of the Training Data
The target data in this part of RF is the training labels generated by the k-means clustering. However, as described earlier, there is a separate MCMC regression for each dataset. To know which regression coefficients belonged to which dataset, it is essential to label the regression coefficients with the dataset. This was done using the RF, as shown in Figure 9. The RF learned the dataset label and assigned the corresponding label to each set of regression coefficients created using the MCMC algorithms.

3) Train the MCMC for Each Cluster
In Equations (5), (6), and (7), each MCMC posterior distribution depend on four βk values. However, in the proposed hybrid algorithm, as the data were split into j = 40 sets to produce 160 values of βkj, the posterior distribution for all 160 values could be rewritten as follows.

\[
\pi(\beta_{k,j} | y) = \prod_{j=1}^{40} \pi_j(\beta_{k,j} | y) \tag{12}
\]

which is the product of all individual posteriors of the jth dataset given by \( \pi_j \).

\[
\pi_j(\beta_{k,j} | y) = L_j(\beta_k, \varphi) \times \pi_j(\beta_k, \varphi) \tag{13}
\]

Where

\[
L_j(\beta_0, \beta_1, \beta_2, \beta_3, \varphi) = \prod_{i=1}^{n} \left[ \frac{\varphi}{\sqrt{2\pi}} \exp \left\{ -\frac{\varphi}{2} (Y_{i,j} - Y_{pr,j})^2 \right\} \right] \tag{14}
\]

And

\[
\pi_j(\beta_0, \beta_1, \beta_2, \beta_3, \varphi) = \varphi^{n-1} \exp(-b\varphi) \prod_{k=0}^{3} \exp \left\{ -\frac{1}{2s_k^2} (\beta_{k,j} - \mu_{k,j})^2 \right\} \tag{15}
\]
4) Label the Corresponding Regression Coefficients with the Labels Generated by the RF

The regression coefficient output of this algorithm works for a \( j \times k \) matrix, where \( j = 40 \) and \( k = 4 \). In Equation (1), \( \beta_0 \) is set as the offset value, \( \beta_1 \) is the air temperature regression coefficient, \( \beta_2 \) is the time of the day regression coefficient, and \( \beta_3 \) is the depth regression coefficient. The selection of \( \beta_k \) from its \( j \) number of sets is dependent on the label of the data determined by RF, as shown in Equations (12) and (13).

Therefore, the data labels generated by the RF were added to the sets of regression coefficients generated by the MCMC. Figure 9 shows the training process for the hybrid algorithm.

The testing prediction procedure for the hybrid algorithm can be shown stepwise as a reading test data, which classifies the test data into one of the 40 clusters using the RF classifier, reading the regression coefficients corresponding to the class, and multiplying the regression coefficients. The result is the predicted asphalt temperature.

**FIGURE 9.** Scatter plot of asphalt temperature (°C) vs depth.

### IV. RESULTS AND DISCUSSION

This section presents the development of the proposed ML models using three input combinations, and the final results are shown in Table 1. The algorithms developed in this study were tested for five random combinations of the test dataset, each with a length of 1500 points. The random forest, MCMC, and proposed RF-MCMC algorithms were tested for the same combination of data. In each trial, the model performance in the testing phase of the prediction was evaluated using four metrics, \( R^2 \), \( RMSE \), \( MAE \), and \( NSE \).

1) Parameter \( R^2 \) is the square of the correlation, where an \( R^2 \) of 1 means the best prediction and the similarity between the predicted data and the observed data.

\[
R^2 = \frac{\sum_{i=1}^{N} (y_0 - \bar{y}_0)(y_p - \bar{y}_p)}{\sqrt{\sum_{i=1}^{N} (y_0 - \bar{y}_0)^2 \sum_{i=1}^{N} (y_p - \bar{y}_p)^2}}
\]

where \( y_0 \) and \( y_p \) are the observed and predicted asphalt temperature, respectively, and \( N \) is the number of test/observation samples.

2) Root mean square error (RMSE) is the square root of the averaged squared errors. Each error in the RMSE is proportional to the squared error size, where more significant errors disproportionately affect RMSE.
\[
RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (y_p - y_0)^2}
\]

3) Mean absolute error (MAE) is the average of the absolute values of the errors. MAE is easier to comprehend than the square root of the average squared errors (RMSE) [62].

\[
MAE = \frac{1}{N} \sum_{i=1}^{N} |y_p - y_0|
\]

4) The Nash-Sutcliffe efficiency (NSE) is used to determine the goodness of fit between the observed and predicted data [63].

\[
NSE = 1 - \frac{\sum_{i=1}^{N} (y_0 - y_p)^2}{\sum_{i=1}^{N} (y_0 - \bar{y}_0)^2}
\]

The results show that the proposed algorithm gave better results for all evaluated parameters in this study. Table 1 shows that the hybrid algorithm with a combination of M3 has the best performance with \(R^2 = 0.9625\), \(RMSE = 0.0378\), \(MAE = 0.0308\), and \(NSE = 0.9605\).

