Comparison the ANN models of Penman-Monteith Potential Evapotranspiration with combination two input of climatological data in Surabaya

D A D Nusantara* and F Nadiar
Department of Civil Engineering, Faculty of Engineering, Universitas Negeri Surabaya, Ketintang, Surabaya, Indonesia

*danayantinusantara@unesa.ac.id

Abstract. In Surabaya, as a part of the equatorial region, determine the rate of daily potential evapotranspiration (PET) turns into a requirement. During the scarcity season, the rate of PET significantly increases for certifying water availability. The PET model founded from several inputs of climatological data that are relative-humidity, wind speed, average daily temperature, and the duration of sun exposure. Modeling a PET through prolonged and complicated steps. To simplify the development of modeling PET, this research using Artificial Neural Network (ANN) based on data-driven modeling with fewer inputs. The PET-ANN model intends to match the PET estimated with Penman-Monteith (PM). This research purpose is learning what the best combination of two climatological data as input. The perform MSE and R on the validating process present how the different results come. The results show the best combination of two climatological data from the entire data set as an input. The conclusion is that the combination of relative humidity and wind speed as an input to the PET-ANN presents the best result than other combinations of climatological data. Besides, it approves that the relative humidity and wind speed as an undoubted input to the PET model even using ANN or not.

1. Introduction
Evapotranspiration is a part of the hydrological cycle that must always be considered when calculating water balance [1,2]. Additionally, for everything that is closely related to the availability of water [3]. For example, in the way of the availability of raw water, water for irrigation [4], water for hydropower operations [5], water for ponds, and others. Evapotranspiration which is a subtracted from the total surface runoff that occurs due to rain, in addition to infiltration or percolation into the soil.

The Evapotranspiration’s amount is the worth of evaporation together with transpiration rates. It more often applied than just figuring evaporation. As another consideration, transpiration occurs in watersheds whose land surface is covered with various vegetation. To calculate the actual evaporation rate obtained by tools such as pan evaporation, while to calculate the actual evapotranspiration rate in an area is difficult [6]. Therefore, for knowing the exact rate that occurs, an empirical approach is carried out, which can be referred to as the potential evapotranspiration (PET) [7].

In recent years, there have been many methods to estimate the magnitude of the PET rate. One of them is the Penman-Monteith (PM) method. The PET rate is estimated from some climatological data, including solar radiation (Rn), temperature (T), wind speed at an altitude of 2 m (u2), saturated vapor
pressure (es), and actual vapor pressure (ea). The required data is not entirely available at BMKG stations in Indonesia. For example, the ea and es data can be obtained by reading the tables that have been issued by the PM. The climatological data that collected at BMKG stations can directly estimate the PET rate using the PM method are average daily temperature, evaporation, rainfall, wind speed, relative humidity, and duration of sun exposure.

Since the 1960s, starting with Penman and Monteith, who discovered an empirical formula in estimating PET rates, this method is still used today [8]. Furthermore, the United Nations Food and Agriculture Organization (FAO) uses the Penman-Monteith (PM) method as a standard calculation of the PET rate [9]. The FAO-PM method has proven suitable for use in the tropical regions, including Indonesia [10].

Calculating the PET rate use the FAO-PM method will go through a sort of complicated process. This method requires some climatological data. Meanwhile, some of the data needed must experience the assumptions and estimates process. It happens because there is no fully complete record at the BMKG station in Indonesia. So the use of data-driven modeling such as Artificial Neural Networks (ANN) is expected to facilitate the tough and long process [11,12].

ANN as a model that is stimulated by the biological system of human brain intelligence. The ANN can classify and recognize a pattern through the learning process [13]. The ANN can also be able to forecasting data that will come based on the learning process of pre-existing data [14, 15]. Application of the ANN models to estimate PET has also been carried out in recent years in the world [12] even in Indonesia [16]. The study showed quite good results compared to conventional calculations [11,12,17].

The fact shows that climatological data as an input FAO-PM PET-ANN models are often not wholly available in some regions in developing countries such as Indonesia, while estimating PET rates is needed. To simplify the development of modeling PET, this research using Artificial Neural Network (ANN) based on data-driven modeling with fewer inputs. The PET-ANN model intends to match the PET estimated with Penman-Monteith (PM). Then in this study, the result comparison of the FAO-PM PET-ANN models with a combination of two climatological data as an input. As a result, the best combination of two climatological data can be found.

