

Review

Recent Advances in Evapotranspiration Estimation Using Artificial Intelligence Approaches with a Focus on Hybridization Techniques—A Review

Min Yan Chia, Yuk Feng Huang * , Chai Hoon Koo and Kit Fai Fung

Department of Civil Engineering, Lee Kong Chian Faculty of Engineering and Science, Universiti Tunku Abdul Rahman, Jalan Sungai Long, Bandar Sungai Long, Kajang 43000, Selangor, Malaysia; chiamy94@gmail.com (M.Y.C.); kooch@utar.edu.my (C.H.K.); fkfpanda@hotmail.com (K.F.F.)

* Correspondence: huangyf@utar.edu.my; Tel.: +603-90860288

Received: 3 December 2019; Accepted: 23 December 2019; Published: 10 January 2020



Abstract: Difficulties are faced when formulating hydrological processes, including that of evapotranspiration (ET). Conventional empirical methods for formulating these possess some shortcomings. The artificial intelligence approach emerges as the best possible solution to map the relationships between climatic parameters and ET, even with limited knowledge of the interactions between variables. This review presents the state-of-the-art application of artificial intelligence models in ET estimation, along with different types and sources of data. This paper discovers the most significant climatic parameters for different climate patterns. The characteristics of the basic artificial intelligence models are also explored in this review. To overcome the pitfalls of the individual models, hybrid models which use techniques such as data fusion and ensemble modeling, data decomposition as well as remote sensing-based hybridization, are introduced. In particular, the principles and applications of the hybridization techniques, as well as their combinations with basic models, are explained. The review covers most of the related and excellent papers published from 2011 to 2019 to keep its relevancy in terms of time frame and field of study. Guidelines for the future prospects of ET estimation in research are advocated. It is anticipated that such work could contribute to the development of agriculture-based economy.

Keywords: hydrological process; hybrid model; data fusion; ensemble modeling; data decomposition; remote sensing; bootstrap aggregating; Bayesian modeling; boosting algorithm; nonlinear neural ensemble

1. Introduction

In 2019, the United Nations [1] reported that world population had reached 7.7 billion. In the same report, it was predicted that the world population would continue to grow, and the forecasted figures are 8.5 billion, 9.7 billion and 10.9 billion for the years 2030, 2050, and 2100, respectively. Consequently, agricultural activity that contributes to the population food supply increases progressively and becoming more important. Agricultural activity is regarded as the anthropic activity that depletes the highest amount of water [2]. Therefore, a good estimation of the water cycle can assist in efficient agricultural planning, water catchment and irrigation strategy, and thus optimize the utilization of water. Evapotranspiration (ET), as suggested by the term itself, is the combination of evaporation of water from land and plant surfaces and transpiration from vegetation through the leaves' stomata [3]. ET is a natural event that affects the hydrological cycle, which is believed to be highly complex that involves several nonlinear processes [4]. There are numerous factors that govern the rate of evapotranspiration and these include temperature, solar radiation, air humidity, and wind speed [5].

ET is classified as a physical phenomenon. Hence, the rate of ET can be measured and represented by a numerical value. Traditionally, lysimeters are used to measure ET directly without any assumptions [6]. It works by measuring the rate of water percolation through soil on the basis of mass transfer [7]. Non-weighable lysimeters are normally used for long-term observation, whereas weighable lysimeters can provide readings with greater temporal resolution [8]. It was claimed that lysimeters could provide the most accurate measurement of ET. In fact, several studies that involved ET estimation in the early years, utilized readings of lysimeters as their calibration standards [9,10]. Unfortunately, the construction, maintenance, and use of lysimeters involve high financial burden and ecological footprints. The limited amount of lysimeters also hindered the measurement of ET at distinct locations [3]. In view of such situation, the development of other more convenient estimation tools to estimate ET with higher accuracy and lower cost becomes the mode of choice.

Before the age of artificial intelligence, empirical equations were constantly developed to cater the need of accurately estimating ET. However, in the absence of point measurement, direct acquisition of ET value is virtually impossible. According to Pereira et al. [11], the term of reference evapotranspiration (ET_0) was introduced to overcome the problem. ET_0 is an estimate of the amount of water loss or consumption based on the weather primary effect. In the coefficient-reference system, a crop coefficient (K_c) will be multiplied with the ET_0 to obtain the potential evapotranspiration (PET) for that particular growth period.

Over the years, numerous efforts had been pursued to obtain ET_0 to the extent of higher accuracy with lower computational complexity. Among the vast number of conventional models for ET_0 estimation, some of the most notable empirical approaches are discussed in this review. The Penman–Monteith (PM) model [12], which had its early beginnings in 1948, is regarded as one of the most widely employed models in the estimation of ET_0 . Furthermore, the Food and Agriculture Organisation (FAO) of the United Nations, in their publication “Crop evapotranspiration—Guidelines for computing crop water requirements—FAO Irrigation and Drainage Paper 56” (FAO56 in short), had followed up and revised the calculation of ET_0 and PET based on the PM model [13]. This indirectly made the PM model as a standard in estimating ET_0 , and it was used in a number of research works as a standard for comparison [14–16]. However, the sheer number of parameters and the complexity of their investigation, derivation, and calculation do put paid to its use in a world of increasing knowledge advancement and digital advance.

Ever since, the effort of providing simpler solution than the PM model continued. The Hargreaves–Samani (HS) model is another method to estimate ET_0 , which was proposed by Hargreaves and Samani [17]. The HS model had been employed in various occasions to estimate ET_0 . However, the constants and coefficients involved in the HS model can be site-specific. Hence, efforts were also done to calibrate the coefficients in HS model to suit local needs and this can be laborious. Luo et al. [18] validated the utilization of calibrated the HS model in Guilin, Kaifeng, Ganyu, and Yinchuan to predict ET_0 by using forecasted temperature. The results showed that, although the prediction accuracy was sufficiently high (87.54% to 96.90%), when the ET_0 is relatively high or relatively low, the Hargreaves model would fail as it did not consider the effects of wind speed and relative humidity. In Veneto, Italy (sub-humid climate), a study was done to compare the performance of calibrated and uncalibrated the HS model [19]. It was found that the standard HS model would overestimate ET_0 value, which led to the tendency of excess water requirements. Calibrating the empirical parameters of the HS model successfully reduced the overestimation from 18.9% to 2.6% thus justified the importance of calibration.

The Priestly–Taylor (PT) Model was proposed by Priestley and Taylor [20]. Similar to the HS model, the PT model instead ignores the sensitivity of ET_0 towards vapor pressure and air movement. On top of that, the PT model even further simplified the PM model where temperature parameters are also removed. Due to the nature of the model that neglects the effect of temperature, the PT model tends to underestimate ET_0 when compared to the PM model [21]. This suggestion was further supported

by the poor performance of PT model at dry climate regions. Moreover, the missing aerodynamic component in the PT model also limited its spatial applicability.

Apart from the three models discussed previously, there are some other less popular models to estimate ET_0 . These include the radiation-based Turc model [22] and temperature-based Thornthwaite model [23]. They are seldom used as a standard for comparison unlike the previous three models. As shown in the literature, conventional models of ET_0 estimation generally have two main shortcomings: highly data intensive and strongly dependent on geographical location or not spatially robust [24,25]. Researchers attempted to solve these problems by modifying or calibrating available models to suit their needs. In addition, researchers also tried to produce a more powerful prediction tool that can handle highly complex and nonlinear processes like ET. This review provides an outlook on the emerging prediction tool, which is the artificial intelligence model.

The “black-box” nature of artificial intelligence model makes it very useful to map the relationship between inputs and outputs even in the absence of relevant scientific knowledge. This leads to its application in ET estimation in order to replace the data intensive and relatively less adaptive empirical models. Increasing popularity of artificial intelligence application in ET estimation can be seen in Figure 1, which is concluded from Scopus. From 2011 to 2015, the number of publications related to ET prediction using artificial intelligence model appears to be quite stagnant. However, from 2016 onwards (except for 2017), there is a steady increase in the number of publications especially in 2018 and 2019. This shows the wide acceptance of such technique by global researchers that affirms the usefulness of artificial intelligence modeling. Having that said, there are certain flaws that exist in the basic artificial intelligence models. Fortunately, this could be overcome by using hybridization techniques to combine different models together in order to produce predictions with better accuracy as well as consistency.

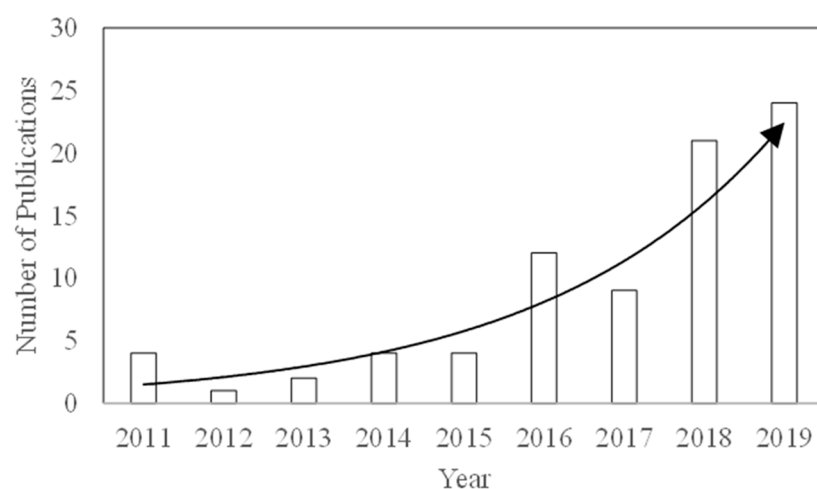


Figure 1. Number of publications related to evapotranspiration estimation using artificial intelligence from 2011 to 2019.

Nonetheless, a review article that focuses on the discussion of the application of artificial intelligence modeling is found to be absent in current literature. This situation has attracted the concern and focus of the authors of this review paper. It is of paramount importance to produce a compilation of related articles together with critical review as a reference and guidelines for future researchers in this field. This shall be regarded as a noble contribution and a starting point to facilitate the progress of relevant scientists and researchers, either as a starting or in their continuous efforts. To the agricultural water and water resources practitioners, the outcome of their research can be used for the decision-making process, which includes the design of irrigation schedules, water resources allocation, and management [26]. Careful, precise, and appropriate decisions are ultimately important for the sustainability of anthropic activities, especially in the context of agricultural economy. This can

also generate more data for the use of research in years to follow. On top of that, published research articles are archived in scientific journals for future references. Figure 2 shows the virtuous cycle between review papers, new research, and decision-making processes.

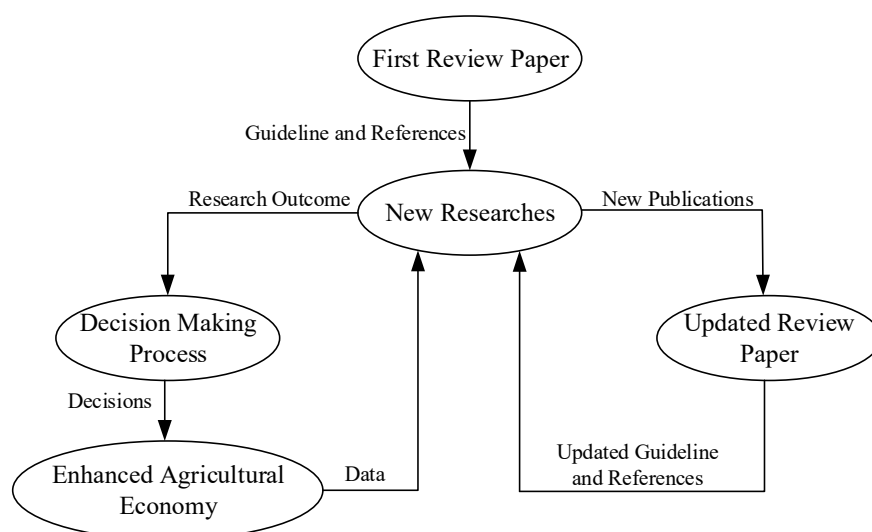


Figure 2. Virtuous cycle between review papers, new research and the decision-making process.

Therefore, this paper aims to play the role of a comprehensive guidance for future research, from several aspects. In Section 2, different types of data and their sources are discussed thoroughly. The advantages and disadvantages of the data types are also summarized in this paper. Section 3 focuses on the explanation of basic artificial intelligence models in terms of their characteristics and application in ET estimation. Hybridization techniques for artificial intelligence models are reviewed in Section 4. Details of each and every technique are explained with respect to their principle and suitability in different situations. Finally, the future prospect of the use of artificial intelligence in ET estimation is presented in Section 5 as a prelude to the concluding remarks of this review paper.

2. Data Types

In order to proceed to the training of artificial intelligence models, the selection and acquisition of data are inseparable processes. Suitable data sets or parameters vary according to different regions as well as climate patterns. In order to aid future researchers during the process of data acquisition, significant parameters (though not exhaustive) to estimate ET for different climate patterns are summarized in Table 1 [27–37]. From Table 1, it can be seen that temperature and radiation data are indispensable for ET estimation. This is strongly in agreement with the background theory whereby temperature (indicating heat energy) and radiation are the two main driving forces of the energy consuming ET process [7].

