LONG SHORT-TERM MEMORY WITH GATED RECURRENT UNIT BASED ON HYPERPARAMETER SETTINGS AND HYBRIDIZATION FOR REFERENCE EVAPOTRANSPIRATION RATE PREDICTION

MUHAMMAD ZAHID BIN HILMI[1], TONI ANWAR[2], DAYANG ROHAYA BINTI AWANG RAMBLI [3]

[1]Computer, Information and Science Department of Universiti Teknologi PETRONAS, Seri Iskandar, 32610, Malaysia.

[2]Computer, Information and Science Department of Universiti Teknologi PETRONAS, Seri Iskandar, 32610, Malaysia.

[3]Computer, Information and Science Department of Universiti Teknologi PETRONAS, Seri Iskandar, 32610, Malaysia.

E-mail: [1]zahid_19000298@utp.edu.my, [2]toni.anwar@utp.edu.my, [3]dayangrohaya.ar@utp.edu.my

ABSTRACT

The evapotranspiration rate can be used to estimate water loss. However, there are 31 equations available to be chosen, and randomly choosing the equation might not project the actual results. This is very crucial because, without the equation, we cannot proceed with the parameter selection. These findings can justify the parameter chosen for the prediction model development. Long Short-Term Memory (LSTM) is known for its ability to retain memory better than Recurrent Neural Network (RNN). This is due to LSTM architecture, where the memory cell is available to store memory for long-term dependency. RNN suffers from a vanishing gradient that can affect the prediction, whether in accuracy, precision, etc. LSTM was developed specifically to address the issue of RNN. Even though LSTM is better overall, it can be further enhanced. The proposed method is to adjust the Hyperparameter Settings and combine them with Hybridization. Our findings indicate that the prediction accuracy improved significantly. The hybrid model chosen was Gated Recurrent Unit (GRU), combined with LSTM and Hyperparameter Settings, resulting in the best and highest prediction accuracy compared to the LSTM Vanilla and LSTM with Hyperparameter Settings. LSTM Hyperparameter Settings and Hybridization dominate the top three scores. The scoring stretched until 11th place before the LSTM Hyperparameter Settings score came in. The top three scores were for Case 99, Case 36, and Case 90 with 0.0626, 0.06446, 0.06606 MAE, 0.00667, 0.00706, 0.00759 MSE, 0.0817, 0.084, 0.0871 RMSE and 0.99261, 0.99219, 0.9916 R^2, respectively. As for the LSTM Hyperparameter Settings score, 0.0712 MAE, 0.00861 MSE, 0.09278 RMSE, and 0.99047 R^2.

Keywords: *Hyperparameter, Hybridization, Deep Learning, LSTM, Evapotranspiration*

1. INTRODUCTION

1.1 Background

Water loss is a natural occurrence in agriculture, where crops lose water through transpiration and evaporation. Evaporation is when soil moisture and water surface level evaporate into the air due to the heat of vaporization [1]. Transpiration is when the vegetation loses water to cool down its temperature, especially during hot days [2]. Both evaporation and transpiration convert the liquid into water vapour. Water loss can be measured or estimated in multiple ways, using a lysimeter, eddy correlation, and water balance in a basin or using energy balance, mass transfer, and crop coefficients. However, Gocic and Trajkovic (2014) found that implementing the Reference Evapotranspiration (ETo) model is the most accepted method [3]. Evapotranspiration is a term that refers to a process that combines water surface and soil evaporation and vegetation transpiration [4]. It plays a vital role in water-related to agriculture, such as determining the next irrigation needed, water surface level, water loss level, and more [3].

This research aims to predict water loss by using the accepted method of estimating water loss. The method selects parameters based on the Reference Evapotranspiration (ETo) model. There are more than 30 ETo models developed by researchers worldwide. Researching the appropriate ETo model should be one of the main priorities. From the ETo model itself, the needed parameters can be laid out instead of randomly choosing any parameters or relying on commonly used ones applied in other research. Wind speed, air humidity, minimum and maximum temperature, solar radiation, and relative humidity are the most used parameters in any ETo model because they are widely available in most cases [5]. It can provide helpful information on whether there is a need for irrigation or not for that particular day [6]. Finding appropriate parameters is challenging because of the wide range of available ETo models that can be chosen. Still, it does not necessarily translate into better predicting the evapotranspiration rate. So, the task is to identify only the most helpful parameter based on the evapotranspiration rate model.

Several ETo models can be applied apart from the FAO-56 Penman-Monteith (PMF56) model. The ETo model can be divided into three categories [7]. The first category is temperature-based which is based on temperature parameters. The second category is radiation-based which is based on radiation parameters. The last category is mass transfer-based or water balance-based, based on temperature and humidity. Forth category has been added, a combination-based type, the PMF56 model [3][8]. Each ETo model has its data requirement even though the parameters are the same. The parameters can be derived into a new parameter by combining them or turning them into a different format, such as degrees Celsius to degrees Fahrenheit. Additional data can be derived according to the ETo model requirements. Most researchers agreed that the best ETo model is the combination-type PMF56 model because FAO was explicitly developed. PMF56 can estimate the ETo rate with the most accurate result. PMF56 model is the best model period to date based on their findings [3][9][10].

As for the prediction model, finding the most accurate prediction model will be a challenge because many models can be chosen. There are many prediction models available such as Feed Forward Neural Network (FFNN), Convolutional Neural Network (CNN), Recurrent Neural Network (RNN), and many more that can be applied. So, finding a suitable prediction model for this study will be challenging. The most used prediction model is the RNN variant, Long Short-Term Memory (LSTM) variant, and Gated Recurrent Unit (GRU). RNN can store temporal dependencies but suffer from a vanishing gradient in which LSTM was developed to overcome the shortcomings of RNN. GRU was created as an improved version of the current LSTM that tackles the processing time and computing power, resulting in faster processing predicted output. However, there were mixed results between LSTM being better than GRU and vice versa [11][12].