2. Methods

ANN modeling executes by using MATLAB software. There is a neural network toolbox, namely NFTOOL (Neural Network Fitting Tool), which can help the PET-ANN modeling process [18, 19]. The study outset with the preparation and determination of input data and target models.

This research conduct using four climatological data, which are average temperature, wind speed, relative humidity, and duration of sun exposure as input in ANN modeling. All that data recorded daily by BMKG Juanda Surabaya station. The total is 1095 for each climatology data collected over three years.

In the modeling ANN, besides having to prepare input variables, target variables must also be ready. The target model of PET-ANN prepares with calculating PET manually using the FAO-PM method. It is intended that the results of the PET-ANN model approach the results of PET-FAO-PM.

The progress of preparing the data in the ANN model is ended by grouping input and targets into three data groups are data for training, validating, and testing. Data for training used as much as 70% of the total amount of data that is about 767 data. As the data for validating and testing, each of them is 15% of the total amount of data, it is about 164 data. The basis of the process of determining the percentage distribution of data above base on the standard of the assistance program used for training must be 70% so that the remaining 30% is for validation and testing.

In ANN modeling, it needs an architectural model consisting of three nodes, namely: input, hidden, and output. This study aims to determine the best combination of two climatological data as input to the model. Caused by there are four climatological data that used, so it will be six architectural PET-ANN models as outlined in Table 1. For the count of hidden nodes, a trial will perform to get the best model with a certain count of hidden nodes on each model. The architecture of the model present in Figure 1 below.
Table 1. The Combination 2-Input of Climatological Data for PET-ANN Model

| Model | The Climatological Data Used As An Input Layer |
|-------|-----------------------------------------------|
| 2C    | average temperature; duration of sun exposure |
| 2D    | wind speed; duration of sun exposure          |
| 2G    | relative-humidity; duration of sun exposure   |
| 2J    | relative-humidity; wind speed                |
| 2N    | average temperature; wind speed              |
| 2O    | average temperature; relative-humidity       |

Figure 1. The Architecture of The PET-ANN Model With Two Combination of Climatological Data As Input.

The next process is to do the learning process of the ANN model by using Levenberg-Marquardt backpropagation. At this stage, data-based modeling will carry out its own learning until it stops. Automatic learning stops when the overall model has increased error. The indication is there is an increase in the value of Mean Square Error (MSE) on the validating process about six times. A model might have a better result with the retraining process. In this study, each architecture of the network model has restrictions of five - four times a retrain process to obtain the best model for that architecture.

The best model in this study can know by the validation process. In that process, the value of MSE (Mean Square Error) will be compared. If some of the models have the same MSE value, then consider the value of R (Regression) generated by the model. For the best model, the MSE-value is close to 0, which means there is almost no error, while the R-value is close to 1, which illustrates the trend similarity between output and target. This process can illustrate in Figure 2 below.

Figure 2. Validating Process on PET-ANN Model
3. Results and discussion
In this study, there are six models, as previously explained, an analysis performs on each MSE and R performance of each model, which the results of MSE and R performance in the validation process of each model present in Table 2.

The best model architecture for each model is different depending on the number of different hidden nodes in each model [17]. For example, Model 2C with temperature and the duration of sun exposure as a combination input of climatological data, the best architecture PET-ANN model is 2-4-1. The architecture of the model means that the PET-ANN model has 2-input nodes, 4-hidden nodes, and 1-output node. The performance of the best Model 2C has an MSE-value 3.05, which is the lowest value among the results of the performance of the MSE other Model 2C that has a different count of hidden nodes. Furthermore, the R-value of the best Model 2C is 0.72. It is the highest R-value among other Model 2C with different architecture.

Another explanation, the Model 2D has the best model with the architecture of the model is 2-2-1, while Model 2G is 2-10-1. It can also show that Model 2J and Model 2N nearly have the same architecture, Model 2J is 2-7-1 whereas Model 2N is 2-8-1. For the last, Model 2O has the best model with the highest count of hidden nodes, so the architecture of the model 2O is 2-12-1.