Table 1. Significant parameters for different climate patterns to estimate evapotranspiration

Climate Pattern	Significant Parameters
Arid	Temperature, Radiation
Semi-Arid	Temperature, Radiation
Humid	Temperature, Radiation, Evaporation
Sub-Humid	Temperature, Radiation, Evaporation
Warm-Humid	Temperature, Radiation
Humid Subtropical	Radiation
Subtropical Monsoon	Temperature
Mediterranean	Radiation

In general, the raw data can be obtained from two main sources, namely the ground observation data and the remote sensing data [38]. In essence, these two types of datasets act as complimentary data to one another. To obtain ground observation data, meteorological and weather stations are set up for continuous collections. FLUXNET, which consists of micrometeorological tower sites, was set up as an initiative to cope with the large demand of ground observation data. The micrometeorological tower sites are mostly concentrated in North America, Europe and Asia, whilst some tower sites are also available in South America, Oceania, and Africa regions. Various types of data could be obtained from the FLUXNET database, including different locations, durations, and time scales. It was claimed that FLUXNET would continue its effort to expand in order to increase the geographical coverage [39]. The main advantage of using ground observation data is that it provides direct measurement, which does not need further imputation or processing in order to retrieve the intrinsic information. Explicitly, the data acquired from weather stations or flux towers are ready to be used without any pre-processing steps. However, measurements of weather station only represent the conditions of the tower's location and its close proximity. In other words, deduction of the weather conditions for a larger region can be a challenging task [40].

Remote sensing emerges as a solution to cover the problem stated. This can be realized by either subtracting the sensible heat flux from net radiation (residual method) or computing the surface resistance with a vegetation index–surface temperature scatterplot [41]. Successful estimation of ET from satellite images opened the door for its forecasting using artificial intelligence models. The major sources of satellite images could be obtained from Landsat, Moderate Resolution Imaging Spectroradiometer (MODIS) and the Global Land Surface Satellite (GLASS). Remote sensing data allow for more information to be captured by satellite images. In fact, the remote sensing method can be used to derive vegetation information as well as different types of radiation, which are useful for ET estimation. Nonetheless, estimation using remote sensing data has to be well calibrated in order to reflect accurate readings. This can be done through the integration of land data assimilation system (LDAS). LDAS forces land surface models with the observed fields and removes biases in forcing based on atmospheric models. In this way, unrealistic model states can be corrected. By merging measurements (from satellite and ground observations) with model estimations, imperfections in observations and errors in model prediction can be minimized [42]. Table 2 summarizes the types of data for ET estimation and their sources, along with the obtainable parameters, advantages, and disadvantages.

Careful selection of data is important for an excellent model training process. Identification and determination of data sources, parameters, and data points are keys to developing an outstanding artificial intelligence model. Hence, prior to model selection, one should always first consider the data available. In the next section, this review will lead the readers to discover some of the more common artificial intelligence models used in ET estimation.

Table 2. Types of data for evapotranspiration estimation.

Data Types	Sources	Available Parameters	Advantages	Disadvantages
Ground Observation	Meteorological Stations, FLUXNET	Temperature, Wind Speed, Radiation, Humidity Sunshine Hours, Vapour Pressure Deficit, Evaporation	Available in different time steps (hourly, daily, monthly)	Only provide point measurement
			Provide direct measurement data	Low spatial coverage
				Less variations of parameters
				Missing data
Remote Sensing	Landsat, MODIS, GLASS	Temperature, Radiation, Vegetation Index, Leaf Area Index, Albedo	More variations of parameters can be derived from satellite images	Require calibration of satellite images for data retrieval
			Able to provide data at different spatial and temporal resolutions	Quality of data affected by weather conditions (cloud coverage) and image resolution
			Higher spatial coverage	
			Real time monitoring	

3. Artificial Intelligence Models

3.1. Artificial Neural Network (ANN)

Artificial neural network (ANN), as suggested by the name, is a variation of the machine learning model that resembles the neural network of human brain. In the latter, neurons are connected to each other via synapses. In ANN, synapses are replaced by weights and biases connections. This helps to map the relationship between inputs and outputs [43]. A Multilayer perceptron (MLP) is one of the earliest types of ANN, and it was introduced by Rosenblatt [44]. However, it was not until the year 1989 that MLP was proven to be able to approximate functions after training [45]. In 1992, in tandem with the advancement of computer development, the MLP showed better performance than traditional statistical method for the first time [46].

The application of the MLP in estimation of ET_0 was initiated by Kumar et al. [47]. In the study, the authors collected the six essential parameters for estimating ET_0 using the PM model in Davis, California. This set of data was dated from 1 January 1990 to 30 June 2000. At the same time, a second set of data dated from 1 January 1960 to 31 December 1963 was obtained together with their corresponding lysimeter readings. The authors intended to compare the performance of the MLP with different architecture trained with different target data. The outcome of the study showed that, when all the six parameters were fed as input of the MLP model, a single layer of seven hidden neurons with 5000 learning cycles was ample to represent the nonlinear process of ET. The training of the model using lysimeter measurement as target produced slightly more accurate estimations than using the PM model. This not only justified the ability of the MLP to map inputs to output in the absence of clear relationship, but also at the same time laid down a foundation that estimation of the PM model is sufficiently good to be used as a target for the model training.

The success achieved by Kumar et al. [47] attracted the attention of researchers to further study the capability of the MLP in estimating ET_0 . Attempts to reduce the number of required parameters were continuously done. Rahimikhoo [48] trained the MLP with only temperature and radiation data in eight different stations in southern coast of Caspian Sea located in northern Iran (humid subtropical climate). By using only maximum temperature, minimum temperature, and global radiation, the trained models were compared with the HS model and the calibrated HS model. The authors had wished to make a comparison between the MLP and empirical models with limited climatic data (both were temperature and radiation based). The study proved that, even in the case where climatic dataset was incomplete, the prediction of the MLP model was still promising, as the lack of flexibility in the empirical model had the tendency of either underestimating or overestimating. Antonopoulos and Antonopoulos [49] conducted their study in a mountainous area in West Macedonia of Greece. The authors continuously removed variables one by one from the MLP model in order to investigate the ability of the MLP in estimating ET_0 with limited climatic parameters. The performance of the MLP contrasted with that of the PT model, the Makkink model, and the HS model. The study showed that, even in the case of two parameters (temperature and radiation), the MLP still outperformed Makkink model and HS model while having comparable performance with the PT model. Some other studies reported by other literature also showed that the MLP could give better estimation than equivalent conventional empirical models in the case of limited climatic parameters, in the four main classes of climate regions such as semi-arid [50,51], arid [31,52,53], humid, and semi humid regions [54].

The introduction of the MLP had also encouraged the establishment of other forms of the ANN model. Some of the examples are the radial basis function network (RBF) [55], generalized regression neural network (GRNN) [56], back-propagation neural network (BPNN) [45] and extreme learning machine (ELM) [57]. These algorithms achieved promising performances in ET_0 estimation. The characteristics of each ANN model are provided in Table 3.

Table 3. Characteristics of artificial neural network models.

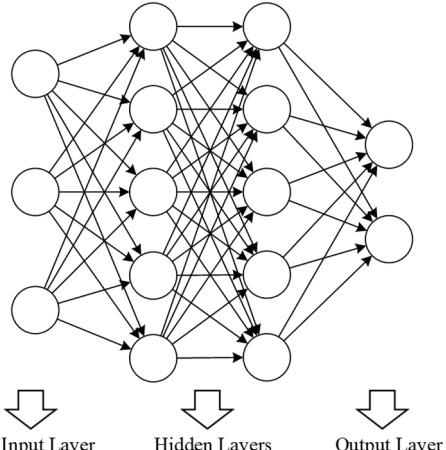
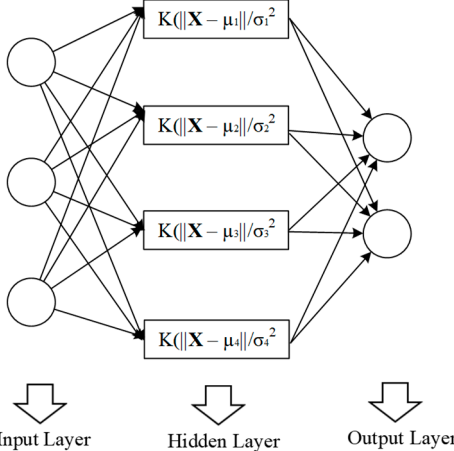
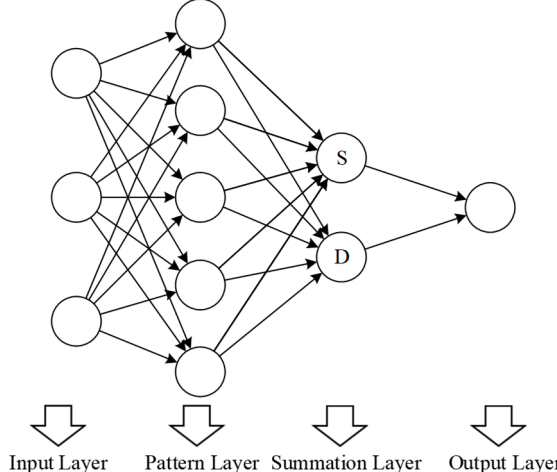
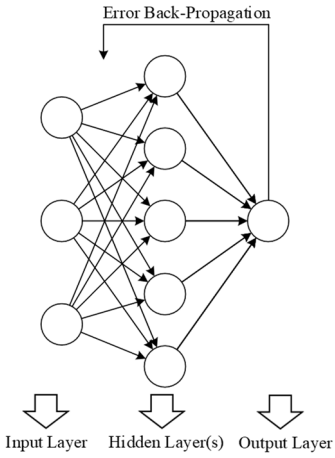
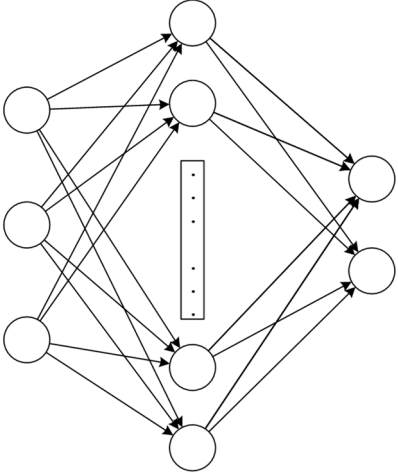
Artificial Neural Network Models	Characteristics
<p>Multilayer Perceptron</p>  <p>Input Layer Hidden Layers Output Layer</p>	<ul style="list-style-type: none"> - Consist of one input layer, one or more hidden layers and one output layer - Signals are passed from input layer to output layer in the forward direction - Normally use sigmoid activation function to map input to output
<p>Radial Basis Function</p>  <p>Input Layer Hidden Layer Output Layer</p>	<ul style="list-style-type: none"> - Consist of one input layer, one hidden layer and one output layer - Gaussian activation function is computed for every nodes in the hidden layer
<p>Generalised Regression Neural Network</p>  <p>Input Layer Pattern Layer Summation Layer Output Layer</p>	<ul style="list-style-type: none"> - A probabilistic based model - Consist of one input layer, one pattern layer, one summation layer and one output layer - Pattern layer is used to cluster the data and train the model - Results of the summation layer nodes are normalised in the output layer

Table 3. Cont.

Artificial Neural Network Models	Characteristics
<p>Back-Propagation Neural Network</p>  <p>Input Layer Hidden Layer(s) Output Layer</p>	<ul style="list-style-type: none"> - Consist of one input layer, one or more hidden layers and one output layer - Include a back-propagation algorithm to feedback the output error in order to optimise the model performance by adjusting weights and biases
<p>Extreme Learning Machine</p>  <p>Input Layer Single Hidden Layer Output Layer</p>	<ul style="list-style-type: none"> - Consist of only one input layer, one hidden layer and one output layer - Number of nodes in hidden layer are randomly generated - Only the number of nodes in the hidden layer have to be tuned to optimize the performance of model

The RBF was first used to convert pan evaporation data into ET_0 [58]. This study proved that the major obstacle of empirical estimation, which is the data dependency can be resolved. The RBF network used in the study required only pan evaporation and radiation data; however, it was able to achieve higher accuracy than both the Christen model and the PM model.

Ladlani et al. [36] did a comparative study on the performance of the RBF and the GRNN to predict ET_0 in Algiers of Algeria. Concurrently, the two ANN models were contrasted with empirical models (PT and HS model). In comparison, the GRNN model had the best performance in terms of low error and high correlation. The GRNN, which was evolved from the RBF, was proved, for the first time, to have more superior ability than the RBF. This could be due to the inclusion of a summation layer in the GRNN, which could enhance the estimation by the RBF. Performance of the GRNN in computing ET_0 is constantly compared with other machine learning models [59,60]. From the results in the literature, the GRNN did not possess any prominent advantageous over ELM, but was deemed to be a good alternative for conventional models.

Traore et al. [61] took a different approach to apply machine learning model in ET_0 estimation where the BPNN model was used. In their study, meteorological data were collected from the

Sudano-Sahelian zone. The HS model, which is a temperature based empirical model, was compared to BPNN model that was trained with only temperature data. The results showed that the artificial intelligence model outperformed the conventional HS model. On top of that, the authors also revealed that inclusion of wind speed data could effectively enhance the accuracy as compared to radiation and relative humidity. Other research works related to the BPNN included comparison of BPNN with gene programming in arid region [62], comparison with tree models [63], and reduction of input parameters [48].

The latest development of artificial intelligence models resulted in the introduction of the ELM as an option of ANN. In 2015, this variation of ANN was first used to estimate ET_0 in Iraq where the authors claimed that this region represented general atmospheric and geographical conditions [30]. Similar to most of the available literature, the authors trained their ELM model by using the PM model estimation as target. Several different combinations of input parameters that consisted of temperature, wind speed, relative humidity, and radiation were tested to study the most favorable combination. Although the ELM and the BPNN showed comparable results, the authors opined that the ELM was preferable due to its efficient computation and great generalization ability. The fast iteration of ELM is due to the fact that only the number of hidden layer nodes have to be tuned and this in turn reduces the risk of overfitting. Their work was followed up by Gocic et al. [64], where they trained the ELM using empirical models with lesser input parameters. In their study, it was found that the ELM trained with the HS model was more superior to those that were trained with the PT model and the Turc model. With that being said, the difference between individual the ELM models were marginal. In comparison with the PM model estimated ET_0 , the ELM predictions had good correlation which justified that the ELM were feasible for such purpose. Subsequently, another study was also carried out to reduce the required input for ELM training [65].