1.2 Problem Statements

Irrigation has always been one of the most crucial agricultural processes for growing crops. It can cause a significant problem for crops dependent on water, such as paddy. Water availability can affect growth and productivity (Wibowo, Rizaldi & Siregar, 2019). Over and under-irrigating can affect overall production yield in terms of the quality and quantity of the product. So, the goal is only to irrigate when needed. The problem was that there was no scientific way to estimate the water loss apart from depending on the experience handling irrigation that could assist the irrigation process without manually irrigating the field daily to maintain the paddy water level.

However, researchers developed a way to estimate water loss by creating the ETo model. Researchers explicitly designed to measure water loss using the ETo model for agricultural purposes. But, another problem is which of the ETo model can be applied. Hargreaves model might be the best case for location A, but it might not fit in location B. More than 30 ETo models can be chosen according to the suitability and data available for that particular area (Muhammad et al., 2019). Deciding which ETo model to use is quite challenging since many factors need to be considered because we need to establish the parameter required to develop the prediction model. Even though this main research focuses on improving prediction accuracy, the need to develop the parameter selection needs to be justified first.

The next deciding factor is finding the prediction model to achieve high accuracy. Recurrent Neural Network (RNN) is an excellent example of the best-performing algorithm. However, it had issued on vanishing gradient, affecting the prediction accuracy. So, another variant of RNN was developed to counter that issue: Long Short-Term

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Memory (LSTM). LSTM tackles the vanishing gradient issue but needs high computing power. A simpler version of LSTM, the Gated Recurrent Unit (GRU), tackles the processing time and computing power of LSTM. However, suppose prediction accuracy is the goal, and processing time and computing power are not considered. In that case, the focus should be on improving the LSTM prediction accuracy to be as highest as possible. In theory, adjusting the hyperparameter using the optimization method or manual tuning should improve the prediction accuracy (Xinrui et al., 2021; Alqushaibi et al., 2021). GRU Hybridization of the prediction model also improves prediction accuracy (Yu et al., 2021; Offiong et al., 2021; Liu et al., 2021; Yin et al., 2020). There were also combining optimization tuning and GRU Hybridization (Ahmed et al., 2021.) However, it is another challenge since it can affect either outperforming the existing prediction model or underperforming.

1.3 Research Questions

Several questions will be answered in this research to tackle the problems:

1. What are the factors that can improve the Vanilla LSTM prediction accuracy?
2. How to develop an improved Vanilla LSTM model that can achieve better prediction accuracy?
3. How to validate and verify the prediction accuracy of the proposed model?

2. LITERATURE REVIEW

2.1 Reference Evapotranspiration (ETo)

Evapotranspiration is a term that refers to a process that combines water surface and soil evaporation and vegetation transpiration [4]. With evapotranspiration, estimating the evapotranspiration rate would be possible. By having a proper method to estimate the evapotranspiration rate, predicting would be easier because it is quite a challenge due to data selection, let alone adding predicting the rate itself.

There are many ways to estimate water loss in the form of evapotranspiration which, according to Xiang et al. (2020), is divided into three which are Actual Evapotranspiration (ETa), Potential Evapotranspiration (ETp), and Reference Evapotranspiration (ETo) [13]. The three evapotranspiration are closely related; however, the concepts and models are different, differentiating them from each other. For an extended period, the uses of ETp and ETo were muddled, and it has been cleared out in proper usage [13]. Based on their findings, the appropriate use of ETp is in hydrology, meteorology, and climatology, and as for ETo, the proper usage is in agronomy, agriculture, irrigation, and ecology. For that reason, ETo has been chosen as the base of collecting parameters because the research field is agriculture for irrigation management. Xiang et al.'s (2020) statement was also supported by Muhammad et al.'s (2019) findings which stated that "The accepted method for estimating water loss is by applying Reference Evapotranspiration (ETo)." [3][13].

ETo is a method scientifically developed by the Food and Agriculture Organization (FAO) of the United Nations to estimate the water loss rate in the form of an evapotranspiration rate. ETo is also designed for grass and alfalfa crops and is unfortunately unsuitable for other vegetation types. FAO developed a model that has been a standard benchmark: the FAO-56 Penman-Monteith (PMF56) model.

As for estimating the evapotranspiration rate using the ETo model, FAO-56 Penman-Monteith (PMF56) model is the most accurate model in estimating the evapotranspiration rate [3][8]. PMF56 is a standard benchmark model that FAO specifically developed to estimate the ETo rate for grass crops. In their findings, Muhammad et al. (2019) supported it, ranking number one with the best evapotranspiration rate [3]. However, due to its high data requirements, some researchers opt for alternative ETo, which results in several other researchers developing existing models or finding which alternative model suits their research's location and geography while considering data availability.

There were several attempts to find an alternative model to be applied. For example, for Rio de Janeiro, Brazil, Fernandes et al. (2012) experimented to find which six alternative ETo models have the nearest evapotranspiration rate estimation to PMF56 and found that the Hargreaves, Priestley & Taylor (P&T), and Makkink model had the closest to the PMF56 [7]. The ETo model evaluation was done based on the coefficient of determination (R2), root-mean-square-error (RMSE), relative error (RelRMSE), and index of agreement (d).

For the Huai River Basin, eastern China, Li et al. (2018) experimented to find which 13 alternative ETo models have the nearest evapotranspiration rate estimation to PMF56 and found the Valiantzas3 model had the closest to the PMF56 considering if the parameter needed is not an issue [14]. However, Valiantzas1 and Valiantzas2 were recommended from April-October in Huai River Basin and other similar regions. The ETo model evaluation was done based on comparative analysis based on the PMF56 result to the other 13 ETo models and statistical metrics based on relative root-mean-square error (RRMSE), mean absolute error (MAE), and the Nash–Sutcliffe coefficient (NS).

