Using ANN to Evaluate the Climate Data that High Affect on Calculate Daily Potential Evapotranspiration with Modified-Penman Method in The Tropical Regions

DAD Nusantara*, F Nadiar
Department of Civil Engineering, Faculty of Engineering, Universitas Negeri Surabaya, Ketintang, Surabaya, Indonesia.

*danayantinusantara@unesa.ac.id

Abstract. In Indonesia, as part of tropical regions, calculating the amount of daily potential evapotranspiration (PET) becomes significant to ascertain the water balance. Applied Artificial Neural Network (ANN) as data-driven modeling in the period of Industrial Revolution 4.0 to develop a better world. Using ANN simplifies the process of modeling the various input climate data to count up the amount of daily PET. There six climate data that will be input in the Modified Penman Method, which are average-temperature, wind velocity, rainfall, relative-humidity, evaporation, and the time of solar radiation. Therefore, for further modeling using fewer input data than typically used in this method, so the high affect climatology data must be known. The methodology is to compare the best model, which has input six of the climate data with the model developed by just five of the climate data. The results of this research are the sequence of climate data that influence the daily PET model from the highest to the smallest: wind velocity, relative humidity, the time of solar radiation, temperature, rainfall, evaporation.

1. Introduction
The Potential Evapotranspiration (PET) commit to being interpreting as the estimated count of evapotranspiration (ET). The ET happens to a land area that covered with full crops, which has the probability of lessening the water quantity in the watershed [1]. PET models have been advanced from simple to complex models to measure evapotranspiration that appears. The simple models that have an uncomplicated and easy scheme in a calculation, the example are Lowry-Johnson method, Thornwaite method, and Blaney-Criddle method. They have only two climate data used to figure the PET. The results of the simple models are less factual than others. Otherwise, the results display from the complex models more decisive [2].

The Penman-Monteith model is distinguished for being the most definite model; however, its demands an entire climate data set that is not always securable at all weather stations, notably in most of the developing countries. That model has been standard for estimating PET in the world [3] [4], even in Indonesia as one of the tropical regions. The method used numerous climate data such are average temperature, relative-humidity, wind velocity, rainfall, evaporation, and the duration of solar radiation. There is preceding research to obtain the simple formulas with an identical outcome to the Penman equation for computing the PET even though from an uncomplete-set of the frequently meteorological variables [5] [6]. This research conducted to learn the climate data that strong influence to determine the daily PET with Modified-Penman Method. As a result, climate data that have the highest involved on
the PET considered as a very crucial variable and must be directed as input to figure out the PET. It can be advantageous for further modeling with limited of the climate data that happen in Indonesia.

The ordinary calculating process of the Modified-Penman Method complicated enough leads this research approached using the data-driven modeling for simplifying the process. Besides, in this industrial revolution era 4.0, one of the excuses and challenges that contribute to developing an improved world is applied data-driven modeling to generate the calculation process more practical and more manageable. So, too, the data-driven modeling established from the relationship between inputs and target variables [7] [8]. In hydrology, the practice of data-driven based modeling has begun to be developed in the last two decades [9], the examples are rainfall-runoff modeling [10] [11], stream-flow predicting [12] [13], stream-flow forecasting [14] [15], water-level fluctuations [16], etc. From all of that, data-driven modeling based can be approved and validated as a brand new action to estimate the hydrological data. In the case of the PET, the example of adopting data-driven modeling (e.g., Adaptive Neuro-fuzzy inference system (ANFIS) [17] [18] [19] [20], multivariate adaptive regression splines (MARS) [21], gene expression programming (GEP) [17] [21] [22], artificial neural network (ANN) [23] [24] [20], support vector machine (SVM) [23]) shows that can be applied and have a good modeling results to the PET with Modified-Penman Method.

The ANN is a computational model, which is found and influenced by the living process of the human brain functions. By training the network to grow the connections and rules that integrated within the data, so the neural network models developing. The network always has three layers consist of input, hidden, and output. The system enters the data through the input units classified as an input layer. These data are then continued to reach succeeding layers, including hidden layers in the core, then appear as the output layer. The inputs can be several combinations of the parameter of climate data, which considered to be notable toward forecasting the output [8].

2. Method

This research used daily climate data from Meteorological-Climatological-Geophysical-Agency (BMKG) Juanda, Surabaya, Indonesia, during 2018 as input. They are average-temperature, relative-humidity, rainfall, evaporation, wind velocity, and the duration of solar radiation, which investigated to know the highest affect to the ANN model. The targets are everyday PET measured by BMKG Juanda, Surabaya, Indonesia, during 2018 as input. They are average-temperature, relative-humidity, temperature, wind velocity, rainfall, evaporation, and the duration of solar radiation (hours) that recorded in BMKG Juanda Surabaya.

2.1. Data Collecting

In the data-driven modeling, it is the first move to assure data for the inputs and the targets of the model. Since the modeling process depends on the genuineness of the data, to ensure the source of collecting data are reliable. As previously revealed, the climate data for inputs for the ANN model is the average-temperature (°C), relative-humidity (%), rainfall (mm), evaporation (mm), wind velocity (km/h), and the duration of solar radiation (hours) that recorded in BMKG Juanda Surabaya.