Besides reducing the number of parameters required, researchers were also working towards identifying the optimum training algorithm of the ANN for ET_0 prediction [28]. The study compared six learning algorithms of MLP, namely the Levenberg–Marquardt, Delta–Bar–Delta, Step, Momentum, ConjugateGradient and the QuickProp. For each algorithm, different combinations of input parameters were tested. At the same time, different activation functions such as the hyperbolic tangent, sigmoid, and linear functions. The investigation revealed that irrespective of input parameters, the Levenberg–Marquardt learning algorithm coupled with hyperbolic tangent function was the optimum setting for ET_0 estimation using the MLP. The major distinction of the Levenberg–Marquardt algorithm is that it includes the Gauss–Newton algorithm in its iterative process, which would lead to the search of global minimum, unlike the other algorithms which have higher risk to converge to local minima.

Recently, the investigations related to the ANN prediction of ET were focused on specific case studies. Instead of estimating ET_0 , researchers began to utilize the MLP to predict PET directly. Since PET relies heavily on the types of crops (different K_c), this resulted in the study to be very specific in terms of regions and plantations. For instance, Hashemi and Sepaskhah [66] obtained lysimeter readings from Kooshak Agricultural Research Station in southwest of Iran. They compared the performance of the MLP with the PM model and the radial basis function (RBF) model to estimate PET at a barley plantation. By only feeding sunshine hours, mean humidity, mean temperature, and wind speed as input, both the MLP and the RBF achieved better performance than the PM model. This breakthrough reduced the need of collecting data for K_c computation which could be tedious. Similar work was also carried out on wheat and maize plantations, which demonstrated the advantages of MLP over conventional methods [67]. However, such studies required lysimeter reading and perhaps leaf area index for training purposes. These data are not widely or easily available and thus suffice to say that it would affect the premise of utilizing the ANN in forecasting PET directly for plantations.

An assessment of the various papers that have been reviewed in this subsection reveals the following:

1. The evolution of the ANN from the MLP to the ELM was due to the constant need of improving training methods and algorithms in order to obtain effective predictions with greater accuracy, better generalization and lesser dependency on input parameters.
2. Within each and every variation of the ANN model, one could safely and easily deduce that the trend and focus of study would not deviate much from the following four aspects:
 - a. Minimization of required input parameters,
 - b. Generalization of the ANN for wider spatial application,
 - c. Introduction of new input parameters,
 - d. Enhancement of ANN prediction ability.

It is believed that these four aspects stated above could revolutionise the prediction of ET_0 , with a more general model—without the need for much climatic data.

3. Longer forecasting horizon could provide a good pre-requisite for effective water allocation strategy. The use of the ANN alone could sometimes be insufficient to provide a solution for the above aspects.
4. The black-box operation of ANN could not offer an explanation to the complex ET process.

Therefore, the upcoming subsections will continue to review other artificial intelligence models used to estimate ET_0 in order to provide a complimentary solution to the shortcomings of ANN.

3.2. Support Vector Machine (SVM)

The support vector machine (SVM) is another popular algorithm used in machine learning modeling, especially when it is claimed to be powerful and robust in regression and classification tasks [68]. Cortes and Vapnik [69] laid down the basic and foundation of the current SVM model. Instead of involving large number of neurons and iterations to infer the relationship between inputs and outputs, the SVM plots the datasets into a feature space. The relationship between inputs and outputs is predicted using kernel function, where problem complexity and accuracy can be optimized concurrently.

Since the ET_0 prediction is more likely a regression problem rather than a classification problem, a variation of the SVM, which is the support vector regression (SVR), is normally used. In the SVR, a loss function is used to define the deviation allowance as well as the function to approximate the targeted output [70]. The working principle of SVM is shown in Figure 3.

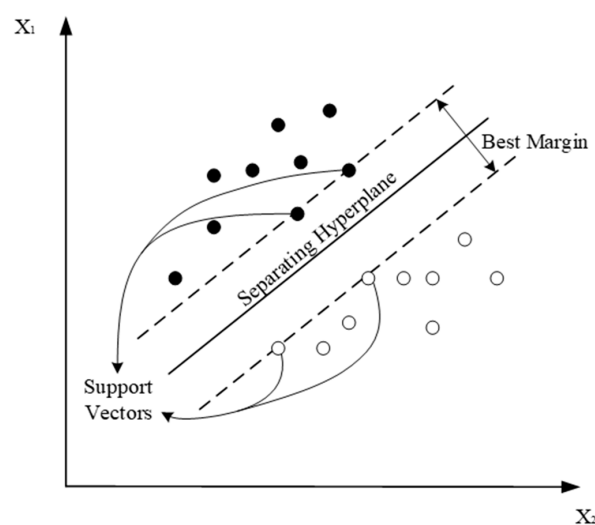


Figure 3. Working principle of support vector machine.

According to Raghavendra and Deka [70], the SVM was widely used in hydrology application, including the estimation of ET_0 . The advantages and strengths of the SVM include high robustness, capability to solve complex problems, less susceptible to overfitting, and it can provide a compact description of the model [71]. The network structure of the SVM is illustrated in Figure 4.

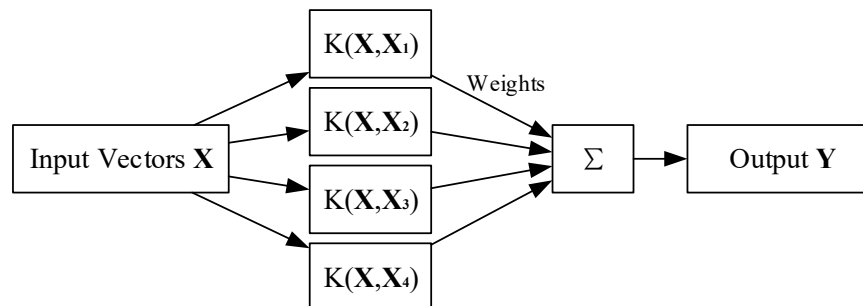


Figure 4. Network structure of support vector machine

The utilization of the SVM in predicting ET_0 with ground observation data started as early as 2010 [72]. The case study was done in California, which represented a Köppen–Geiger climate system. The predictions of the SVM were compared with the CIMIS Penman, HS model, Ritchie model, and the Turc model. The authors discovered that when all climatic parameters were available, the SVM outperformed all other models in all the stations studied. When wind speed and relative humidity were removed during the training of the SVM, the model underestimated ET_0 but still had satisfying outcomes, whereby it only performed slightly worse than conventional the HS model. In fact, the authors also claimed that the SVM had better performance than the ANN, where the former incurred lower error and higher correlation with the standard PM model.

As stated earlier, the performance of the SVM greatly relies on the type of kernel functions chosen to transform the datasets before plotting them into the feature space. Selection of kernel functions can be done using a trial-and-error method. Mehdizadeh et al. [73] did a comprehensive study to compare the performance of the SVM using the RBF and polynomial kernel functions. The study revealed that the RBF kernel function could obtain more accurate results, but did not provide further explanation. A simple deduction that can be made is that the RBF function which represents a Gaussian distribution can be fitted well to the ET problem and datasets of the particular study. This suggestion is supported by the results of Mohammadrezapour et al. [74], where they showed that selection of kernel functions to estimate ET_0 varied from case to case. In other words, there is no universal kernel function that is suited for all problems. Researchers who wished to estimate ET using the SVM should be prepared to include a tuning stage in order to identify suitable kernel functions.

Continuous efforts were done to study the limit of the SVM as well as comparing the SVM to other artificial intelligence models. While the pioneers that applied the SVM in estimating ET_0 reported that the SVM performed better than ANN, some other literature opined otherwise [75,76]. This is due to the nature of the SVM where a global optimum has to be located instead of converging to local optima as in ANN. This makes SVM a very generalizable model, but it would incur higher residuals. More often, the strengths of the SVM would be visualized in the case of limited climatic parameters. In a work done by Fan et al. [77], when only temperature and radiation data were available, the performance of the SVM could be on par with the ELM. In fact, in terms of accuracy and correlation, the SVM achieved better score than most of the hybrid models such as extreme gradient boosted model, random forest, and gradient boosted decision tree. Further discussions on hybrid models is given in a later part of this review.

Similar to the ANN, the SVM had been used to predict PET directly as this could reduce the extra effort of measuring/estimating K_c . This attempt was done by Shrestha and Shukla [78], where they trained their SVM models against lysimeter readings for pepper and watermelon crops. Instead of using conventional climatic parameters, the authors opted some interesting features which included

days after transplant, irrigation frequency, water table depth, soil moisture, rainfall, and rainfall event as well as drainage and runoff frequency. According to the authors, the trained SVM model should be able to predict K_c and ET_0 , thereby making the computation of PET possible. It was observed that the SVM model was robust and well-generalized due to the fact that it could be successfully applied to both vine and erect plantations, and work well in distinct seasons (spring and fall) as well as different irrigation system (drip and sub-irrigation). The SVM not only produced closer estimations to lysimeter readings as compared to the standard estimation procedure [13]; it also beat the ANN and the relevance vector machine (RVM). The performance of the SVM was stable and consistent at each growth stage of the plantations. As a side note, the authors suggested that their SVM models identified that the evaporation and transpiration partition of plantations' PET could be represented by days after transplant, water table depth, rainfall events, and soil surface moisture.

From the reviewed literature, it can be inferred that the SVM has the potential to be reliable for accurate estimation of both ET_0 and PET. However, the literature also revealed that the performance of SVM could be strongly affected by the selection of kernel functions and quality of input data [70]. This could also be justified by the contradicting findings of researchers on the comparison between the SVM and ANN. Computational cost is another concern of SVM application particularly when high dimensionality is involved.

3.3. Fuzzy Models

Introduced by Zadeh [79], fuzzy logic allows the description of data in such a way that a “degree of likeliness” can be given. In other words, by using fuzzy logic, instead of describing in terms of “either A or B”, one can produce a membership degree between 0 and 1 so that the description looks like “partly A and partly B”. Application of fuzzy logic requires an initial set up by experts to determine the type of distribution by selecting a membership function (usually Gaussian function is chosen). In addition, three major ingredients should be fed to fuzzy inference system (FIS), namely a set of fuzzy rule base, database which contains the membership functions and a mechanism (either Sugeno or Mamdani) to apply the fuzzy rules on input and output [80]. The main difference between Sugeno and Mamdani fuzzy logic is the approach to compute the final output. The overall flow of an FIS is illustrated in Figure 5.

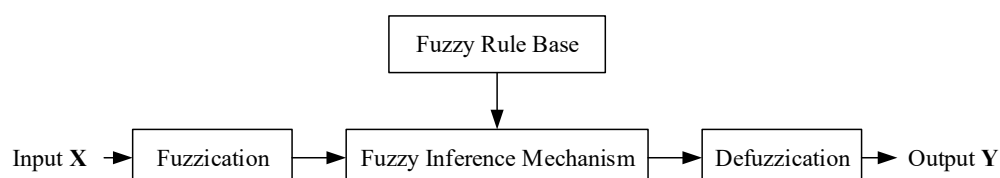


Figure 5. Overall flow of fuzzy inference system

The history of applying fuzzy logic to estimate ET_0 began in 2009. Keskin et al. [81] forecasted the pan evaporation of Lake Eğirdir in Turkey using ground observation climatic data. A comparable study was done in Karso watershed of India [82]. The authors did not only study the feasibility of fuzzy logic predicting pan evaporation, but also the performance of fuzzy logic as compared to the ANN, the least-squared SVR and the adaptive-neuro fuzzy inference system (ANFIS), was also evaluated. The authors remarked that the fuzzy logic model emerged as one of the best models for pan evaporation estimation. This study stressed the importance of fuzzy rules in producing good estimations. Successful application of fuzzy logic shall have good membership functions as foundation. The tuning of membership function not only requires expert knowledge, but is also time-consuming, especially for a complex phenomenon like ET that can be affected by a number of parameters. Hence, the ANN acts as a complimentary to the fuzzy logic to form an ANFIS [83]. Application of the ANFIS for ET_0 estimation was first done by Kisi and Öztürk [84] and there are a number of recent works showing promising results [85–87].

Since the ANFIS is a product of an enhancement based on ANN, hence its performance is frequently compared with ANN in terms of ET_0 estimation. Pour-Ali Baba et al. [85] conducted their experiment in Gwangju and Haenam of South Korea. They realized that the performance of the ANFIS and ANN could vary when the input datasets were different. ANFIS had produced better estimation when solar radiation was fed as input, whereas ANN had better performance when sunshine hours were used. Similarly, the performance of ANFIS and ANN could be affected by geographical location [88]. However, some literature claimed that ANN had slightly better performance than ANFIS, which could be due to ANN's flexibility (not bound by any rules) [89,90].

One interesting study on the ANFIS model is the comparison between two methods of setting up a fuzzy rule, namely the grid partitioning method and subtractive clustering method [91]. The former divides input space in grid-like manner and each region is fuzzy. For subtractive clustering, rules are set up based on the number of clusters found in the input space. It was claimed that subtractive clustering had computational advantage over grid partitioning. Investigation done by Cobaner [91] showed that both approaches had similar performances. However, the ANFIS model using subtractive clustering method could be affected by quality of training data, especially when data are missing [92].