For the Bangladeshi region, Islam & Alam (2021) experimented to find which 15 alternative ETo models have the nearest evapotranspiration rate estimation to PMF56 and found that the Abtew model had the closest to the PMF56 [8]. The ETo model evaluation was based on RRMSE, MAE, and NS.

For Peninsular Malaysia, Muhammad et al. (2019) experimented to find which 31 alternative ETo models have the nearest evapotranspiration rate estimation to PMF56 and found that the P&T model had the closest to PMF56. Instead of ranking out the ETo model based on MAE, MSE, and RMSE, their research ranked out from the best to the worst by heat scatter plots, four types of statistical metrics, and compromise programming of frequency of occurrence. Another strong reason was that the research was conducted in Peninsular Malaysia, where Ipoh, Perak is located in this thesis study. So, in theory, it is more accurate in ranking the ETo models.

Their findings showed that the best ETo types are the combination-based PMF56 model, followed by the radiation-based Priestley and Taylor (P&T) model, mass transfer-based Dalton model, and temperature-based Ivanov model, respectively. Muhammad et al. (2019) discovered that the temperature-based P&T model was the best apart from the PMF56 model acting as the benchmark [3]. They also found that the P&T model is the most suitable replacement for PMF56. The advantage of the PMF56 model over the P&T model is that it has the best estimation rate and period. The primary disadvantage of the PMF56 model is its data requirements, which need more parameters and data than the P&T model. The PMF56 model needs five parameters (temperature, solar radiation, relative humidity, wind speed, and saturated vapour pressure). The P&T model needs only three parameters (mean air temperature, solar radiation, and relative humidity). These are the advantages of the P&T model over the PMF56 model on the data requirements, and the P&T model is the next best model apart from other 29 ETo models such as the Hargreaves model, Turc model, etc.

2.2 Artificial Intelligence (AI)

Researchers favoured RNN and its variants, such as LSTM and GRU, because it is suitable for dealing with time series problems in retaining the iteration state by producing an output as a new input for the next iteration [15]. RNN has low computational complexities, which provide accurate predictions with a short time series [15]. However, RNN had one disadvantage: the vanishing gradient problem in the long run, which the LSTM overcame. LSTM can overcome the long-term dependencies which RNN suffers from handling [16]. As a result, the LSTM is preferred for this study, even though LSTM also has its flaws, such as it consumes high computational resources when used. LSTM is more complex than RNN because LSTM has multiple memory cells with typical LSTM that have nearly four times more parameters than RNN [16][17].

When comparing RNN and its variants, LSTM and GRU, it was found that GRU performed the best, followed by LSTM in second and RNN in last. It was found by Alqushaibi et al. (2020) when comparing the RNN, LSTM, and GRU using MSE, RMSE, and MAE, GRU scored the best with 0.0116, 0.1077, 0.0821, while LSTM scored 0.0121, 0.1102, 0.0121, and RNN scored 0.013, 0.114, 0.0884, respectively even though the result were quite near to each other with 0.001 different which might not be significant [16].

Other researchers also found that when comparing GRU with LSTM prediction, it was found that GRU performed better than LSTM. For instance, GRU outperformed LSTM with hybridizing AdaBoost to predict daily crude oil prices based on MAE with 1.4164 compared to 1.6374, RMSE with 2.4602 compared to 3.0161, SI with 0.0322 compared to 0.0395, MAPE with 0.0354 compared to 0.0540, and WMAPE with 0.0247 compared to 0.0285 [18]. GRU projects a 23.45% higher accuracy ratio, 27.69% recall ratio, 26.95% F1 ratio, and 29.29% faster processing time than LSTM [19]. GRU better predicted the Bitcoin price with 3.97% in MAPE and 381.34 in RMSE [20].

However, there is not much research showing that the LSTM outperforms the GRU but otherwise. For instance, LSTM is better than GRU in truck traffic flow, with a 4.10% higher [21]. LSTM performed the best in speech recognition compared to RNN and GRU, but GRU managed to come quite close to LSTM in less time [22].

LSTM and GRU have their advantages and disadvantages in a certain case study. However, in most cases, GRU had the upper hand over the LSTM. Furthermore, GRU was explicitly developed as an improved and simpler version of LSTM, so there is a need to improve LSTM further. In conclusion, when compared between RNN, LSTM, and GRU, it was found that GRU performed better than LSTM, and LSTM performed better than RNN [16].

2.3 Long Short-Term Memory (LSTM) Hyperparameter Settings and Hybridization

2.3.1 LSTM Hyperparameter Settings Method

Hyperparameter optimization aims to adopt a method that efficiently ensures the optimization at the highest optimal value [23]. Optimization of a prediction model can be divided into two ways, i) manual optimization by trial and error, and ii) applying existing optimization methods such as Bayesian and Random Forest. To decide which tuning to be tuned, firstly, we need to know how many tuning is available for the LSTM.

We decided to focus on three simple hyperparameter settings, which are i) number of hidden layers, ii) training to testing ratio, and iii) epoch size. The reason was that it involves only tuning the number without involving complex mathematical calculations. For example, we are substituting the number of hidden layers between single hidden layers to double hidden layers by adding another hidden layer. The same method can be applied to testing ratio and epoch size training.

2.3.1.1 Number of Hidden Layers

The hidden layer is one of the hyperparameters available in the LSTM. By default, LSTM has a single hidden layer, but it can be added as much as needed, but how much is needed?

Adeyemi et al. (2018) claimed that adding more than one hidden layer is not advisable since it reduces prediction accuracy [24]. However, there are instances where double hidden layers improve the prediction accuracy [25]. Hence, the experiment must be run to compare the results and justify the findings. So, we will conduct the experiments with a single layer [18], double layers, and quad layers to check whether the claims by Adeyemi et al. (2018) can be applied to this research [24].

2.3.1.2 Number of Training to Testing

The training-to-testing ratio is one of the hyperparameters available in the LSTM. It is used to train the model and test the training outcome.