2.2. Building a Network Model

The ANN data-driven modeled by the help of the toolbox neural network on MATLAB that is NFTOOL (Neural Network Fitting Tool). The following process:

2.2.1. Data inputs and targets. Prepare all the climatology data that would be inputs and targets as a workspace so can be straightforwardly imported in MATLAB.

2.2.2. Training, validation, testing. Distribute the data used in this research to be 70% as training, 15% as validation, and 15% as testing.

2.2.3. The architecture of the ANN model. The next step was creating several network model architectures. The first, 6-input ANN model that has all of six climate data used in the modeling. There are average-temperature, relative-humidity, wind velocity, rainfall, evaporation, and the duration of
solar radiation. The 6-1-1 architectural models, namely because of the model, have six inputs, one hidden layer, and one output. Therefore, it created several architectural models that have a number from 1 to 12 of the hidden layers. This step is to consider the best ANN model by comparing the outputs and the targets of several architectural models with a different number of the hidden layer. The best 6-input ANN model will be compared to the 5-input ANN model, so the highest climatology data affect the ANN model can be known. There six of architecture ANN models that only have 5-input of climate data. Each of them has five different input, which has the same target, as follows in Table 1 below:

| Model | The Climatology Data Not Used as An Input Layer |
|-------|-----------------------------------------------|
| 5A    | Temperature                                    |
| 5B    | Relative-humidity                              |
| 5C    | Wind Velocity                                  |
| 5D    | Rainfall                                       |
| 5E    | Evaporation                                    |
| 5F    | The Duration of Solar Radiation                |

2.2.4. Training process. Train the network model reaching out the smallest error, otherwise retrain it continually.

2.2.5. Validation process. Validation obtained the best network model that has the smallest Mean Squared Error (MSE) value, in case 0, it means no error altogether. The MSE value calculated with:

\[ M = \frac{1}{N} \sum_{t=1}^{N} (Y't - Y't)^2 \]

Where Y’t is output from the network model of ANN, Yt is the target as PET calculated with the Modified Penman Method; and N is the number of series data be analyzed. In addition to the MSE value, regression result (R) commits to consider. R is the correlation between input and target when one it means to have a high correlation, then 0 is the opposite.

2.2.6. Testing process and analysis. Compare the results of the network model between 6-input and 5-input so that the climatology data that has a high even smaller affect on the model found. These research results will relate to determining the climatology data used as input in further simplifying the network model.

3. Results and Discussion

During the learning process of the ANN model, the best model achieved by testing all of several architecture models. From the report of the validation process, the architecture model 6-5-1 shows the most dependable results instead of the other 6-input architecture models. On the validation of the architecture model 6-5-1, it has the smallest MSE-value, which is 0.365 and the highest R-value, which is 0.981. That can be noticed in Table 2 below:

| 6-Input ANN Model | TRAINING | VALIDATION | TESTING | ALL |
|-------------------|----------|------------|---------|-----|
|       | MSE      | R          | MSE     | R   | MSE | R   | R   |

Table 2. Result of MSE and R-value of 6-Input Architecture ANN Model
| The Architecture of ANN Model | TRAINING | VALIDATION | TESTING | ALL |
|-----------------------------|----------|-----------|---------|-----|
|                            | MSE      | R         | MSE     | R   |
| 6 - 5 - 1                   | 1.731    | 0.875     | 0.903   |
| 5A                         | 1.333    | 0.908     | 0.889   |
| ΔA                         | -0.398   | 0.034     | -0.014  |
| 5B                         | 2.290    | 0.853     | 0.820   |
| ΔB                         | -0.497   | 0.047     | -0.028  |
| 5C                         | 2.825    | 0.813     | 0.903   |
| ΔC                         | -0.559   | -0.022    | -0.083  |
| 5D                         | 1.093    | -0.061    | 0.000   |
| ΔD                         | 0.912    | -0.126    | 0.006   |
| 5E                         | 1.323    | 0.913     | 0.895   |
| ΔE                         | -0.408   | 0.039     | -0.009  |
| 5F                         | 0.824    | 0.890     | 0.906   |
| ΔF                         | -0.201   | 0.029     | -0.034  |

The 5-input model that has the most significant difference (Δ) MSE and R values with the 6-input model shows the climate data not used in this model has the highest affect on the ANN model. Based on Table 3 above the sequence of the 5-input architecture model are C – B – F – A – D – E. So the order of the climate data that affect in calculating daily PET with Modified-Penman Method from the highest to the smallest are: wind velocity, relative-humidity, the duration of solar radiation, temperature, rainfall, evaporation.
4. Conclusion
The sequence of the climate data that influence the daily PET model in the Tropical Region from the highest to the smallest is wind velocity, relative-humidity, the duration of solar radiation, temperature, rainfall, evaporation. The results can be different in another area that has no similarity characteristic to the origin area of the climate data taken for the development of the ANN Model. Therefore, the results become beneficial for further modeling with ANN with limitless of the climate data.