The review of publications in this subsection revealed that, unlike nonlinear learning in ANN and kernel tricks applied in SVM, fuzzy logic provides another way for a machine to learn the rather complex phenomenon of evapotranspiration. The main advantage of the fuzzy logic-based models over the ANN and the SVM is that it actually allows for a more linguistic way of describing the data. In other words, based on the fuzzy rules and membership functions, one can more or less deduce the relationship that maps the inputs to the outputs.

3.4. Tree Based Models

Breiman [93] was the first person to compile decision trees into two main categories, which were the classification tree and regression tree. However, it was Quinlan [94] who provided a better understanding on the operation of tree models. In Quinlan's work, it was stated that the decision would continue to split and grow as long as the data within the nodes of the trees were still considered as impure. In the case of ET, using a tree model for regression analysis is favored over classification. Within this context, Pal and Deswal [95] introduced a widely accepted splitting criterion for the M5 tree model. They claimed that, in order to produce better splits with highest computation efficiency, data within any nodes should be split in such a way that the standard deviation reduction could be maximized. It was observed that the M5 tree model could produce high correlation results to the ET_0 value, although the errors of estimations gradually increased when input climatic parameters were reduced.

There are several research works published that followed up the study of Pal and Deswal [95]. Rahimikhoob et al. [96] attempted to convert pan evaporation data in ET_0 while using other climatic parameters as complementary data. Subsequently, the performance of M5 tree model in predicting ET_0 was compared with the ANN [52]. The study was done in Iran where wind speed and radiation data were found to be absent. The author concluded that, under such circumstances, the M5 tree model could achieve similar performance to the ANN. It was also suggested that the M5 tree model should be favored over the ANN due to its simplicity in terms of computation.

Elsewhere, Kisi and Kilic [97] also studied the difference in prediction performance of the M5 tree model and ANN. In their concluding remarks, the authors revealed that both the M5 tree model and ANN could produce outstanding ET_0 estimation when trained and tested locally. This was however not true when the machine learning models were trained and tested at different stations. The M5 tree model had the worst performance, especially where lesser climatic parameters were available. In fact, the performance of M5 tree model was worse than the empirical models. However, in another study, the results showed disagreement where the M5 tree model was claimed to be having better forecasting accuracy when trained locally as well as using external data [98]. In other words, the M5

tree model could be very dependent on the quality of training data to determine its spatial robustness and generalisability.

According to the papers reviewed regarding to tree-based modeling, it is clear to us that tree-based models exhibit a clear advantage of simple and fast computation. In spite of that, the pitfall of such a model is also obvious. As the tree in the model would have to grow until there are no any other possible splits (the data is deemed to be pure by then), there is a risk of overgrowing the tree. In such circumstance, overfitting could occur which is undesirable for regression analysis. To overcome the problem, a strategy known as pruning is needed to remove unnecessary parts of the tree and replaced them with linear functions. Moreover, the sequence of tree's splitting could end up with different results even though with the same set of training data. To compensate the effect of randomness, trees are sometimes bundled to form a random forest. This will be discussed in detail in later parts of this review.

Basic artificial intelligence models have their own advantages as well as disadvantages. The ANN can be efficient in fitting nonlinear relationship, but it is less explanatory and prone to overfit. The SVM has good generalizability at the expense of costly computation especially for high dimensionality problems. The fuzzy logic provides interpretable rules but that would require initial set up with expert knowledge. The Tree models are computationally efficient, however, would incur high errors. Hence, using basic artificial intelligence models alone is insufficient to accommodate the increasing expectations of their performances. In the following next section, different hybridization techniques of artificial intelligence models are explored as an effective solution to overcome the problems encountered above.

4. Hybrid Models

Hybrid modeling which combines two or more models may somehow improve model performance by merging their individual strengths [99,100]. As demonstrated in the research works mentioned previously, researchers are ambitious for artificial intelligence models to be developed in the future that can work in harsher conditions. For example, in environments with limited climatic parameters, wide region of interest, or longer prediction horizon, among others. Therefore, this section of review will be focusing on uncovering some more commonly used techniques to develop hybrid artificial intelligence models.

4.1. Data Fusion and Ensemble Modelling

4.1.1. Averaging

The idea of ensemble modeling was suggested in 2005, where it was used to forecast weather to overlay predictions of multiple models [101]. The simplest possible ensemble model is by plainly averaging the product of the members of the ensemble. Simple averaging obtains the mean of the models. In such a way, all involved models will be treated as though they have equal performance. In order to correct the absurdity in the assumption of simple averaging, some studies preferred to use weighted averaging. The weight values assigned to the models are ranked based on certain performance measure. For example, Nourani et al. [33] proposed to use the coefficient of determination as a ranking reference. However, these two methods were not comprehensive enough to provide accurate insights for individual models in an ensemble.

Taylor [102] proposed an alternative measure known as the simple Taylor skill. For each individual model, a Taylor skill score will be assigned as the weight value. The Taylor skill score is deemed to be more comprehensive as it takes correlation coefficient and relative standard deviation into consideration. This approach is used by Yao et al. [103], where it was proven that the ensemble model produced from the simple Taylor skill fusion could produce spatial estimation which was comparable to the remote sensing technique. Nonetheless, the authors raised the concern that the simple Taylor skill fusion lacks the ability to describe the ET phenomena physically. This led to the rather low popularity of this method among researchers worldwide.

4.1.2. Bootstrap Aggregating

One of the most common techniques to hybridize artificial intelligence models is the data fusion technique (ensemble modeling). There are various strategies that can lead to the desired output. The first method to be reviewed in the bootstrap aggregating (bagging) method. Generally, bootstrap aggregating involves two main parts: resampling and aggregation. Bootstrap aggregating is especially useful when one has a smaller sample size. During the stage of resampling, the collected samples will be treated as an “apparent population”. Bags of “samples” will be produced from the “apparent population” by using resampling with replacement method. The bags of “samples” will be having an equivalent size with its “apparent population” [104]. Application of bootstrap aggregating in estimating ET_0 is common. Kim et al. [105] applied bootstrap aggregating on the GRNN to study the performance of soft computing in forecasting ET_0 . The study showed that using bootstrapping alone to solely extend the size of training data was insufficient to produce significant improvement to the GRNN models. Instead, the authors suggested training multiple models in order to obtain their aggregated output. It was opined that the latter could effectively reduce the generalization error. This study was the pioneer of utilising bootstrap aggregating for improvement of artificial intelligence models when calculating ET_0 .

The success of Kim et al. [105] attracted the attention of global researchers to conduct similar studies. Besides the GRNN, bootstrap aggregating can be applied on other machine learning models such as tree models. In fact, performing the tree model analysis using bootstrapped samples can lead to the formation of a random forest that was mentioned in passing in the previous section. Feng et al. [59] reported that the random forest model could perform better than the GRNN. In the study of Granata [5], the author compared the results of bagged random forests with individual regression tree models. However, this study reported another finding which claimed that bagging did not significantly improve the performance of single regression tree. Although the author did not provide explanation to his discovery, nonetheless, it is strongly believed that the contradictions between the works of Kim et al. [105], Feng et al. [59], and Granata [5] originated from differences in the datasets. The former two opted to use monthly and annual data respectively, whereas the latter was using daily time step data. Bootstrap aggregating is clearly providing positive effects when the sample size is smaller.

The unique characteristics of bootstrap aggregating is that it does not only perform data pre-processing on the raw datasets. At the same time, it offers an algorithm to aggregate and average out the output of individual models. This is especially useful as an approach to enlarge the limited collected data while offsetting the bias and variance that was aroused from the randomness of model training.

4.1.3. Bayesian Modeling Approaches

Apart from averaging and bootstrap aggregating, another very useful technique to create an ensemble model is via the Bayesian modeling approaches. The Bayesian modeling approaches utilizes the Bayes rules in statistical studies. There are two main strategies when applying the Bayesian modeling approaches in modeling hydrological processes, namely the Bayesian model selection and the Bayesian model averaging [106]. Although both approaches originated from the same fundamentals, their intuitions could still show remarkable differences. In Bayesian model averaging (“team-of-rivals” approach), the main theory underneath is that the model is convinced that there is a truth to be told by models. However, the degree of correctness is strongly dependent on the uncertainties incurred.

Bayesian model averaging works on the basis that it considers truthfulness of the members in an ensemble as their weights. Explicitly, this is realized through the computation of the posterior probability of each model [107]. In this way, the ensemble model would not take excessive risk to exclude models that could be true as well. In the case when data and observations are massive enough to confidently deduce a conclusion, or when there is a particular model that considered virtually able to be true, this model will be promoted to Bayesian model selection (“winner-takes-all” approach).

In other words, Bayesian modeling approaches will keep updating weight of models by imputing their posterior probabilities, until an exceptional model emerges as a “winner” to end the search.

Bayesian modeling approaches were widely used in ET_0 estimation research. Zhu et al. [108] studied the posterior distributions of factors affecting ET_0 for different periods, which varied in terms of leaf area index. In a later study, Zhu et al. [109] produce an ET_0 estimation ensemble model which included the likes of the PM model, the advection-aridity model, the Shuttleworth–Wallace model, and the modified PT model. The outcome of the study proved that, as compared to the simple averaging, the Bayesian model averaging had more positive influences during the process of developing the ensemble model. The authors is of the opinion that the probability density function proposed in the Bayesian theory was well suited to ET phenomena. Despite its good performance, the authors also stressed that output of the Bayesian model averaging was strongly linked to the selection of input parameters.

Chen et al. [110] took a more aggressive approach whereby they used the Bayesian model averaging to combine empirical and artificial intelligence models. The research team suggested two different schemes to create the ensemble model. The authors observed that including all models in the ensemble would result in poorer performance as the Bayesian model averaging assign some weights to poor-performing models. In view of such circumstance, another ensemble was created to include only the models that were performing well. The hypothesis of the authors was verified to be correct.

The usefulness of the Bayesian modeling approaches has resulted in the introduction of various related algorithms such as the Bayesian joint probability [111] and the Bayesian regression [112]. Bayesian regression was thought to be able to provide an insight to the selection of input parameters as well as their relationship to ET_0 . This could provide a solution to policy makers to prioritize collection of data in the near future.

4.1.4. Boosting Algorithm

Boosting is a technique whereby the prediction accuracy is improved by compounding estimations of several weak learners [113]. Unlike the Bayesian model averaging, the boosting algorithm works in a step-wise method, where a learner is added at a time to minimize the loss function. In the boosting algorithm, the first learner will try to search for an optimum loss function value. Subsequently, the following models will be fitted into the ensemble and work on the residuals of their predecessors. Over the years, many versions of boosting algorithms had been established, each with its own novel distinction. Some commonly known boosting methods include the gradient boosting [114], adaptive boosting [115], extreme gradient boosting [116], and the categorical boosting [117].

In recent years, the use of the boosting algorithm in estimating ET_0 has emerged increasingly popular. Fan et al. [77], in particular, had provided the comparison of two types of boosting algorithms on tree models, namely gradient boosting and extreme gradient boosting. The two algorithms differ in the sense that the gradient boosting uses nodes of tree models as weak learners, whereas the extreme gradient boosting uses set of trees as weak learners. The authors found that, generally, the extreme gradient boosting overpowered the gradient boosting. This could be due to the fact that extreme gradient boosting combined the averaged-out results of trees in set, which reduced the variance in the output. On top of that, the design of extreme gradient boosting allows for parallel computation, which could reduce the time taken for analysis.

Ponraj and Vigneswaran [118] proposed the use of the gradient boosting regression to estimate ET_0 at Borrego Springs, California. In the same study, the authors compared the performance of the gradient boosting regression with the conventional multiple linear regression and random forest methods. The results of the study showed that the gradient boosting regression showed higher correlation with the standard PM model. The authors also suggested the use of the gradient boosting machine as an alternative in future investigations.

Recently, Fan et al. [119] used another variation of the boosting algorithm, which was the light gradient boosting algorithm to estimate ET_0 . The operating principle of the light gradient boosting

is that it integrates the essence of gradient boosting and extreme gradient boosting. It performs leaf-wise optimization instead of level-wise optimization. This could effectively reduce the memory and time taken for computation. However, as shown by the results, the light gradient boosting machine required sufficient data in order to be trained well. During the training stage, the performance of light gradient boosting machine was generally weaker than the random forest as well as the M5 tree model. The situation was reversed during the testing phase where light gradient boosting machine performed better when it was well trained.

The core of the boosting algorithm is to assemble several weak learners to form a strong learner. By doing so, the strengths and experience of the weak learners can be utilized well by the hybrid artificial intelligence model. Most important of all, the boosting algorithm can be used as a strategy that can reduce the risk of overfitting. Nevertheless, the development of the boosting algorithm is still at the early stages and more advanced methods shall be anticipated in the near future.

4.1.5. Nonlinear Neural Ensemble

Previously discussed data fusion techniques are developed based on certain statistical logics. There is a kind of data fusion technique that depends on the black-box theory known as the nonlinear neural ensemble. To summarize, outputs of individual artificial intelligence models are fed into a secondary neural network to be trained once more. In other words, an ANN will be used to assemble individual artificial intelligence models. This method had been applied by Nourani et al. [33] through the combined ANN, SVM, ANFIS, and multiple linear regression. When compared with the simple averaging and the weighted averaging, the nonlinear neural ensemble yielded better performance. Similar observations were obtained when they used nonlinear neural ensemble to combine empirical models. This proved that, for a highly nonlinear process like ET, averaging might be insufficient to capture the complexity.

In another study, the individual ANN was added one at a time to produce an ensemble [120]. The addition of ANN was continued until the termination condition was met (tolerable error was achieved). In this way, the architecture and activation function of individual ANN can be constantly modified in order to be considered acceptable into the ensemble. This approach was also used by El-Shafie et al. [121]. The resulting ensemble model would be consisting of only excellent models, which in turn led to the accurate prediction of seasonal ET_0 .