The training and testing ratio can significantly impact prediction accuracy. It can improve the accuracy or turn out worse than the existing benchmark. If the training ratio is lower than the testing ratio, the accuracy will be high because it suffers from underfitting. If the training ratio is too high, the accuracy will be low because it suffers from overfitting. So, the experiment is done to find which training-to-testing ratios balance is the best fit. As for the experiment, there are multiple training-to-testing ratios, so we selected three types of ratios which are 70:30 [24], 75:25 [18], and 80:20 [26][27].

2.3.1.3 Number of Epochs

Epoch is one of the hyperparameters available in the LSTM. Epoch is the number of times the algorithm will work through the entire training dataset before displaying the result [28]. It can be either 10 or 100 epochs before finding the perfect fit.

Wang et al. (2022) tested the prediction model with 50 epochs. They discovered that the epoch only ran the best up to 25 epochs. So, finding just the right number of epochs is crucial because it can also affect prediction accuracy. As for the experiment, there is a multiple epochs range, so we selected three epochs which are 25 epochs [21], 50 epochs [21], and 100 epochs [18].

2.3.2 LSTM Hybridizing Method

In a study conducted by Yin et al. (2020), Hybridization between Bidirectional-LSTM (Bi-LSTM) with Artificial Neural Network (ANN) results showed the best forecast performance for short-term daily ETo with three meteorological data (maximum temperature, minimum temperature, and sunshine duration). The method of Hybridization done was Bi-LSTM will produce the forecast, and ANN will process it into new output. The ANN consists of the input, hidden, and output layers. The number of neurons was determined by 12 input and 12 output neurons divided by 12 months and 36 hidden neurons based on trial-and-error with an ANN network structure of 12-36-12. The results were validated using RMSE, MAE, R, and NSE in 1-day, 4-day and 7-day lead times. For 1-day lead time, RMSE = 0.159, MAE = 0.039, R = 0.992, and NSE = 0.988. As for 4-day lead time, RMSE = 0.247, MAE = 0.075, R = 0.972, and NSE = 0.985. As for 7-day lead time, RMSE = 0.323, MAE = 0.089, R = 0.943, and NSE = 0.982. There was a slight decrease in performance as the number of day leads increased.

In a study conducted by Ferreira and da Cunha (2020), Hybridization between Convolution Neural

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Network-LSTM (CNN-LSTM) results showed that it performed slightly better than machine learning models, Artificial Neural Network (ANN) and Random Forest (RF) [30]. CNN-LSTM2 performed better for the local scenario among the hybridization models, and CNN-LSTM3 performed better in the regional scenario. It was recommended to apply the local model, CNN-LSTM2. The models were developed based on three input data: 1) lagged ETo; 2) lagged ETo + day of the year each step of the time lag considered; and 3) lagged ETo + day of the year of each step of the time lag considered + lagged meteorological variables used to compute ETo based on PMF56 model (maximum and minimum temperature, maximum and minimum relative humidity, wind speed, solar radiation, and extraterrestrial radiation). The results were validated using RMSE, where CNN-LSTM2 scored 0.87 while CNN-LSTM3 scored 0.88.

In a study conducted by Sharma et al. (2021), Hybridization between Convolution-LSTM (Conv-LSTM) and Convolution Neural Network-LSTM (CNN-LSTM) showed that both models outperform Hargreaves, Makkink, and Ritchie ETo models, and Conv-LSTM performs the best among ETo models [31]. The hybrid models applied daily maximum temperature (Tmax), minimum temperature (Tmin), wind speed measured at the height of 2 m (U2), solar radiation (Rs), relative humidity (Rh), vapour pressure (Vp), and sunshine hours (Ssh) data.

2.4 Summary

In conclusion, deep learning models possess superiority over machine learning models and the ETo models [24, 31]. To develop the prediction model, we need to decide what parameters to be selected. Determining the best parameters to be chosen is a challenge since a wide range of parameters can be selected. Applying the ETo model as the basis of parameter selection eliminates the unwanted or unnecessary parameters which can affect the prediction accuracy. Hence, using the alternative ETo model, the P&T model is being done by considering the data availability and the findings of Muhammad et al. (2019). The LSTM model can be developed based on the P&T model parameters. As for improving the LSTM model, Hyperparameter Settings with GRU Hybridization had possibilities that could improve the prediction accuracy.

3. METHODOLOGY

3.1.1 Parameter Selection based on Reference Evapotranspiration (ETo)

In a study conducted by Ferreira and da Cunha (2020), Hybridization between Convolution Neural Network-LSTM (CNN-LSTM) results showed that it performed slightly better than machine learning models, Artificial Neural Network (ANN) and Random Forest (RF) [30]. CNN-LSTM2 performed better for the local scenario among the hybridization models, and CNN-LSTM3 performed better in the regional scenario. It was recommended to apply the local model, CNN-LSTM2. The models were developed based on three input data: 1) lagged ETo; 2) lagged ETo + day of the year each step of the time lag considered; and 3) lagged ETo + day of the year of each step of the time lag considered + lagged meteorological variables used to compute ETo based on PMF56 model (maximum and minimum temperature, maximum and minimum relative humidity, wind speed, solar radiation, and extraterrestrial radiation). The results were validated using RMSE, where CNN-LSTM2 scored 0.87 while CNN-LSTM3 scored 0.88.