Acknowledgments
Thanks to BMKG Juanda Surabaya for providing the climate data used in this research.

References
[1] T. A. McMahon, M. C. Peel, L. Lowe, R. Srikanthan, and T. R. McVicar, “Estimating actual, potential, reference crop and pan evaporation using standard meteorological data: A pragmatic synthesis,” Hydrol. Earth Syst. Sci., vol. 17, no. 4, pp. 1331–1363, 2013.
[2] S. Li et al., “Evaluation of six potential evapotranspiration models for estimating crop potential and actual evapotranspiration in arid regions,” J. Hydrol., 2016.
[3] H. S. Division, “MODIFIED PENMAN METHOD FOR POTENTIAL EVAPOTRANSPIRA- ,” vol. 18, pp. 349–365, 1973.
[4] K. Beven, “A sensitivity analysis of the penman--monteith actual evapotranspiration estimates,” vol. 44, pp. 169–190, 1979.
[5] J. D. Valiantzas, “Simplified forms for the standardized FAO-56 Penman-Moneith reference evapotranspiration using limited weather data,” J. Hydrol., 2013.
[6] J. D. Valiantzas, “Simplified limited data penman’s ET formulas adapted for humid locations,” J. Hydrol., vol. 524, pp. 701–707, 2015.
[7] D. P. Solomatine and A. Ostfeld, “Data-driven modelling: some past experiences and new approaches,” J. Hydroinformatics, vol. 10, no. 1, pp. 3–22, 2007.
[8] D. Solomatine, L. M. See, and R. J. Abrahart, “Data-Driven Modelling: Concepts, Approaches and Experiences,” Pract. Hydroinformatics, pp. 17–30, 2008.
[9] V. Nourani, A. H. Baghanam, J. Adamowski, and O. Kisi, “Applications of hybrid Wavelet-Artificial Intelligence models in hydrology. A review,” J. Hydrol., 2014.
[10] P. C. Nayak, B. Venkatesh, B. Krishna, and S. K. Jain, “Rainfall-runoff modeling using conceptual, data driven, and wavelet based computing approach,” J. Hydrol., vol. 493, pp. 57–67, 2013.
[11] M. J. Alizadeh, M. R. Kavianpour, O. Kisi, and V. Nourani, “A new approach for simulating and forecasting the rainfall-runoff process within the next two months,” J. Hydrol., 2017.
[12] N. Noori and L. Kalin, “Coupling SWAT and ANN models for enhanced daily streamflow prediction,” J. Hydrol., vol. 533, pp. 141–151, 2016.
[13] H. Zhang, V. P. Singh, B. Wang, and Y. Yu, “A hybrid data-driven model for forecasting annual streamflow from a socio-hydrological system,” J. Hydrol., vol. 540, pp. 246–256, 2016.
[14] Z. M. Yaseen, A. El-shafie, O. Jaafar, H. A. Afan, and K. N. Sayl, “Artificial intelligence based models for stream-flow forecasting : 2000 – 2015 ,” J. Hydrol., vol. 530, pp. 829–844, 2015.
[15] G. B. Humphrey, M. S. Gibbs, G. C. Dandy, and H. R. Maier, “A hybrid approach to monthly streamflow forecasting : integrating hydrological model outputs into a Bayesian artificial neural network .,” J. Hydrol., 2016.
[16] J. Jeong and E. Park, “Comparative applications of data-driven models representing water table fluctuations,” J. Hydrol., vol. 572, no. March, pp. 261–273, 2019.
[17] J. Shiri, Ô. Kis, G. Landeras, J. Javier, and A. Hossein, “Daily reference evapotranspiration modeling by using genetic programming approach in the Basque Country ( Northern Spain ),” vol. 415, pp. 302–316, 2012.
[18] M. Cobaner, “Evapotranspiration estimation by two different neuro-fuzzy inference systems,” J.
[19] J. Shiri et al., “Global cross-station assessment of neuro-fuzzy models for estimating daily reference evapotranspiration,” *J. Hydrol.*, vol. 480, pp. 46–57, 2013.

[20] A. Guven and O. Kisi, “Monthly pan evaporation modeling using linear genetic programming,” *J. Hydrol.*, no. August, 2013.

[21] S. Mehdizadeh, “Estimation of daily reference evapotranspiration (ET0) using artificial intelligence methods: Offering a new approach for lagged ET0 data-based modeling,” *J. Hydrol.*, 2018.

[22] J. Shiri et al., “Generalizability of Gene Expression Programming-based approaches for estimating daily reference evapotranspiration in coastal stations of Iran,” *J. Hydrol.*, vol. 508, pp. 1–11, 2014.

[23] L. B. Ferreira, F. França, R. A. De Oliveira, E. Inácio, and F. Filho, “Estimation of reference evapotranspiration in Brazil with limited meteorological data using ANN and SVM; a new approach,” *J. Hydrol.*, 2019.

[24] S. Sultan, M. A. Malek, N. Sultan, O. Kisi, and K. Siah, “Extreme Learning Machines : A new approach for prediction of reference evapotranspiration,” *J. Hydrol.*, vol. 527, pp. 184–195, 2015.