**TABLE 1. Performance of the applied predictive model in the testing stage**

| Performance Evaluation | Test Number | ML | \(R^2\) | RMSE | MAE | NASH |
|------------------------|-------------|----|--------|-------|-----|------|
|                        | M1          | RF | 0.9249 | 0.0534 | 0.0387 | 0.9228 |
|                        |             | MCMC | 0.9384 | 0.0481 | 0.0389 | 0.9340 |
|                        |             | HYBRID | 0.9606 | 0.0385 | 0.0316 | 0.9578 |
|                        | M2          | RF | 0.9251 | 0.0526 | 0.0383 | 0.9237 |
|                        |             | MCMC | 0.9362 | 0.0481 | 0.0388 | 0.9329 |
|                        |             | HYBRID | 0.9604 | 0.0379 | 0.0312 | 0.9577 |
|                        | M3          | RF | 0.9224 | 0.0540 | 0.0394 | 0.9197 |
|                        |             | MCMC | 0.9382 | 0.0480 | 0.0388 | 0.9335 |
|                        |             | HYBRID | 0.9616 | 0.0379 | 0.0311 | 0.9584 |
|                        | M4          | RF | 0.9268 | 0.0556 | 0.0395 | 0.9259 |
|                        |             | MCMC | 0.9399 | 0.0479 | 0.0387 | 0.9369 |
|                        |             | HYBRID | 0.9628 | 0.0377 | 0.0310 | 0.9608 |
|                        | M5          | RF | 0.9272 | 0.0529 | 0.0385 | 0.9253 |
|                        |             | MCMC | 0.9381 | 0.0485 | 0.0388 | 0.9346 |
|                        |             | HYBRID | 0.9619 | 0.0381 | 0.0312 | 0.9590 |

A graphical representation of the correlation coefficient (\(R\)) and \(RMSE\) with the standard deviation is shown as the Taylor diagram to intuitively compare the performance of the models. The Taylor diagrams for the five sets of data have four comparison points. The first point is from the original data (observed data) with \(R = 1\), \(RMSE = 0\), and next is the data points predicted by the different models. The data points are expected to show a similar maximum \(R\), minimum \(RMSE\), and standard deviation as the observed data. Based on the Taylor diagrams, it is concluded that the hybrid algorithm data point is the closest to the actual data, followed by MCMC and RF. Furthermore, the RF, MCMC, and hybrid algorithms showed significant performance differences.
FIGURE 11. Taylor diagram for the hybrid models of the two-dimensional testing of the standard deviation, correlation and RMSE.

A boxplot is a graph that shows how the values in the data are distributed. Although boxplots appear rudimentary compared to a histogram or density plot, they have the advantage of taking up less space, which is beneficial for comparing distributions across several groups or datasets. A boxplot may be used to determine whether a distribution is skewed and whether the data collection contains any potentially exceptional observations (outliers). The box is a rectangle with a vertical line inside it at the median. The top and the bottom of the box represent the 75th and 25th percentiles (3rd and 1st quartiles). The distribution is skewed if the median is not in the centre of the box. It is positively skewed if the median is closer to the bottom and negatively skewed if the median is closer to the top of the box. Extreme values and outliers are frequently indicated with a red ‘+’. Figure 12 shows the box plot analysis for the RF, MCMC, and hybrid algorithms for the estimated and measured asphalt temperature. In both cases, the hybrid algorithm showed the best performance with none or minimum outlier points. The
RF and MCMC have the highest and second-highest number of outliers in the prediction error, respectively.

![Box plot of the estimated and measured asphalt temperature validation based on the RF, MCMC, and RF-MCMC models.](image)

**FIGURE 12.** Box plot of the estimated and measured asphalt temperature validation based on the RF, MCMC, and RF-MCMC models.

V. COMPARISON WITH PREVIOUS STUDY

The literature review showed that several researchers have developed regression models for predicting pavement temperature [4]. Furthermore, some related studies on the regression models were investigated and compared with the proposed method in Table 2. The observations of the presented formulas revealed that there was no general mechanism used by all of the other research. Nonetheless, all models used regression processing to evaluate the pavement temperature on site. However, the validation of the proposed hybrid algorithm models against other regression model results from the practical perspective was essential. Hence, a valid justification could be reached as noted that the objective authentication performance of $RMSE$ was the best, considering the other machine learning methods. However, this reduction in the objective function reflected excellent results for asphalt pavement temperature prediction.