PMF56 model parameters, according to Muhammad et al. (2019), are [3]:

$$ET\text{o} = \frac{0.408(Rs - G) + \gamma \frac{900}{Tmean + 273} u^2{}_{\text{m}}(e_{\text{s}} - e_{\text{a}})}{\Delta + \gamma(1 + 0.34u^2{}_{\text{m}})} \quad (1)$$

Table 1: PMF50 Model

Symbol	Description
$ET\text{o}$	FAO-56 Penman-Monteith evapotranspiration in (mm/day)
Rs	Solar radiation (MJ/m²/day)
G	Soil heat flux density (MJ/m²/day)
$Tmean$	Air temperature at 2 m height (°C)
$u_{2\text{m}}$	Wind speed at 2 m height (m/s)
e_{s}	Saturation vapour pressure (kPa)
e_{a}	Actual vapour pressure (kPa)
$e_{\text{s}} - e_{\text{a}}$	Saturation vapour pressure deficit (kPa)

Δ	Slope of the saturation vapour pressure-temperature curve (kPa/°C)
γ	The psychrometric constant (kPa/°C)

P&T model parameters, according to Muhammad et al. (2019), are [3]:

$$ET\text{o} = \alpha\left(\frac{\Delta}{\Delta+\gamma}\right)\frac{Rn}{\lambda} \tag{2}$$

Table 2: P&T Model

Symbol	Description
$ET\text{o}$	Priestley and Taylor evapotranspiration in units of (mm/day)
α	Empirical Constant (1.26) - Muhammad et al. (2019)
Δ	Slope Vapour Pressure (Δ) for different Temperature (T) in (kPa/°C)
γ	Psychometric Constant (γ) for different Altitude (z) in (kPa/°C)
Rs	Solar radiation (MJ/m²/day)
λ	Latent heat of vapourization (λ) at different Temperature (T) in (MJ/kg)

In theory, the higher the number of parameters, the better the evapotranspiration rate estimation, as shown by the PMF56 model (Muhammad et al., 2019) [3]. It was the best ETo model since it combines all the needed parameters, significantly impacting the evapotranspiration rate. PMF56 model is the most accurate in estimating the evapotranspiration rate to date. One of the reasons was that it has the highest data requirement compared to the other ETo models covering a more significant dimension and depth, resulting in a better ETo rate.

However, not every parameter is readily available for use, and this situation is the same as the Ipoh location. Since this research will predict the next irrigation, ready and easily accessible data will be applied. We decided to opt for an alternative model, the P&T model.

We decided on the P&T model without compromising any substantial difference because of its data requirement based on Muhammad et al.'s (2019) recommendation [3]. The P&T model was explicitly developed for humid climates, in which Peninsular Malaysia is suitable for the model. According to their findings, the P&T model requirement is not as high as the PMF56 model and is ranked the second-best equation.

With only three parameters, the P&T model can achieve the next-best result to the PMF56 model. The P&T model also reduces the computation time since less data means less time needed to compute. In conclusion, we decided to implement the P&T model due to its data requirements.

In order to develop the prediction model, daily meteorological data were gathered from Jabatan Meteorologi Malaysia from 1st January 2015 until 31st December 2019. The parameters were based on the P&T model for parameter selection.

3.1.2 Proposed Model Development

The proposed model would combine the Hyperparameter Settings with Hybridization. Hyperparameter Settings will consist of three hyperparameter settings which are i) the hidden layer/s, ii) training to testing ratio, and iii) epoch size. As for Hybridization, the hybrid algorithm of choice would be the Gated Recurrent Unit (GRU).

For hidden layer/s, we choose three numbers of hidden layer/s, which are i) single, ii) double, and iii) quadruple. As for the training and testing ratio, we choose three sets of ratios which are i) 70:30, ii) 75:25, and iii) 80:20. As for the epoch size, we choose i) 25, ii) 50, and iii) 100.

For GRU, we will incorporate it into the double and quadruple hidden layers by positioning the GRU hidden layer between the first and the last hidden layer for double hidden layers and added into the second and the third hidden layer for quadruple hidden layers.

Both methods will be combined and made up into 153 cases cross. Below are all the cross available between Vanilla Settings, Hyperparameter Settings, and Hyperparameter Settings with Hybridization.

Table 3: Combination Settings

Case No	Hidden Layer/s	GRA Layer	Training to Testing	Epochs
1	Single		70:30	25
2				50
3				100
4			75:25	25
5				50
6				100
7			80:20	25
8				50

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

9				100
10	Double		70:30	25
11				50
12				100
13			75:25	25
14				50
15				100
16			80:20	25
17				50
18				100
19	Quadruple		70:30	25
20				50
21				100
22			75:25	25
23				50
24				100
25			80:20	25
26				50
27				100
28	Double	1st	70:30	25
29				50
30				100
31			75:25	25
32				50
33				100
34			80:20	25
35				50
36				100
37	Double	2nd	70:30	25
38				50
39				100
40			75:25	25
41				50
42				100
43			80:20	25
44				50
45				100
46	Quadruple	1st	70:30	25
47				50
48				100
49			75:25	25
50				50
51				100
52			80:20	25
53				50
54				100
55	Quadruple	2nd	70:30	25
56				50
57				100
58			75:25	25
59				50
60				100
61			80:20	25
62				50
63				100
64	Quadruple	3rd	70:30	25
65				50
66				100
67			75:25	25
68				50
69				100
70			80:20	25
71				50
72				100
73	Quadruple	4th	70:30	25
74				50
75				100
76			75:25	25
77				50
78				100
79			80:20	25
80				50
81				100
82	Quadruple	1st, 2nd	70:30	25
83				50
84				100
85			75:25	25
86				50
87				100
88			80:20	25
89				50
90				100

91	Quadruple	1st, 3rd	70:30	25
92				50
93				100
94			75:25	25
95				50
96				100
97			80:20	25
98				50
99				100
100	Quadruple	1st, 4th	70:30	25
101				50
102				100
103			75:25	25
104				50
105				100
106			80:20	25
107				50
108				100
109	Quadruple	2nd, 3rd	70:30	25
110				50
111				100
112			75:25	25
113				50
114				100
115			80:20	25
116				50
117				100
118	Quadruple	2nd, 4th	70:30	25
119				50
120				100
121			75:25	25
122				50
123				100
124			80:20	25
125				50
126				100
127	Quadruple	3rd, 4th	70:30	25
128				50
129				100
130			75:25	25
131				50
132				100
133			80:20	25
134				50
135				100
136	Quadruple	1st, 2nd, 3rd	70:30	25
137				50
138				100
139			75:25	25
140				50
141				100
142			80:20	25
143				50
144				100
145	Quadruple	2nd, 3rd, 4th	70:30	25
146				50
147				100
148			75:25	25
149				50
150				100
151			80:20	25
152				50
153				100

3.1.3 Evaluation Metrics

1.1.1 Mean Absolute Error (MAE)

Mean Absolute Error (MAE) is the average of the absolute difference between the actual and predicted values in the dataset. It measures the average of the residuals in the dataset—the lower the value, the better the results.