Table 2. Comparison The Validation Result of MSE-Value and R-Value on PET-ANN Model

| Count Of Hidden Nodes | The Validation Result on Each PET-ANN Model |
|-----------------------|--------------------------------------------|
|                       | 2C  | 2D  | 2G  | 2J  | 2N  | 2O  |
| MSE                  |     |     |     |     |     |     |
| R                    | 0.67| 0.81| 0.79| 0.85| 0.77| 0.86|
| MSE                  | 2.46| 2.02| 1.82| 1.67| 1.58| 1.75|
| MSE                  | 1.55| 1.25| 0.84| 0.85| 0.74| 0.84|
| R                    | 0.87| 0.90| 0.93| 0.92| 0.93| 0.93|
| MSE                  | 1.23| 1.25| 0.93| 0.93| 0.93| 0.93|
| MSE                  | 0.90| 0.90| 0.93| 0.92| 0.92| 0.92|
| MSE                  | 2.55| 2.88| 2.67| 2.62| 2.78| 2.92|
| MSE                  | 0.77| 0.76| 0.76| 0.79| 0.79| 0.79|
| MSE                  | 1.58| 1.75| 1.82| 1.82| 1.81| 1.81|
| R                    | 0.86| 0.86| 0.84| 0.85| 0.74| 0.74|
| MSE                  | 3.41| 4.28| 3.05| 2.85| 2.67| 2.88|
| MSE                  | 2.46| 2.02| 1.82| 1.67| 1.58| 1.75|
| MSE                  | 1.55| 1.25| 0.93| 0.93| 0.93| 0.93|
| R                    | 0.87| 0.90| 0.93| 0.92| 0.92| 0.92|
| MSE                  | 1.23| 1.25| 0.93| 0.93| 0.93| 0.93|
| MSE                  | 0.90| 0.90| 0.93| 0.92| 0.92| 0.92|
| MSE                  | 2.55| 2.88| 2.67| 2.62| 2.78| 2.92|
| MSE                  | 0.77| 0.76| 0.76| 0.79| 0.79| 0.79|
| MSE                  | 1.58| 1.75| 1.82| 1.82| 1.81| 1.81|
| MSE                  | 0.86| 0.86| 0.84| 0.85| 0.74| 0.74|
| MSE                  | 3.41| 4.28| 3.05| 2.85| 2.67| 2.88|
| MSE                  | 2.46| 2.02| 1.82| 1.67| 1.58| 1.75|
| MSE                  | 1.55| 1.25| 0.93| 0.93| 0.93| 0.93|
| MSE                  | 0.90| 0.90| 0.93| 0.92| 0.92| 0.92|
| MSE                  | 2.55| 2.88| 2.67| 2.62| 2.78| 2.92|
| MSE                  | 0.77| 0.76| 0.76| 0.79| 0.79| 0.79|
| MSE                  | 1.58| 1.75| 1.82| 1.82| 1.81| 1.81|
| MSE                  | 0.86| 0.86| 0.84| 0.85| 0.74| 0.74|
| MSE                  | 3.41| 4.28| 3.05| 2.85| 2.67| 2.88|
| MSE                  | 2.46| 2.02| 1.82| 1.67| 1.58| 1.75|
| MSE                  | 1.55| 1.25| 0.93| 0.93| 0.93| 0.93|
| MSE                  | 0.90| 0.90| 0.93| 0.92| 0.92| 0.92|
| MSE                  | 2.55| 2.88| 2.67| 2.62| 2.78| 2.92|
| MSE                  | 0.77| 0.76| 0.76| 0.79| 0.79| 0.79|
| MSE                  | 1.58| 1.75| 1.82| 1.82| 1.81| 1.81|
| MSE                  | 0.86| 0.86| 0.84| 0.85| 0.74| 0.74|

After determining the best architecture of each model with the same combination input climatological data, then the next step is to compare the results of MSE-value and R-value of the best model with the different combination inputs. Based on Table 2 above, it shows that Model 2C has the worst result than other models. The next worst model is Model 2N, which has MSE-value 2.47. Following that, Model 2D, Model 2G, and Model 2O have slightly different of MSE-value. Their MSE-value is about more than 1.0. For Model 2J, it is the leading model that beats another model. Model 2J, which has wind speed and relative-humidity as input on the PET-ANN model, is the only model that has MSE-value less than 1.0 and it has the highest R-value.

4. Conclusion
This research concludes with the result is the combination of relative humidity and wind speed as an input to the PET-ANN model presents the best result than other combinations of climatological data. Besides, it approves that the relative humidity and wind speed as an undoubted input to the PET model even using ANN or not. The results can be varied in other places if where has different characteristic to
the Surabaya as the origin area of the climatological data applied for the development of the PET-ANN model. Hence, the results become helpful for another modeling with ANN with limitable of the climatological data.

Acknowledgments
Gratitude to BMKG Juanda Surabaya for contributing to the climatological data applied in this research.