**TABLE 2.** Comparison of the machine learning performance when predicting pavement temperature.

| Reference    | Methods | $R^2$ | RMSE   |
|--------------|---------|-------|--------|
| Liu et al. [37] | MLR     | 0.931 | 3.8471 |
|              | RF      | 0.957 | 3.0281 |
|              | NN      | 0.913 | 4.1944 |
| Nojumi et al. [38] | SVR     | 0.950 | 2.92   |
Figure 13 summarises the RMSE of the hybrid performance against the results published in the literature. The proposed hybrid model showed an excellent prediction ability compared to others. For instance, the RF showed a 95% better accuracy than Liu et al. [37], GPR showed a 99% better accuracy than Nojumi et al. [38], and 96% better accuracy than Asefzadeh et al. [34], and 81.8% better accuracy than GBELM. These percentages proved that the proposed hybrid algorithm model has a promising potential, high reliability and high accuracy for predicting the pavement temperature. This result explains the potential of advanced machine learning (AML) technologies in facilitating engineering applications, reducing human interactions, preventing regression model failure, and providing generalisation capability.

VI. CONCLUSIONS

This study developed a new hybrid RF-MCMC model for predicting the temperature of asphalt pavements in the Gaza Strip. In order to utilise the machine learning application of the RF–MCMC model, measurement and determine the regularity of temperature on the pavement surface to measure the asphalt pavement temperature every two hours at a certain depth, which were 0cm, 2cm, 5.5cm, and 7cm, at different time intervals during one year with different variables. Additionally, the static and dynamic prediction methods in the ML methods were introduced to improve the prediction accuracy at the alteration points of the temperature range daily. This study merged two algorithm methods to develop the hybrid RF-MCMC model. These concepts have different main aspects and strengths and weaknesses. However, combining these concepts led to spectacular results. Even though basic machine learning models such as RF and MCMC can make good predictions of asphalt pavement temperature, the performance of the hybrid RF-MCMC is better in terms of the $R^2$ values. The hybrid RF-MCMC algorithm showed the best prediction capability after the hyper-parameter tuning with $R^2$ values of 0.9249, 0.9384, and 0.9606 for RF, MCMC, and hybrid RF-MCMC, respectively. Moreover, the RMSE, MAE and NSE were used to assess the performance of the proposed model. The results of the hybrid machine learning training and regression and the testing of the two-dimensional Taylor diagram for the hybrid models gave a standard deviation and excellent correlation. Moreover, the reduced RMSE error is effective in predicting the optimal patterns of asphalt pavement temperature. Furthermore, the results showed that the boxplot of the estimated and measured asphalt temperature for the validation comparison of the other techniques and the proposed hybrid exhibited magnificent performance.

REFERENCES

[1] H. Pérez-Acebo, et al., “IRI performance models for flexible pavements in two-lane roads until first maintenance and/or rehabilitation work”. Coatings, vol. 10, no. 2, pp. 97, 2020. DOI: 10.3390/coatings10020097.

[2] M. S. Yamany, et al., “Characterizing the performance of interstate flexible pavements using artificial neural networks and random parameters regression” J. Infrastruct. Syst., vol. 26, no. 2, pp. 4020010, 2020. DOI: 10.1061/(asce)is.1943-555x.0000542.

[3] Y. Wang, K. C. Mahboub, and D. E. Hancher. “Survival analysis of fatigue cracking for flexible pavements based on long-term pavement performance data”, J. Transp. Eng., vol. 131, no. 8, pp. 608–616, 2005. DOI: 10.1061/(asce)0733-947x(2005)131:8(608).
recycling, pavement design, waste materials in asphalt, asphalt technology, transportation traffic, transport engineering, transportation planning, driver behaviour and road safety. He is involved in research and asphalt-related projects. He is currently a reviewer for 13 international journals.

**HEND ALI OMAR** is an assistant professor. Her research interests are pavement engineering, pavement materials and pavement management system.

**MAHER G. M. ABDOLRASOL** received his BSc in electrical engineering and communication from the Omer Al-Mukhtar University, Tobruk, Libya in 2006, a Masters of Engineering Science in power electronics from the University of Malaya in 2011, and a PhD from Universiti Kebangsaan Malaysia UKM. He is currently researching microgrid controller, renewable energy management and artificial intelligent algorithms. He has several IEEE conferences and journal papers in his field of study.

**ALIYU USMAN** is a registered Engineer in Nigeria and a member of The Nigerian Society of Engineers (NSE). He is a lecturer at the Ahmadu Bello University Zaria Nigeria. He is currently pursuing his PhD in Highway Materials at the Universiti Teknologi PETRONAS, Malaysia. He obtained a bachelor degree in Civil Engineering from Bayero University Kano-Nigeria in 2009 and a Master of Science in Civil Engineering from Ahmadu Bello University Zaria-Nigeria in 2016. His fields of interest are concrete and pavement materials, highway engineering and sustainable highways.

**NUR IZZI MD YUSOFF** is a senior lecturer at the Department of Civil and Structural Engineering, Universiti Kebangsaan Malaysia (UKM). His research interests are pavement engineering and pavement materials.