Equation 1: Mean Absolute Error (MAE)

$$MAE = \frac{1}{N}\sum_{i=1}^{N} |y_i - \hat{y}| \quad (3)$$

1.1.2 Mean Square Error (MSE)

Mean Square Error (MSE) is the average squared difference between the original and predicted values in the data set. It measures the variance of the residuals—the lower the value, the better the results.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Equation 2: Mean Square Error (MSE)

$$MSE = \frac{1}{n}\sum_{i=1}^{N}(y_i - \hat{y})^2 \quad (4)$$

1.1.3 Root Mean Square Error (RMSE)

Root Mean Square Error ($RMSE$) is the square root of Mean Squared error. It measures the standard deviation of residuals—the lower the value, the better the results.

Equation 3: Root Mean Square Error (RMSE)

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{N}(y_i - \hat{y})^2} \quad (5)$$

1.1.4 Coefficient of Determination (R^2)

Coefficient of Determination (R^2) is the proportion of the variance in the dependent variable explained by the linear regression model—the higher the value to 1, the better the results.

Equation 4: Coefficient of Determination (R^2)

$$R^2 = \frac{Sum\ of\ Squared\ Regression\ (SSR)}{Total\ Variation\ (SST)} \quad (6)$$

4. RESULTS

We pick the first rank in LSTM Hyperparameter Settings and the top three ranks in LSTM Hyperparameter Settings with Hybridization to be compared with the benchmark LSTM, the Vanilla Settings.

Below are the comparison results:

Table 4: Vanilla Settings vs Hyperparameter Settings

Case No.	MAE	MSE	RMSE	R^2
1	0.27523	0.11559	0.33998	0.88278
9	0.0712	0.00861	0.09278	0.99047
Improvement (%)	74.1307	92.5513	72.7102	12.199

Table 4 shows the comparison between Vanilla Settings with Hyperparameter Settings. It showed that with Hyperparameter Settings, the result improved significantly with more than 72% in MAE, MSE, and RMSE. As for R^2, the improvement was at 12%.

Table 5: Vanilla Settings vs Hyperparameter Settings with Hybridization

Case No.	MAE	MSE	RMSE	R^2
1	0.27523	0.11559	0.33998	0.88278
90	0.06606	0.00759	0.0871	0.9916
Improvement (%)	75.9983	93.4337	74.3808	12.327

Table 5 shows the comparison between Vanilla Settings with Hyperparameter Settings. It showed that with Hyperparameter Settings with Hybridization, the result improved significantly with more than 74% in MAE, MSE, and RMSE. As for R^2, the improvement was at 12%.

Table 6: 1st Rank Hyperparameter Settings vs 1st Rank Hyperparameter Settings with Hybridization

Case No.	MAE	MSE	RMSE	R^2
9	0.0712	0.00861	0.09278	0.99047
99	0.0626	0.00667	0.0817	0.99261
Improvement (%)	12.0787	22.5319	11.9422	0.21606

Table 6 compares the 1st rank in Hyperparameter Settings with the 1st rank in Hyperparameter Settings with Hybridization. It showed that with Hyperparameter Settings with Hybridization, the result improved significantly with more than 11% in MAE, MSE, and RMSE. As for R^2, the improvement was at 0.2%.

Table 7: 1st Rank Hyperparameter Settings vs 2nd Rank Hyperparameter Settings with Hybridization

Case No.	MAE	MSE	RMSE	R^2
9	0.0712	0.00861	0.09278	0.99047
36	0.06446	0.00706	0.084	0.99219
Improvement (%)	9.46629	18.0023	9.46325	0.17366

Journal of Theoretical and Applied Information Technology
30th November 2022. Vol.100. No 22

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Table 7 compares the 1st rank in Hyperparameter Settings with the 2nd rank in Hyperparameter Settings with Hybridization. It showed that with Hyperparameter Settings with Hybridization, the result improved significantly with more than 9% in MAE, MSE, and RMSE. As for R^2, the improvement was at 0.1%.

Table 8: 1st Rank Hyperparameter Settings vs 3rd Rank Hyperparameter Settings with Hybridization

Case No.	MAE	MSE	RMSE	R^2
9	0.0712	0.00861	0.09278	0.99047
90	0.06606	0.00759	0.0871	0.9916
Improvement (%)	7.2191	11.8467	6.12201	0.11409

Table 8 compares the 1st rank in Hyperparameter Settings with the 3rd rank in Hyperparameter Settings with Hybridization. It showed that with Hyperparameter Settings with Hybridization, the result improved significantly with more than 6% in MAE, MSE, and RMSE. As for R^2, the improvement was at 0.1%.

All experiment cases combined between LSTM for i) Vanilla Settings, ii) Hyperparameter Settings, and iii) Hyperparameter Settings with Hybridization indicate that Hybridization did improve the prediction accuracy quite significantly. Even with just Hybridization, the accuracy is on par with the Hyperparameter Settings.

The proposed LSTM model combines Hyperparameter Settings with Hybridization and displays the highest prediction accuracy result. The results in Hyperparameter Settings with Hybridization outperformed the best results in Hyperparameter Settings. Hence, we can conclude that Hybridization significantly improves prediction accuracy.