References
[1] Herman M R, Nejadhashemi A P, Abouali M, Hernandez-Suarez J S, Daneshvar F, Zhang Z, Anderson M C, Sadeghi A M, Hain C R and Sharifi A 2018 Evaluating the role of evapotranspiration remote sensing data in improving hydrological modeling predictability J. Hydrol. 556 39–49
[2] Liu W, Wang L, Zhou J, Li Y, Sun F, Fu G, Li X and Sang Y F 2016 A worldwide evaluation of basin-scale evapotranspiration estimates against the water balance method J. Hydrol. 538 82–95
[3] Dezsi Ş, Mindrescu M, Petrea D, Rai P K, Hamann A and Nistor M M 2018 High-resolution projections of evapotranspiration and water availability for Europe under climate change Int. J. Climatol. 38 3832–41
[4] Xu J, Liu X, Yang S, Qi Z and Wang Y 2017 Modeling rice evapotranspiration under water-saving irrigation by calibrating canopy resistance model parameters in the Penman-Monteith equation Agric. Water Manag. 182 55–66
[5] Salema M G 2012 Water and hydropower for sustainable development of qattara depression as a national project in Egypt Energy Procedia 18 994–1004
[6] McMahon T A, Peel M C, Lowe L, Srijanthan R and McVicar T R 2013 Estimating actual, potential, reference crop and pan evaporation using standard meteorological data: A pragmatic synthesis Hydrol. Earth Syst. Sci. 17 1331–63
[7] Li S, Kang S, Zhang L, Zhang J, Du T, Tong L and Ding R 2016 Evaluation of six potential evapotranspiration models for estimating crop potential and actual evapotranspiration in arid regions J. Hydrol. 543 450–61
[8] Debnath S, Adamala S and Raghuvanshi N S 2015 Sensitivity Analysis of FAO-56 Penman-Monteith Method for Different Agro-ecological Regions of India Environ. Process. 2 689–704
[9] Córdova M, Carrillo-Rojas G, Crespo P, Wilcox B and Célleri R 2015 Evaluation of the Penman-Monteith (FAO 56 PM) Method for Calculating Reference Evapotranspiration Using Limited Data Mt. Res. Dev. 35 230
[10] Manik T K, Rosadi R B and Karyanto A 2012 Evaluation of Penman-Monteith Method in Estimating Standard Evapotranspiration (ET0) in Lowland Area of Lampung Province, Indonesia J. Keteknikan Pertan. 182 121–8
[11] Sharma S and Regulvar D G 2016 Prediction of Evapotranspiration by Artificial Neural Network and Conventional Prediction of Evapotranspiration by Artificial Neural Network and Conventional Methods Int. J. Eng. Res. 5 184–7
[12] Antonopoulos V Z, Giannini S K and Antonopoulos A V 2016 Artificial neural networks and empirical equations to estimate daily evaporation: application to Lake Vegoritis, Greece Hydrol. Sci. J. 61 2590–9
[13] Moasheri, S. Ali. , Afrasiabi, P. , sarani, Sh. , Sarani N 2012 Estimating of Reference Evapotranspiration by Using Artificial Neural Networks Int. Conf. Transp. Environ. Civ. Eng. August 25-26, 2012 80–4
[14] Alves W B, Rolim G D S and Aparecido L E D O 2017 Reference evapotranspiration forecasting by artificial neural networks Eng. Agric. 37 1116–25
[15] Sheikh S K and Unde M G 2012 Short-term load forecasting using ANN technique Int. J. Eng. Sci. Emerg. Technol. 1 97–107
[16] Nusantara D A D, Nadiar F and Ansori M B 2019 The Behaviour Of Various Climates Data As A Single Input To The Ann Model Of Daily Potential Evapotranspiration’s Penman-Monteith Reka Buana J. Ilm. Tek. Sipil dan Tek. Kim. 4 63

[17] Laaboudi A, Mouhouche B and Draoui B 2012 Neural network approach to reference evapotranspiration modeling from limited climatic data in arid regions Int. J. Biometeorol. 56 831–41

[18] Tayyab M, Zhou J, Zeng X and Adnan R 2016 Discharge Forecasting By Applying Artificial Neural Networks At The Jinsha River Basin, China Eur. Sci. Journal, ESJ 12 108

[19] Al Shamisi M H, H. A and N. Hejase H A 2011 Using MATLAB to Develop Artificial Neural Network Models for Predicting Global Solar Radiation in Al Ain City – UAE Engineering Education and Research Using MATLAB ed A Assi (United Arab Emirates: InTech) pp 219–38