The nonlinear neural ensemble is very useful in the scenario where statistical data fusion methods could not produce improvements as compared to the original artificial intelligence models. It can hybridize the artificial intelligence models by mapping another black-box relationship between inputs and outputs. However, the results would be less interpretable as the intrinsic relationship within the black box could not be observed.

4.1.6. Ensemble Models for Remote Sensing

One of the first attempt to use the machine learning model to estimate ET_0 with remote sensing data was done in the United States, where AmeriFlux sites were available [40]. In the study, land surface temperature, enhanced vegetation index, shortwave radiation, and land cover data are recovered from satellite images that provided 1 km by 1 km coverage and eight-day time step. The ET_0 were estimated using the SVM, ANN and multiple regression. A similar approach was done by Zhang et al. [122] in China where they extended the application of remote sensing data in the BPNN and the ANFIS for estimating ET_0 . Further studies were done to include more artificial intelligence models such as the M5 tree model, bagging, random forest [5], ELM [123], and boosted tree [124]. However, the accuracy of these studies was constrained by the quality of the images for retrieving the estimated meteorological data. It was claimed that the images shall be within microwave band whereas cloud free condition is preferred [125,126].

The major advantages of the remote sensing data over conventional ground observation data include the wide selection of spatiotemporal range as stated earlier. Furthermore, satellite images can

provide a massive variety of parameters to be used for ET_0 estimation. As a result, multiple models can be used to train the artificial intelligence models such as land surface model, energy balance (based on eddy covariance and Bowen ratio) and equations for ET_0 predictions. In spite of its advantages, the shortcoming of remote sensing is that the estimation of ground data could be inaccurate. Moreover, the homogeneity of the satellite images could also affect the process of estimation. Hence, numerous efforts were done to apply data assimilation techniques on ET_0 estimation artificial intelligence models based on remote sensing data. This could be done by merging several satellite images that capture different information and feed them to the artificial intelligence models during the training stage.

The spatial and temporal adaptive reflectance fusion model (STARFM) is regarded as one of the most commonly used data assimilation technique when dealing with remote sensing data. The basic idea of the STARFM is that, if a Landsat-MODIS image pair is available, the algorithm can calculate the systematic error on each pixel for MODIS image in order to retrieve a Landsat-like image [127]. However, this method assumes that both Landsat and MODIS images observe the same amount of reflectance and incur by a constant bias error. Enhanced STARFM (ESTARFM) was introduced later to overcome some problems of STARFM [128]. ESTARFM can handle heterogeneous regions unlike STARFM method. In other words, when the pixel resolutions of the satellite images are not uniform, ESTARFM would be favored over STARFM [129].

The operating principles of both the STARFM and ESTARFM are similar. Available MODIS images are matched with Landsat images of different overpass dates. Optimum base pairs will be used for training the model. Based on the given MODIS image, the two data fusion techniques will retrieve a predicted Landsat image so that the computation of ET_0 is made possible. Cammalleri et al. [130] proposed to apply data fusion technique on remote sensing based ET_0 prediction so that images that carry multiple information can be combined. In their study, Landsat images (30 m spatial resolution, 16-day temporal resolution) and MODIS images (1 km spatial resolution, 1-day temporal resolution) were used to estimate ET_0 using ALEXI and DisALEXI land surface models. By using the STARFM, ET_0 can be estimated from both sets of data. Landsat based ET_0 was compared with Landsat-MODIS based ET_0 . It was found the latter had higher accuracy, especially in the presence of discriminant factors such as rainfall events.

A similar method was applied by Cammalleri et al. [131] to the field scale, where they studied the ET_0 of corn and cotton crops. Semmens et al. [132] extended the application to viticulture system where Landsat, MODIS and multi-sensor data were fused by the STARFM method. Recently, the ALEXI and DisALEXI fused by STARFM was also used by Knipper et al. [126]. Their study was built on the basis of Semmens et al. [132] with the expansion of study to multiple years in order to have an insight of seasonal dynamics of ET_0 . Ma et al. [133] deployed ESTARFM for ET_0 estimation from satellite imaging for the first time. They used three sets of MODIS data and two sets of Landsat data with dissimilar spatial and temporal resolutions. Instead of using the ALEXI and DisALEXI models, the surface energy balance system (SEBS) model was used in the study to calculate ET_0 . The authors claimed that their results produced high resolution ET_0 estimation with good accuracy. It was stated that the estimated ET_0 could produce similar trends with the observed ET_0 with slight underestimation. Nevertheless, a direct comparison between STARFM and ESTARFM has not been studied. Hence, it is believable that the upper hand of ESTARFM is that it allows for inclusion of more satellite images with different resolutions.

Besides the STARFM and ESTARFM, another very popular data assimilation technique used by researchers worldwide is the Kalman based ensemble. The definition of the observable model and state model are essential to use the Kalman algorithm. Alavi et al. [134] demonstrated the usefulness of the Kalman filter based ET_0 estimation. To be exact, the work estimated missing heat flux by treating as a function of time and temperature which acted as observable models. The Kalman filter-based algorithm was compared with the conventional mean diurnal variation, multiple regression, two-week average PT coefficient, and the multiple imputation. It was found that, although the Kalman filter-based algorithm did not show outstanding accuracy among the other methods, the slight difference provided

better estimation of ET_0 , especially during short gap periods with volatile ET_0 fluctuations where sensitivity was a decisive factor.

Ever since then, the ensemble Kalman filter approach was widely used in estimating ET_0 . For instance, Peters-Lidard et al. [135] evaluated several data assimilation systems that employed the Kalman filter approach to predict the latent heat flux. The prediction was done using FLUXNET as well as MODIS data as inputs. It was reported that data assimilation was able to provide more accurate results, indicating wide data structure application of the ensemble Kalman filter. In the Shahe River Basin of China, Yin et al. [136] assimilated a hydrological model (data) with remote sensing-based evapotranspiration. The outcome of the research work showed a promising potential of the ensemble Kalman filter as a predictor when the state model is available. The advantages of using the ensemble Kalman filter are that it permits the realization of multisource data as well as increases the precision of estimation by incorporating suitable models. However, the determination of the observation and state models can be challenging, and this throws the spanner in the works for the popularity of the ensemble Kalman filter.

Utilization of the remote sensing approach in estimating ET_0 removes the constraint of spatial coverage. Satellite images of varying resolutions can be processed to recover valuable information during the prediction. The remote sensing method also enables the provision of real-time data to become possible and allows continuous monitoring of ET of certain regions. The development of data fusion algorithms successfully combines different satellite images and this in turn results in more information to be used in ET_0 prediction. Despite all these, the use of remote sensing is still at its early stage and hence more robust as well as powerful tools can be expected in the near future.

4.2. Data Decomposition

The previous discussions are mainly focused on the exploitation of historical data as ingredients for creating an estimating model. However, temporal trends and variations of ET are of utmost importance as they can be a predictive tool to assist the decision-making of the stakeholders. Therefore, a good artificial intelligence model shall be able to provide such information. Data related to ET could be highly dynamic and contain unnecessary noises. Decomposition of data is needed to filter out the noises in order to retrieve useful information.

According to Partal [137], the wavelet transformation had been successfully applied in many hydrological processes research. In fact, a combination of ANN with wavelet transformation was proved to be feasible in many other studies. Therefore, Partal [137] attempted to perform wavelet transformation of data series of several climate data using different temporal resolutions to obtain useful decompositions. These sub-series were reconstructed and then be fed into the BPNN, multiple linear regression and HS model. The resultant wavelet neural network (WNN) had better performance than the other two models and proved that the wavelet-transformed data were useful to retain only useful information as well as trends. The application of wavelet transformation is not constrained to only BPNN, but is also applied in other ANNs such as RBF [138], ELM [29], GRNN [139,140], and ANFIS [141].

Cobaner [142] converted Class A pan evaporation data into ET_0 by using wavelet decomposition. The study only focused on the effect of wavelet transformation; therefore, instead of using an ANN, the author selected a regression model for analysis. By using the Mallat discrete wavelet transformation, the complex time series was broken down into several sub-time series that exhibited daily, monthly, and annual features of the process. Each time series was weighted based on the strength of its correlation. It was concluded that although the wavelet regression model had slightly lower accuracy than the standard FAO-24 model for pan evaporation conversion; however, the drastic reduction of required parameters was sufficiently proved to be a success of this study.

Apart from the wavelet transformation, there are also other variations of data decomposition. For instance, Adarsh et al. [143] used a multivariate empirical mode decomposition to pre-treat the raw data (temperature, solar radiation, relative humidity and wind speed). In this method, intrinsic mode

functions were generated after the decomposition of the data using varying temporal scale. Using multivariate empirical mode decomposition on the data did not provide significant improvement on the obtained predictions. Future investigations can be conducted to study the effect of such decomposition method when climatic parameters were scarce.

Misaghian et al. [144] provided another form of data decomposition a priori to estimate ET_0 . The ET_0 data were represented in a multi-dimensional or tensor vector space. By using the Tucker decomposition (a variation of singular value decomposition), the three-way relationship of month, year and ET_0 was unfolded. The core tensor can be computed by the prediction machine and reconstructing the predicted original tensor. The authors compared the values of ET_0 computed with the empirical models with those generated by tensor decomposition prediction. The predicted outcome of tensor decomposition model was close to the estimations using the PM model, PT model, HS model, Blaney–Criddle model, and the Jensen–Haise model.

Data decomposition offers another perspective of the ET_0 prediction whereby future ET_0 can be forecasted based on historical trends. In order to obtain a clearer picture of how ET_0 is behaving at different time scales, data decomposition could do the work by filtering noise and generate profiles of ET_0 trends to be analyzed by artificial intelligence models. Data decomposition works as a pre-processing technique that serves to reduce redundant data to the artificial intelligence model in order to produce more meaningful and useful estimations.

Figure 6 outlines the pathways to develop hybrid models using different modeling approaches. In addition, in Table 4, an overview of different hybridization methods is provided. The variations of hybridization models are discussed in details in terms of their background principles and suitable applications.

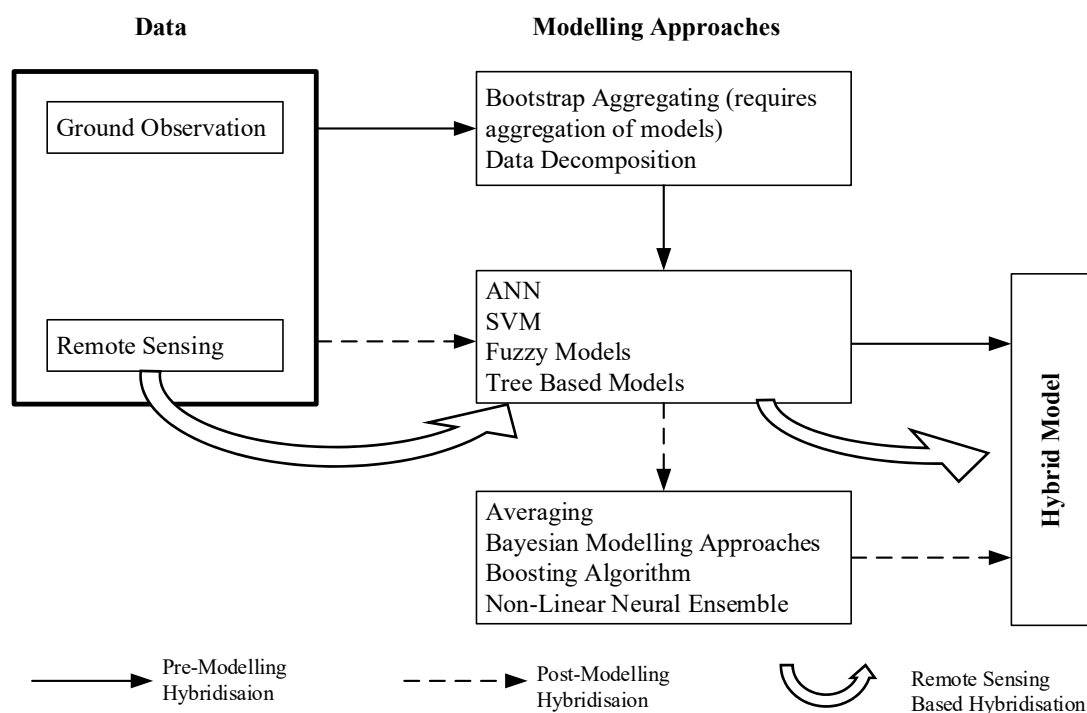


Figure 6. Pathways for hybrid model development. ANN – Artificial Neural Network; SVM – Support Vector Machine.

Table 4. Overview of different hybridization techniques.