5. CONCLUSIONS

This research aims to predict water loss using the evapotranspiration rate and prediction model. There will be three components which are i) parameter selections, ii) prediction model selection, and iii) proposed improvement method. The research goal was to improve the prediction accuracy for the evapotranspiration rate; hence the three components were laid out.

Firstly, parameter selections are crucial because many parameters can be chosen. The range can be from a general type available for free to a specific parameter type that can only be obtained from the local meteorological department. This research goal is to predict water loss, so finding what parameters related to water loss are justifiable. The standard method to estimate water loss is applying the Reference Evapotranspiration (ETo) to calculate the evapotranspiration rate. We used Priestley and Taylor (P&T) as the chosen model.

Secondly, to decide what algorithm to be applied for this research. Long Short-Term Memory (LSTM) is the most suitable algorithm for this research based on the extensive literature review. LSTM is better than Recurrent Neural Network (RNN) because it is a better version made from RNN. LSTM offers long-term dependencies and is not vulnerable to the vanishing gradient problem from which RNN suffers.

The results were measured based on the statistical metrics MAE, MSE, RMSE, and R^2. The results were also compared based on the Improvement (%). The result proved that Hybridization improves prediction accuracy. Comparisons were made between LSTM for i) Vanilla Settings, Hyperparameter Settings, and Hyperparameter Settings with Hybridization. The best result was Case 99, where the MAE, MSE, RMSE, and R^2, 0.0626, 0.00667, 0.0817, and 0.99261, respectively. Case 99 combination was quadruple hidden layers with a combination between the LSTM at the second and fourth layers and the GRU at the first and third layers. The training-to-testing ratio was 80:20 ratio, and the epoch size was 100 epochs.

As for the second-ranked and third-ranked cases, Case 36 and Case 90, the only different settings from Case 99 were the numbers of hidden layers. Case 36 has double hidden layers, combining the LSTM at the second hidden layer and the GRU at the first hidden layer. Case 90 has the same number of hidden layers as Case 99 but with different LSTM at the third and fourth hidden layers and GRU at the first and second hidden layers.

The best Hyperparameter Settings only ranked 12 in MAE, MSE, and RMSE while ranking 13 in R^2 for Case 9. The combination was a single hidden layer with an 80:20 ratio and 100 epochs. As for the Vanilla Settings, acting as the benchmark, Case 1 ranked 150 for MSE and 152 for MAE, RMSE, and R^2. It consists of a single hidden layer with a 70:30 ratio and 25 epochs.

The similarity between Case 99, Case 30, Case 90, and Case 9 was in training to testing ratio, 80:20

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

and 100 epochs. It gave the best results for Hyperparameter Settings and Hyperparameter Settings with Hybridization. As for Hyperparameter Settings with Hybridization cases, Case 99, Case 30, and Case 90, another similarity was that the LSTM will always be at the last hidden layer, and GRU will always be at the first hidden layer, producing the best results.

The findings indicate that to improve the existing prediction model, besides researching the most suitable hyperparameter combinations, adding the Hybridization can significantly improve the prediction accuracy. We can see the result when comparing the hyperparameter settings with and without Hybridization.

Even though Hyperparameter Settings did manage to improve the prediction accuracy, selecting only the best might be a challenge. As in the initial stage of this research, the challenge was to decide which hyperparameter settings to and to find which combination suits the best in which combination testing needs to be done. With a deeper literature review, applying the evapotranspiration rate managed to narrow down only the needed parameter based on the evapotranspiration rate model. The parameter based on the model can be applied in prediction model development.

REFERENCES:

[1] Penman, H. L. (1956). Estimating Evaporation. Eos, Transactions American Geophysical Union, 37(1), 43-50.

[2] Lin, H., Chen, Y., Zhang, H., Fu, P., & Fan, Z. (2017). Stronger Cooling Effects Of Transpiration And Leaf Physical Traits Of Plants From A Hot Dry Habitat Than From A Hot Wet Habitat. Functional Ecology, 31(12), 2202-2211.

[3] Gocic, M., & Trajkovic, S. (2014). Analysis Of Trends In Reference Evapotranspiration Data In A Humid Climate. Hydrological Sciences Journal, 59(1), 165-180.

[4] Muhammad, M. K. I., Nashwan, M. S., Shahid, S., Ismail, T. B., Song, Y. H., & Chung, E. S. (2019). Evaluation Of Empirical Reference Evapotranspiration Models Using Compromise Programming: A Case Study Of Peninsular Malaysia. Sustainability, 11(16), 4267.

[5] Xinrui, L., Ping, L., Juncai, S., Yunfei, X., & Xiangning, S. (2021, May). Simulation Of Forest Evapotranspiration Based On Attention-Lstm. In 2021 International Conference On Communications, Information System And Computer Engineering (Cisce) (Pp. 641-648). Ieee.

[6] Kumar, S., Goyal, P., & Sharda, R. (2021, March). Ai Approaches For Estimation Of Reference Evapotranspiration-A Review. In Virtual National Conference On Strategic Reorientation For Climate Smart Agriculture V-Agmet 2021 (Vol. 2021, P. 345).

[7] Gulati, A., & Thakur, S. (2018, January). Smart Irrigation Using Internet Of Things. In 2018 8th International Conference On Cloud Computing, Data Science & Engineering (Confluence) (Pp. 819-823). Ieee.

[8] Fernandes, L. C., Paiva, C. M., & Rotunno Filho, O. C. (2012). Evaluation Of Six Empirical Evapotranspiration Equations-Case Study: Campos Dos Goytacazes/Rj. Revista Brasileira De Meteorologia, 27, 272-280.

[9] Islam, S., & Alam, A. R. (2021). Performance Evaluation Of Fao Penman-Monteith And Best Alternative Models For Estimating Reference Evapotranspiration In Bangladesh. Heliyon, 7(7), E07487.

[10] Ponraj, A. S., & Vigneswaran, T. (2020). Daily Evapotranspiration Prediction Using Gradient Boost Regression Model For Irrigation Planning. The Journal Of Supercomputing, 76(8), 5732-5744.

[11] Djaman, K., Balde, A. B., Sow, A., Muller, B., Irmak, S., N'diaye, M. K., ... & Saito, K. (2015). Evaluation Of Sixteen Reference Evapotranspiration Methods Under Sahelian Conditions In The Senegal River Valley. Journal Of Hydrology: Regional Studies, 3, 139-159.

[12] Shahid, F., Zameer, A., & Muneeb, M. (2020). Predictions For Covid-19 With Deep Learning Models Of Lstm, Gru And Bi-Lstm. Chaos, Solitons & Fractals, 140, 110212.

[13] Yang, S., Yu, X., & Zhou, Y. (2020, June). Lstm And Gru Neural Network Performance Comparison Study: Taking Yelp Review Dataset As An Example. In 2020 International Workshop On Electronic Communication And Artificial Intelligence (Iwecai) (Pp. 98-101). Ieee.

[14] Xiang, K., Li, Y., Horton, R., & Feng, H. (2020). Similarity And Difference Of Potential Evapotranspiration And Reference Crop Evapotranspiration–A Review. Agricultural Water Management, 232, 106043.

[15] Li, M., Chu, R., Islam, A. R. M., & Shen, S. (2018). Reference Evapotranspiration Variation Analysis And Its Approaches Evaluation Of 13 Empirical Models In Sub-Humid And Humid Regions: A Case Study Of The Huai River Basin, Eastern China. Water, 10(4), 493.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

[16] Granata, F., & Di Nunno, F. (2021). Forecasting Evapotranspiration In Different Climates Using Ensembles Of Recurrent Neural Networks. Agricultural Water Management, 255, 107040.

[17] Alqushaibi, A., Abdulkadir, S. J., Rais, H. M., Al-Tashi, Q., & Ragab, M. G. (2020, October). An Optimized Recurrent Neural Network For Metocean Forecasting. In 2020 International Conference On Computational Intelligence (Icci) (Pp. 190-195). Ieee.

[18] Ni, R., & Cao, H. (2020, July). Sentiment Analysis Based On Glove And Lstm-Gru. In 2020 39th Chinese Control Conference (Ccc) (Pp. 7492-7497). Ieee.

[19] Busari, G. A., & Lim, D. H. (2021). Crude Oil Price Prediction: A Comparison Between Adaboost-Lstm And Adaboost-Gru For Improving Forecasting Performance. Computers & Chemical Engineering, 155, 107513.

[20] Yang, S., Yu, X., & Zhou, Y. (2020, June). Lstm And Gru Neural Network Performance Comparison Study: Taking Yelp Review Dataset As An Example. In 2020 International Workshop On Electronic Communication And Artificial Intelligence (Iwecai) (Pp. 98-101). Ieee.

[21] Yamak, P. T., Yujian, L., & Gadosey, P. K. (2019, December). A Comparison Between Arima, Lstm, And Gru For Time Series Forecasting. In Proceedings Of The 2019 2nd International Conference On Algorithms, Computing And Artificial Intelligence (Pp. 49-55).

[22] Wang, S., Zhao, J., Shao, C., Dong, C. D., & Yin, C. (2020). Truck Traffic Flow Prediction Based On Lstm And Gru Methods With Sampled Gps Data. Ieee Access, 8, 208158-208169.

[23] Shewalkar, A. (2019). Performance Evaluation Of Deep Neural Networks Applied To Speech Recognition: Rnn, Lstm And Gru. Journal Of Artificial Intelligence And Soft Computing Research, 9(4), 235-245.

[24] Nakisa, B., Rastgoo, M. N., Rakotonirainy, A., Maire, F., & Chandran, V. (2018). Long Short Term Memory Hyperparameter Optimization For A Neural Network Based Emotion Recognition Framework. Ieee Access, 6, 49325-49338.

[25] Adeyemi, O., Grove, I., Peets, S., Domun, Y., & Norton, T. (2018). Dynamic Neural Network Modelling Of Soil Moisture Content For Predictive Irrigation Scheduling. Sensors, 18(10), 3408.

[26] Alibabaei, K., Gaspar, P. D., & Lima, T. M. (2021). Modeling Soil Water Content And Reference Evapotranspiration From Climate Data Using Deep Learning Method. Applied Sciences, 11(11), 5029.

[27] Fraihat, H., Almbaideen, A. A., Al-Odienat, A., Al-Naami, B., De Fazio, R., & Visconti, P. (2022). Solar Radiation Forecasting By Pearson Correlation Using Lstm Neural Network And Anfis Method: Application In The West-Central Jordan. Future Internet, 14(3), 79.

[28] Wegayehu, E. B., & Muluneh, F. B. (2022). Short-Term Daily Univariate Streamflow Forecasting Using Deep Learning Models. Advances In Meteorology, 2022.

[29] Brownlee, J. (2018). What Is The Difference Between A Batch And An Epoch In A Neural Network. Machine Learning Mastery, 20.

[30] Yin, J., Deng, Z., Ines, A. V., Wu, J., & Rasu, E. (2020). Forecast Of Short-Term Daily Reference Evapotranspiration Under Limited Meteorological Variables Using A Hybrid Bi-Directional Long Short-Term Memory Model (Bi-Lstm). Agricultural Water Management, 242, 106386.

[31] Ferreira, L. B., & Da Cunha, F. F. (2020). Multi-Step Ahead Forecasting Of Daily Reference Evapotranspiration Using Deep Learning. Computers And Electronics In Agriculture, 178, 105728.

[32] Sharma, G., Singh, A., & Jain, S. (2022). A Hybrid Deep Neural Network Approach To Estimate Reference Evapotranspiration Using Limited Climate Data. Neural Computing And Applications, 34(5), 4013-4032.