Hybridization Techniques	Variations	Principle	Application
Averaging	Simple Averaging	Treat each and every artificial intelligence models as equally good models by obtaining their output mean value	Suitable for less complex problems where outputs of several models can be averaged directly
	Weighted Averaging	Assign weights to each artificial intelligence models based on certain performance measure prior averaging their results	
	Simple Taylor Skill	Weights assigned to each artificial intelligence models are derived by considering more than one performance measures	
Bootstrap Aggregating		Bags of samples are created from original sample (“apparent population”) so that more than one model can be trained and the results are aggregated	Suitable when original sample size is too small or results have high variance and bias
Bayesian Modelling Approaches	Bayesian Model Averaging; Bayesian Model Selection	Weights assigned to each artificial intelligence models are computed based on the posterior probabilities given that the model accurately explain the problems	Can be used to assess the suitability or ability of a model to describe a problem
Boosting Algorithm	Gradient Boosting; Extreme Gradient; Boosting; Light Gradient Boosting	Combine several weak learners (poor performing artificial intelligence models) to form a strong model	Suitable to be used when there are numerous weak learners of different aspects are available
Nonlinear Neural Ensemble		Feed the output of several models into a secondary ANN and rely on black-box operation to obtain the ensemble	Shall be the last resort when no other more intuitive hybridization method is suitable to create an ensemble
Data Decomposition	Wavelet Decomposition; Multivariate Empirical Mode Decomposition; Tensor Decomposition	De-noise the time series data to obtain the trends of different temporal resolutions in order to forecast the future trends	Can be used when time series data are available and when there is a need of forecasting future events
STARFM		Learn using satellite image pairs to compute predictions based on only one image	For remote sensing data application
ESTARFM		Improvement on STARFM in order to handle images with non-uniform pixels or resolutions	For remote sensing data application
Kalman Filter Based Ensemble		Estimate state model by only using observable model as input	Can be used when there are clear definitions of observable and state models

5. Future Prospects

Shifting from conventional empirical models to artificial intelligence models for ET estimation should be regarded as an indubitable trend. This is in line with the introduction of the Fourth Industrial Revolution where artificial intelligence will take over non-value-added activities such as forecasting and estimation. This would assist in the reduction of errors or mistakes when policy makers are making decisions based on highly precise, accurate, and effective predictions. In addition, it is inevitably important for researchers worldwide to seek for solutions and reduce the number of meteorological parameters needed for ET prediction for all its attendant costs and time savings and efficiency reasons. The black-box operating nature of artificial intelligence models is currently the solution to this problem. With that being said, ET data from ground observation or physical measurement would still remain imperative during this transition while a robust artificial intelligence model is being developed concurrently. On the other hand, advancement in satellite technologies allows the use of remote sensing in ET monitoring. In other words, this provides another form of data that would not have been collected from ground weather stations. Application of remote sensing technology in ET estimation reduces the dependency of ET estimation from ground observation data as it offers a new basis to compute ET. Nevertheless, ground observation data are still important in order to calibrate raw satellite images for better prediction in coming years.

In short, the future prospects of this field of study can be summarized as follows:

1. Effective data assimilation.

Data fusion techniques shall be well utilized to accurately map ground observation data to remote sensing data. This can make the satellite images become more informative in terms of accuracy as well as temporal and spatial resolutions, if well calibrated.

2. Creation of new hybrid models.

This can be done by changing the combinations of currently available artificial intelligence models and hybridization techniques. Meanwhile, development of new algorithms or enhancement of present algorithms can be attempted in the future. It is anticipated that the “committee of decision” formed from hybrid models can produce predictions with greater accuracy and shorter computation time.

3. Be cautious of climate changes.

Artificial intelligence models are highly dependent on the training (historical) data. Volatile climate poses a serious challenge where past trends might not be applicable in the future. Studies in the coming years can be focused on retrieving information which take climate changes into consideration. For example, data selection and sampling shall be done with care in order to ensure the homogeneity of the data where the effect of climate change is minimal. In addition, models should be kept as updated as possible. Dynamic modeling can be done to cater to this need while artificial intelligence can support it with fast calculation and real-time data.

4. Relationship discovery from new association rules.

Making use of the “Big Data” allows us to explore various possibilities which associate input parameters to ET. By using the developed artificial intelligence models, one can explore a vast number of variables or parameters and study their association with ET within a short period of time. Parameters that are highly correlated with ET can be further studied to reveal their relationships and scientific interactions.

5. Widening of forecasting horizons.

Related studies were still in the infancy stage where the forecasting windows were too narrow. Increasing forecasting lead time can assist in the design of efficient water resource management plans. This would be important, especially in crop plantations that require a longer time to schedule irrigation their plan.

6. Conclusions

The estimation of ET is of paramount importance, especially when dealing with agricultural activities. This review has outlined the pitfalls of conventional models based on energy balance which included the high dependency on climatic parameters. In addition, empirical models could be specific to certain regions and this in turn requires further calibration before the models could be used. Emergence of artificial intelligence models, which operated on the premise of a black-box principle aims to overcome these problems. The integration of artificial intelligence reignites the possibility of reduction of the now much needed climatic parameters for estimation of ET. Since artificial intelligence models are data-driven, this review has pointed out some sources of data, and also the significance of different parameters in various climate patterns. ANN, SVM, fuzzy models, and tree-based models had been studied extensively in the past and their feasibilities were tried and tested. Nevertheless, studies had revealed that, in the case of limited meteorological parameters or data, performance of these artificial intelligence models would deteriorate.

In view of such circumstance, data fusion techniques had been developed as a solution. Bootstrap aggregating is useful when available data size is too little to train a good model. Bayesian modeling approaches rely on the imputation of posterior probabilities to weigh the correctness of individual models. The boosting algorithm works by combining several weak learners to form a strong learner. Moreover, the nonlinear neural ensemble relies on the black-box operation in order to create an ensemble which produced better results than its constituent individual models. Data decomposition has distinct characteristics whereby it extracts useful information at different resolutions via certain forms of transformation. This could assist in the removal of unwanted noise prior analysis. The purpose of performing data decomposition, especially wavelet transformation is to draw certain trends from historical data in order to predict future behaviour of ET. At the end of this review, a compilation of suggested hybridization techniques for each base artificial intelligence models are provided in Table 5. This could serve as a guideline in terms of parameters selections and ensemble strategies for future research workers who wish to have a fresh start on ET estimation using the hybrid artificial intelligence models.

Remote sensing technology appears to be able to remove the limitation of spatial coverage when estimating ET. It also serves to provide real-time data in order to increase the dynamicity of analysis. Remote sensing-based ET estimation is always integral with land surface model where energy balance and radiation play important roles. Data assimilation can also be performed on remote sensing data where satellite images from different sources could be combined in artificial intelligence models. This enables the combination of different information being carried by different sources of satellite images (including resolution and band range). This can be realized by some of the commonly used techniques such as the STARFM, ESTARFM, and the Kalman filter-based ensemble.

Besides providing the chronological development and guidelines to select methods or algorithms for ET estimation using artificial intelligence models, this review also suggested the future trends of the development of artificial intelligence in ET prediction. In upcoming studies, it is anticipated that data fusion or assimilation would be the major subject alongside with the development of more robust artificial intelligence models. It is in our interest that ground observation data can be merged with remote sensing data. New hybrid models are also anticipated in order to increase the prediction accuracy and speed. In the near future, climate change will be a major environmental issue and researchers shall be cautious about its effect. With matured and well-developed models, we could expect that more parameters well associated with ET can be explored to discover their relationship with ET, all in all for a more profound understand of the processes. Finally, forecasting horizons are to be lengthened for achieving water resources allocation with higher efficiency. It can be a useful tool during key steps of the decision-making process for policy makers, especially in water resources management for a successful economic growth and development in the agricultural sector.

Table 5. Mapping of hybridization techniques for artificial intelligence models.

HybridizationTechnique \ BaseModel	Artificial Neural Network	Support Vector Machine	Tree Based Model	Fuzzy Logic
Averaging	√ ¹	√	√	√
Bootstrap Aggregating	√	SVM does not require much data to map the relationship. Instead, it needs good support vector (data) to infer the relationship between inputs and outputs. Therefore, bootstrap aggregating is seldom used for in SVM hybrid models.	√	Fuzzy model itself contains rules that is interpretable by human language. Researchers tend to apply black-box based hybridization method on it.
Bayesian Model Averaging	√	√	√	
Boosting Algorithm	Boosting algorithm is not necessary for ANN as ANN itself is powerful enough to map most of the relationships (not weak learner).	SVM is also a strong learner and therefore the application of boosting algorithm on it is unnecessary.	√	
Data Decomposition	√	√	√	
Nonlinear Neural Ensemble	√	√	Flexibility of tree based model allows multiple hybridization techniques to be used. It is believed that studies which include tree based model in nonlinear neural ensemble will be available in the future.	√

¹ √ mark indicates combination between base models and hybridization techniques is available.

Author Contributions: M.Y.C. participated in the preparation and write up of the manuscript. Y.F.H. and C.H.K. conceptualised the idea and topic of the paper. K.F.F. provided technical assistance and advises during the completion of the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Universiti Tunku Abdul Rahman (UTAR), Malaysia through Universiti Tunku Abdul Rahman Research Fund under project number IPSR/RMC/UTARRF/2018-C2/K03. The APC was fully funded by UTAR.

Conflicts of Interest: The authors declare no conflicts of interest.

Nomenclature

ANFIS	adaptive neuro-fuzzy inference system
ANN	artificial neural network
BPNN	back-propagation neural network
ELM	extreme learning machine
ESTARFM	enhanced spatial and temporal adaptive reflectance fusion model ET evapotranspiration
ET ₀	reference evapotranspiration
FIS	fuzzy inference system
GLASS	global land surface satellite
GRNN	generalised regression neural network
HS	Hargreaves–Samani
K _c	crop coefficient
LDAS	land data assimilation system
MLP	multilayer layer perceptron
MODIS	Moderate Resolution Imaging Spectroradiometer
PET	potential evapotranspiration
PM	Penman–Monteith
PT	Priestley–Taylor
RBF	radial basis function
RVM	relevance vector machine
SEBS	surface energy balance system
STARFM	spatial and temporal adaptive reflectance fusion model
SVM	support vector machines
SVR	support vector regression
WNN	wavelet neural network

References

1. United Nations. *World Population Prospects: The 2019 Highlights*; ST/ESA/SER.A/423; Department of Economic and Social Affairs/Population Division: New York, NY, USA, 2019.
2. Cascone, S.; Coma, J.; Gagliano, A.; Pérez, G. The evapotranspiration process in green roofs: A review. *Build. Environ.* **2019**, *147*, 337–355. [[CrossRef](#)]
3. Stanhill, G. Evapotranspiration. In *Encyclopedia of Soils in the Environment*; Hillel, D., Ed.; Elsevier: Amsterdam, The Netherlands, 2005; pp. 502–506.
4. Jovic, S.; Nedeljkovic, B.; Golubovic, Z.; Kostic, N. Evolutionary algorithm for reference evapotranspiration analysis. *Comput. Electron. Agric.* **2018**, *150*, 1–4. [[CrossRef](#)]
5. Granata, F. Evapotranspiration evaluation models based on machine learning algorithms—A comparative study. *Agric. Water Manag.* **2019**, *217*, 303–315. [[CrossRef](#)]
6. Holmes, J.W. Measuring evapotranspiration by hydrological methods. *Agric. Water Manag.* **1984**, *8*, 29–40. [[CrossRef](#)]
7. Pokorny, J. Evapotranspiration. In *Encyclopedia of Ecology*, 2nd ed.; Fath, B., Ed.; Elsevier: Amsterdam, The Netherlands, 2019; Volume 2, pp. 292–303.
8. Wang, K.; Dickinson, R.E. A review of global terrestrial evapotranspiration: Observation, modeling, climatology, and climatic variability. *Rev. Geophys.* **2012**, *50*, RG2005. [[CrossRef](#)]

9. Anapalli, S.S.; Ahuja, L.R.; Gowda, P.H.; Ma, L.; Marek, G.; Evett, S.R.; Howell, T.A. Simulation of crop evapotranspiration and crop coefficients with data in weighing lysimeters. *Agric. Water Manag.* **2016**, *177*, 274–283. [\[CrossRef\]](#)
10. Liu, X.; Xu, C.; Zhong, X.; Li, Y.; Yuan, X.; Cao, J. Comparison of 16 models for reference crop evapotranspiration against weighing lysimeter measurement. *Agric. Water Manag.* **2017**, *184*, 145–155. [\[CrossRef\]](#)
11. Pereira, L.S.; Allen, R.G.; Smith, M.; Raes, D. Crop evapotranspiration estimation with FAO56: Past and future. *Agric. Water Manag.* **2015**, *147*, 4–20. [\[CrossRef\]](#)
12. Monteith, J.L. Evaporation and the environment in the state and movement of water in living organisms. In Proceedings of the Society for Experimental Biology, Symposium No. 19, Cambridge, UK, 1 January 1965; Cambridge University Press: Cambridge, UK, 1965; pp. 205–234.
13. Allan, R.G.; Pereira, L.; Raes, D.; Smith, M. *Crop Evapotranspiration—Guidelines for Computing Crop Water Requirements—FAO Irrigation and Drainage Paper 56*; Food and Agriculture Organization of the United Nations: Rome, Italy, 1998; Volume 56.
14. Saggi, M.K.; Jain, S. Reference evapotranspiration estimation and modeling of the Punjab Northern India using deep learning. *Comput. Electron. Agric.* **2019**, *156*, 387–398. [\[CrossRef\]](#)
15. Shiri, J.; Marti, P.; Karimi, S.; Landeras, G. Data splitting strategies for improving data driven models for reference evapotranspiration estimation among similar stations. *Comput. Electron. Agric.* **2019**, *162*, 70–81. [\[CrossRef\]](#)
16. Güçlü, Y.S.; Subyani, A.M.; Şen, Z. Regional fuzzy chain model for evapotranspiration estimation. *J. Hydrol.* **2017**, *544*, 233–241. [\[CrossRef\]](#)
17. Hargreaves, G.H.; Samani, Z.A. Reference Crop Evapotranspiration from Temperature. *Appl. Eng. Agric.* **1985**, *1*, 96–99. [\[CrossRef\]](#)
18. Luo, Y.; Chang, X.; Peng, S.; Khan, S.; Wang, W.; Zheng, Q.; Cai, X. Short-term forecasting of daily reference evapotranspiration using the Hargreaves–Samani model and temperature forecasts. *Agric. Water Manag.* **2014**, *136*, 42–51. [\[CrossRef\]](#)
19. Berti, A.; Tardivo, G.; Chiaudani, A.; Rech, F.; Borin, M. Assessing reference evapotranspiration by the Hargreaves method in north-eastern Italy. *Agric. Water Manag.* **2014**, *140*, 20–25. [\[CrossRef\]](#)
20. Priestley, C.H.B.; Taylor, R.J. On the Assessment of Surface Heat Flux and Evaporation Using Large-Scale Parameters. *Mon. Weather Rev.* **1972**, *100*, 81–92. [\[CrossRef\]](#)
21. Liu, J.G.; Zhao, T.S.; Chen, R.; Wong, C.W. The effect of methanol concentration on the performance of a passive DMFC. *Electrochem. Commun.* **2005**, *7*, 288–294. [\[CrossRef\]](#)
22. Turc, L. Water requirements assessment of irrigation, potential evapotranspiration: Simplified and updated climatic formula. *Ann. Agron.* **1961**, *12*, 13–49.
23. Thornthwaite, C.W. An Approach toward a Rational Classification of Climate. *Geogr. Rev.* **1948**, *38*, 55. [\[CrossRef\]](#)
24. Liu, S.; Xu, Z.; Song, L.; Zhao, Q.; Ge, Y.; Xu, T.; Ma, Y.; Zhu, Z.; Jia, Z.; Zhang, F. Upscaling evapotranspiration measurements from multi-site to the satellite pixel scale over heterogeneous land surfaces. *Agric. For. Meteorol.* **2016**, *230–231*, 97–113. [\[CrossRef\]](#)
25. Valipour, M.; Gholami Sefidkouhi, M.A.; Raeini-Sarjaz, M.; Guzman, S.M. A Hybrid Data-Driven Machine Learning Technique for Evapotranspiration Modeling in Various Climates. *Atmosphere* **2019**, *10*, 311. [\[CrossRef\]](#)
26. Wang, S.; Fu, Z.-Y.; Chen, H.-S.; Nie, Y.-P.; Wang, K.-L. Modeling daily reference ET in the karst area of northwest Guangxi (China) using gene expression programming (GEP) and artificial neural network (ANN). *Theor. Appl. Climatol.* **2015**, *126*, 493–504. [\[CrossRef\]](#)
27. Falamarzi, Y.; Palizdan, N.; Huang, Y.F.; Lee, T.S. Estimating evapotranspiration from temperature and wind speed data using artificial and wavelet neural networks (WNNs). *Agric. Water Manag.* **2014**, *140*, 26–36. [\[CrossRef\]](#)
28. Tabari, H.; Hosseinzadeh Talaei, P. Multilayer perceptron for reference evapotranspiration estimation in a semiarid region. *Neural Comput. Appl.* **2012**, *23*, 341–348. [\[CrossRef\]](#)
29. Kisi, O.; Alizamir, M. Modelling reference evapotranspiration using a new wavelet conjunction heuristic method: Wavelet extreme learning machine vs. wavelet neural networks. *Agric. For. Meteorol.* **2018**, *263*, 41–48. [\[CrossRef\]](#)

30. Abdullah, S.S.; Malek, M.A.; Abdullah, N.S.; Kisi, O.; Yap, K.S. Extreme Learning Machines: A new approach for prediction of reference evapotranspiration. *J. Hydrol.* **2015**, *527*, 184–195. [\[CrossRef\]](#)
31. Huo, Z.; Feng, S.; Kang, S.; Dai, X. Artificial neural network models for reference evapotranspiration in an arid area of northwest China. *J. Arid Environ.* **2012**, *82*, 81–90. [\[CrossRef\]](#)
32. Wen, X.; Si, J.; He, Z.; Wu, J.; Shao, H.; Yu, H. Support-Vector-Machine-Based Models for Modeling Daily Reference Evapotranspiration With Limited Climatic Data in Extreme Arid Regions. *Water Resour. Manag.* **2015**, *29*, 3195–3209. [\[CrossRef\]](#)
33. Nourani, V.; Elkiran, G.; Abdullahi, J. Multi-station artificial intelligence based ensemble modeling of reference evapotranspiration using pan evaporation measurements. *J. Hydrol.* **2019**, *577*, 123958. [\[CrossRef\]](#)
34. Feng, Y.; Gong, D.; Mei, X.; Cui, N. Estimation of maize evapotranspiration using extreme learning machine and generalized regression neural network on the China Loess Plateau. *Hydrol. Res.* **2017**, *48*, 1156–1168. [\[CrossRef\]](#)
35. Huang, G.; Wu, L.; Ma, X.; Zhang, W.; Fan, J.; Yu, X.; Zeng, W.; Zhou, H. Evaluation of CatBoost method for prediction of reference evapotranspiration in humid regions. *J. Hydrol.* **2019**, *574*, 1029–1041. [\[CrossRef\]](#)
36. Ladlani, I.; Houichi, L.; Djemili, L.; Heddami, S.; Belouze, K. Modeling daily reference evapotranspiration (ET₀) in the north of Algeria using generalized regression neural networks (GRNN) and radial basis function neural networks (RBFNN): A comparative study. *Meteorol. Atmos. Phys.* **2012**, *118*, 163–178. [\[CrossRef\]](#)
37. Yamaç, S.S.; Todorovic, M. Estimation of daily potato crop evapotranspiration using three different machine learning algorithms and four scenarios of available meteorological data. *Agric. Water Manag.* **2020**, *228*, 105875. [\[CrossRef\]](#)
38. Maeda, E.E.; Wiberg, D.A.; Pellikka, P.K.E. Estimating reference evapotranspiration using remote sensing and empirical models in a region with limited ground data availability in Kenya. *Appl. Geogr.* **2011**, *31*, 251–258. [\[CrossRef\]](#)
39. National Aeronautics and Space Administration FLUXNET. Available online: https://daac.ornl.gov/cgi-bin/dataset_lister.pl?p=9 (accessed on 23 October 2019).
40. Yang, F.; White, M.A.; Michaelis, A.R.; Ichii, K.; Hashimoto, H.; Votava, P.; Zhu, A.X.; Nemani, R.R. Prediction of Continental-Scale Evapotranspiration by Combining MODIS and AmeriFlux Data Through Support Vector Machine. *IEEE Trans. Geosci. Remote Sens.* **2006**, *44*, 3452–3461. [\[CrossRef\]](#)
41. Nagler, P.; Scott, R.; Westenberg, C.; Cleverly, J.; Glenn, E.; Huete, A. Evapotranspiration on western U.S. rivers estimated using the Enhanced Vegetation Index from MODIS and data from eddy covariance and Bowen ratio flux towers. *Remote Sens. Environ.* **2005**, *97*, 337–351. [\[CrossRef\]](#)
42. Rodell, M.; Houser, P.R.; Jambor, U.; Gottschalk, J.; Mitchell, K.; Meng, C.J.; Arsenault, K.; Cosgrove, B.; Radakovich, J.; Bosilovich, M.; et al. The Global Land Data Assimilation System. *Bull. Am. Meteorol. Soc.* **2004**, *85*, 381–394. [\[CrossRef\]](#)
43. Abiodun, O.I.; Jantan, A.; Omolara, A.E.; Dada, K.V.; Mohamed, N.A.; Arshad, H. State-of-the-art in artificial neural network applications: A survey. *Heliyon* **2018**, *4*, e00938. [\[CrossRef\]](#)
44. Rosenblatt, F. The perceptron: A probabilistic model for information storage and organization in the brain. *Psychol. Rev.* **1958**, *65*, 386–408. [\[CrossRef\]](#)
45. Hornik, K.; Stinchcombe, M.; White, H. Multilayer feedforward networks are universal approximators. *Neural Netw.* **1989**, *2*, 359–366. [\[CrossRef\]](#)
46. Lek, S.; Guégan, J.F. Artificial neural networks as a tool in ecological modelling, an introduction. *Ecol. Model.* **1999**, *120*, 65–73. [\[CrossRef\]](#)
47. Kumar, M.; Raghuwanshi, N.S.; Singh, R.; Wallender, W.W.; Pruitt, W.O. Estimating Evapotranspiration using Artificial Neural Network. *J. Irrig. Drain. Eng.* **2002**, *128*, 224–233. [\[CrossRef\]](#)
48. Rahimikhoob, A. Estimation of evapotranspiration based on only air temperature data using artificial neural networks for a subtropical climate in Iran. *Theor. Appl. Climatol.* **2009**, *101*, 83–91. [\[CrossRef\]](#)
49. Antonopoulos, V.Z.; Antonopoulos, A.V. Daily reference evapotranspiration estimates by artificial neural networks technique and empirical equations using limited input climate variables. *Comput. Electron. Agric.* **2017**, *132*, 86–96. [\[CrossRef\]](#)
50. Reis, M.M.; da Silva, A.J.; Zullo Junior, J.; Tuffi Santos, L.D.; Azevedo, A.M.; Lopes, É.M.G. Empirical and learning machine approaches to estimating reference evapotranspiration based on temperature data. *Comput. Electron. Agric.* **2019**, *165*, 104937. [\[CrossRef\]](#)

51. Citakoglu, H.; Cobaner, M.; Haktanir, T.; Kisi, O. Estimation of Monthly Mean Reference Evapotranspiration in Turkey. *Water Resour. Manag.* **2013**, *28*, 99–113. [[CrossRef](#)]
52. Rahimikhoob, A. Comparison between M5 Model Tree and Neural Networks for Estimating Reference Evapotranspiration in an Arid Environment. *Water Resour. Manag.* **2014**, *28*, 657–669. [[CrossRef](#)]
53. Shiri, J.; Nazemi, A.H.; Sadraddini, A.A.; Landaras, G.; Kisi, O.; Fakheri Fard, A.; Marti, P. Comparison of heuristic and empirical approaches for estimating reference evapotranspiration from limited inputs in Iran. *Comput. Electron. Agric.* **2014**, *108*, 230–241. [[CrossRef](#)]
54. Pandey, P.K.; Nyori, T.; Pandey, V. Estimation of reference evapotranspiration using data driven techniques under limited data conditions. *Model. Earth Syst. Environ.* **2017**, *3*, 1449–1461. [[CrossRef](#)]
55. Broomhead, D.S.; Lowe, D. Multivariable functional interpolation and adaptive networks. *Complex Syst.* **1988**, *2*, 321–355.
56. Specht, D.F. A general regression neural network. *IEEE Trans. Neural Netw.* **1991**, *2*, 568–576. [[CrossRef](#)]
57. Huang, G.-B.; Zhu, Q.-Y.; Siew, C.-K. Extreme learning machine: Theory and applications. *Neurocomputing* **2006**, *70*, 489–501. [[CrossRef](#)]
58. Trajkovic, S. Comparison of radial basis function networks and empirical equations for converting from pan evaporation to reference evapotranspiration. *Hydrol. Process.* **2009**, *23*, 874–880. [[CrossRef](#)]
59. Feng, Y.; Cui, N.; Gong, D.; Zhang, Q.; Zhao, L. Evaluation of random forests and generalized regression neural networks for daily reference evapotranspiration modelling. *Agric. Water Manag.* **2017**, *193*, 163–173. [[CrossRef](#)]
60. Feng, Y.; Peng, Y.; Cui, N.; Gong, D.; Zhang, K. Modeling reference evapotranspiration using extreme learning machine and generalized regression neural network only with temperature data. *Comput. Electron. Agric.* **2017**, *136*, 71–78. [[CrossRef](#)]
61. Traore, S.; Wang, Y.-M.; Kerh, T. Artificial neural network for modeling reference evapotranspiration complex process in Sudano-Sahelian zone. *Agric. Water Manag.* **2010**, *97*, 707–714. [[CrossRef](#)]
62. Yassin, M.A.; Alazba, A.A.; Mattar, M.A. Artificial neural networks versus gene expression programming for estimating reference evapotranspiration in arid climate. *Agric. Water Manag.* **2016**, *163*, 110–124. [[CrossRef](#)]
63. Rahimikhoob, A. Comparison of M5 Model Tree and Artificial Neural Network's Methodologies in Modelling Daily Reference Evapotranspiration from NOAA Satellite Images. *Water Resour. Manag.* **2016**, *30*, 3063–3075. [[CrossRef](#)]
64. Gocic, M.; Petković, D.; Shamshirband, S.; Kamsin, A. Comparative analysis of reference evapotranspiration equations modelling by extreme learning machine. *Comput. Electron. Agric.* **2016**, *127*, 56–63. [[CrossRef](#)]
65. Patil, A.P.; Deka, P.C. An extreme learning machine approach for modeling evapotranspiration using extrinsic inputs. *Comput. Electron. Agric.* **2016**, *121*, 385–392. [[CrossRef](#)]
66. Hashemi, M.; Sepaskhah, A.R. Evaluation of artificial neural network and Penman–Monteith equation for the prediction of barley standard evapotranspiration in a semi-arid region. *Theor. Appl. Climatol.* **2019**. [[CrossRef](#)]
67. Abrishami, N.; Sepaskhah, A.R.; Shahrokhnia, M.H. Estimating wheat and maize daily evapotranspiration using artificial neural network. *Theor. Appl. Climatol.* **2018**, *135*, 945–958. [[CrossRef](#)]
68. Vapnik, V. *The Nature of Statistical Learning Theory*, 2nd ed.; Springer-Verlag: New York, NY, USA, 1995.
69. Cortes, C.; Vapnik, V. Support-vector networks. *Mach. Learn.* **1995**, *20*, 273–297. [[CrossRef](#)]
70. Raghavendra, N.S.; Deka, P.C. Support vector machine applications in the field of hydrology: A review. *Appl. Soft Comput.* **2014**, *19*, 372–386. [[CrossRef](#)]
71. Zendehboudi, A.; Baseer, M.A.; Saidur, R. Application of support vector machine models for forecasting solar and wind energy resources: A review. *J. Clean. Prod.* **2018**, *199*, 272–285. [[CrossRef](#)]
72. Kisi, O.; Cimen, M. Evapotranspiration modelling using support vector machines. *Hydrol. Sci. J.* **2010**, *54*, 918–928. [[CrossRef](#)]
73. Mehdizadeh, S.; Behmanesh, J.; Khalili, K. Using MARS, SVM, GEP and empirical equations for estimation of monthly mean reference evapotranspiration. *Comput. Electron. Agric.* **2017**, *139*, 103–114. [[CrossRef](#)]
74. Mohammadrezapour, O.; Piri, J.; Kisi, O. Comparison of SVM, ANFIS and GEP in modeling monthly potential evapotranspiration in an arid region (Case study: Sistan and Baluchestan Province, Iran). *Water Supply* **2019**, *19*, 392–403. [[CrossRef](#)]

75. Ferreira, L.B.; da Cunha, F.F.; de Oliveira, R.A.; Fernandes Filho, E.I. Estimation of reference evapotranspiration in Brazil with limited meteorological data using ANN and SVM—A new approach. *J. Hydrol.* **2019**, *572*, 556–570. [[CrossRef](#)]
76. Kumar, D.; Adamowski, J.; Suresh, R.; Ozga-Zielinski, B. Estimating Evapotranspiration Using an Extreme Learning Machine Model: Case Study in North Bihar, India. *J. Irrig. Drain. Eng.* **2016**, *142*, 04016032. [[CrossRef](#)]
77. Fan, J.; Yue, W.; Wu, L.; Zhang, F.; Cai, H.; Wang, X.; Lu, X.; Xiang, Y. Evaluation of SVM, ELM and four tree-based ensemble models for predicting daily reference evapotranspiration using limited meteorological data in different climates of China. *Agric. For. Meteorol.* **2018**, *263*, 225–241. [[CrossRef](#)]
78. Shrestha, N.K.; Shukla, S. Support vector machine based modeling of evapotranspiration using hydro-climatic variables in a sub-tropical environment. *Agric. For. Meteorol.* **2015**, *200*, 172–184. [[CrossRef](#)]
79. Zadeh, L.A. Fuzzy sets. *Inf. Control* **1965**, *8*, 38–53. [[CrossRef](#)]
80. Kisi, O. Applicability of Mamdani and Sugeno fuzzy genetic approaches for modeling reference evapotranspiration. *J. Hydrol.* **2013**, *504*, 160–170. [[CrossRef](#)]
81. Keskin, M.E.; Terzi, Ö.; Taylan, D. Fuzzy logic model approaches to daily pan evaporation estimation in western Turkey / Estimation de l'évaporation journalière du bac dans l'Ouest de la Turquie par des modèles à base de logique floue. *Hydrol. Sci. J.* **2004**, *49*, 1001–1010. [[CrossRef](#)]
82. Goyal, M.K.; Bharti, B.; Quilty, J.; Adamowski, J.; Pandey, A. Modeling of daily pan evaporation in sub tropical climates using ANN, LS-SVR, Fuzzy Logic, and ANFIS. *Expert Syst. Appl.* **2014**, *41*, 5267–5276. [[CrossRef](#)]
83. Jang, J.R. Self-learning fuzzy controllers based on temporal backpropagation. *IEEE Trans. Neural Netw.* **1992**, *3*, 714–723. [[CrossRef](#)]
84. Kisi, Ö.; Öztürk, Ö. Adaptive Neurofuzzy Computing Technique for Evapotranspiration Estimation. *J. Irrig. Drain. Eng.* **2007**, *133*, 368–379. [[CrossRef](#)]
85. Pour-Ali Baba, A.; Shiri, J.; Kisi, O.; Fard, A.F.; Kim, S.; Amini, R. Estimating daily reference evapotranspiration using available and estimated climatic data by adaptive neuro-fuzzy inference system (ANFIS) and artificial neural network (ANN). *Hydrol. Res.* **2013**, *44*, 131–146. [[CrossRef](#)]
86. Petković, D.; Gocic, M.; Trajkovic, S.; Shamshirband, S.; Motamedi, S.; Hashim, R.; Bonakdari, H. Determination of the most influential weather parameters on reference evapotranspiration by adaptive neuro-fuzzy methodology. *Comput. Electron. Agric.* **2015**, *114*, 277–284. [[CrossRef](#)]
87. Keshtegar, B.; Kisi, O.; Ghohani Arab, H.; Zounemat-Kermani, M. Subset Modeling Basis ANFIS for Prediction of the Reference Evapotranspiration. *Water Resour. Manag.* **2017**, *32*, 1101–1116. [[CrossRef](#)]
88. Kisi, O.; Sanikhani, H.; Zounemat-Kermani, M.; Niazi, F. Long-term monthly evapotranspiration modeling by several data-driven methods without climatic data. *Comput. Electron. Agric.* **2015**, *115*, 66–77. [[CrossRef](#)]
89. Gavili, S.; Sanikhani, H.; Kisi, O.; Mahmoudi, M.H. Evaluation of several soft computing methods in monthly evapotranspiration modelling. *Meteorol. Appl.* **2018**, *25*, 128–138. [[CrossRef](#)]
90. Seifi, A.; Riahi, H. Estimating daily reference evapotranspiration using hybrid gamma test-least square support vector machine, gamma test-ANN, and gamma test-ANFIS models in an arid area of Iran. *J. Water Clim. Chang.* **2018**. [[CrossRef](#)]
91. Cobaner, M. Evapotranspiration estimation by two different neuro-fuzzy inference systems. *J. Hydrol.* **2011**, *398*, 292–302. [[CrossRef](#)]
92. Kisi, O.; Zounemat-Kermani, M. Comparison of Two Different Adaptive Neuro-Fuzzy Inference Systems in Modelling Daily Reference Evapotranspiration. *Water Resour. Manag.* **2014**, *28*, 2655–2675. [[CrossRef](#)]
93. Breiman, L. *Classification and Regression Trees*; Routledge: Abington on the Thames, UK, 1984.
94. Quinlan, J.R. Learning with continuous classes. In *Proceedings of the Australian Joint Conference on Artificial Intelligence*, Singapore; World Scientific Press: Singapore, 1992; pp. 343–348.
95. Pal, M.; Deswal, S. M5 model tree based modelling of reference evapotranspiration. *Hydrol. Process.* **2009**, *23*, 1437–1443. [[CrossRef](#)]
96. Rahimikhoob, A.; Asadi, M.; Mashal, M. A Comparison Between Conventional and M5 Model Tree Methods for Converting Pan Evaporation to Reference Evapotranspiration for Semi-Arid Region. *Water Resour. Manag.* **2013**, *27*, 4815–4826. [[CrossRef](#)]
97. Kisi, O.; Kilic, Y. An investigation on generalization ability of artificial neural networks and M5 model tree in modeling reference evapotranspiration. *Theor. Appl. Climatol.* **2015**, *126*, 413–425. [[CrossRef](#)]

98. Kisi, O. Modeling reference evapotranspiration using three different heuristic regression approaches. *Agric. Water Manag.* **2016**, *169*, 162–172. [\[CrossRef\]](#)
99. Fung, K.F.; Huang, Y.F.; Koo, C.H.; Soh, Y.W. Drought forecasting: A review of modelling approaches 2007–2017. *J. Water Clim. Chang.* **2019**. [\[CrossRef\]](#)
100. Galar, M.; Fernandez, A.; Barrenechea, E.; Bustince, H.; Herrera, F. A Review on Ensembles for the Class Imbalance Problem: Bagging-, Boosting-, and Hybrid-Based Approaches. *IEEE Trans. Syst. Man Cybern. Part C (Appl. Rev.)* **2012**, *42*, 463–484. [\[CrossRef\]](#)
101. Palmer, T.N.; Doblas-Reyes, F.J.; Hagedorn, R.; Weisheimer, A. Probabilistic prediction of climate using multi-model ensembles: From basics to applications. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **2005**, *360*, 1991–1998. [\[CrossRef\]](#) [\[PubMed\]](#)
102. Taylor, K.E. Summarizing multiple aspects of model performance in a single diagram. *J. Geophys. Res. Atmos.* **2001**, *106*, 7183–7192. [\[CrossRef\]](#)
103. Yao, Y.; Liang, S.; Li, X.; Zhang, Y.; Chen, J.; Jia, K.; Zhang, X.; Fisher, J.B.; Wang, X.; Zhang, L.; et al. Estimation of high-resolution terrestrial evapotranspiration from Landsat data using a simple Taylor skill fusion method. *J. Hydrol.* **2017**, *553*, 508–526. [\[CrossRef\]](#)
104. Breiman, L. Bagging predictors. *Mach. Learn.* **1996**, *24*, 123–140. [\[CrossRef\]](#)
105. Kim, S.; Singh, V.P.; Seo, Y.; Kim, H.S. Modeling Nonlinear Monthly Evapotranspiration Using Soft Computing and Data Reconstruction Techniques. *Water Resour. Manag.* **2013**, *28*, 185–206. [\[CrossRef\]](#)
106. Höge, M.; Guthke, A.; Nowak, W. The hydrologist’s guide to Bayesian model selection, averaging and combination. *J. Hydrol.* **2019**, *572*, 96–107. [\[CrossRef\]](#)
107. Draper, D. Assessment and Propagation of Model Uncertainty. *J. R. Stat. Soc. Ser. B (Methodol.)* **1995**, *57*, 45–70. [\[CrossRef\]](#)
108. Zhu, G.; Su, Y.; Li, X.; Zhang, K.; Li, C. Estimating actual evapotranspiration from an alpine grassland on Qinghai-Tibetan plateau using a two-source model and parameter uncertainty analysis by Bayesian approach. *J. Hydrol.* **2013**, *476*, 42–51. [\[CrossRef\]](#)
109. Zhu, G.; Li, X.; Zhang, K.; Ding, Z.; Han, T.; Ma, J.; Huang, C.; He, J.; Ma, T. Multi-model ensemble prediction of terrestrial evapotranspiration across north China using Bayesian model averaging. *Hydrol. Process.* **2016**, *30*, 2861–2879. [\[CrossRef\]](#)
110. Chen, Y.; Yuan, W.; Xia, J.; Fisher, J.B.; Dong, W.; Zhang, X.; Liang, S.; Ye, A.; Cai, W.; Feng, J. Using Bayesian model averaging to estimate terrestrial evapotranspiration in China. *J. Hydrol.* **2015**, *528*, 537–549. [\[CrossRef\]](#)
111. Zhao, T.; Wang, Q.J.; Schepen, A. A Bayesian modelling approach to forecasting short-term reference crop evapotranspiration from GCM outputs. *Agric. For. Meteorol.* **2019**, *269–270*, 88–101. [\[CrossRef\]](#)
112. Khoshraves, M.; Sefidkouhi, M.A.G.; Valipour, M. Estimation of reference evapotranspiration using multivariate fractional polynomial, Bayesian regression, and robust regression models in three arid environments. *Appl. Water Sci.* **2015**, *7*, 1911–1922. [\[CrossRef\]](#)
113. Hassan, M.A.; Khalil, A.; Kaseb, S.; Kassem, M.A. Exploring the potential of tree-based ensemble methods in solar radiation modeling. *Appl. Energy* **2017**, *203*, 897–916. [\[CrossRef\]](#)
114. Friedman, J.H. Greedy function approximation: A gradient boosting machine. *Ann. Stat.* **2001**, *29*, 1189–1232. [\[CrossRef\]](#)
115. Freund, Y.; Schapire, R.E. A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting. *J. Comput. Syst. Sci.* **1997**, *55*, 119–139. [\[CrossRef\]](#)
116. Chen, T.; Guestrin, C. XGBoost: A scalable tree boosting system. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, 2016; pp. 785–794.
117. Prokhorenkova, L.; Gusev, G.; Vorobev, A.; Dorogush, A.V.; Gulin, A. CatBoost: Unbiased boosting with categorical features. In Proceedings of the 32nd International Conference on Neural Information Processing Systems, Montréal, QC, Canada, 3–8 December 2018; pp. 6639–6649.
118. Ponraj, A.S.; Vigneswaran, T. Daily evapotranspiration prediction using gradient boost regression model for irrigation planning. *J. Supercomput.* **2019**. [\[CrossRef\]](#)
119. Fan, J.; Ma, X.; Wu, L.; Zhang, F.; Yu, X.; Zeng, W. Light Gradient Boosting Machine: An efficient soft computing model for estimating daily reference evapotranspiration with local and external meteorological data. *Agric. Water Manag.* **2019**, *225*, 105758. [\[CrossRef\]](#)

120. El-Shafie, A.; Alsulami, H.M.; Jahanbani, H.; Najah, A. Multi-lead ahead prediction model of reference evapotranspiration utilizing ANN with ensemble procedure. *Stoch. Environ. Res. Risk Assess.* **2012**, *27*, 1423–1440. [\[CrossRef\]](#)
121. El-Shafie, A.; Najah, A.; Alsulami, H.M.; Jahanbani, H. Optimized Neural Network Prediction Model for Potential Evapotranspiration Utilizing Ensemble Procedure. *Water Resour. Manag.* **2014**, *28*, 947–967. [\[CrossRef\]](#)
122. Zhang, Z.; Gong, Y.; Wang, Z. Accessible remote sensing data based reference evapotranspiration estimation modelling. *Agric. Water Manag.* **2018**, *210*, 59–69. [\[CrossRef\]](#)
123. Dou, X.; Yang, Y. Evapotranspiration estimation using four different machine learning approaches in different terrestrial ecosystems. *Comput. Electron. Agric.* **2018**, *148*, 95–106. [\[CrossRef\]](#)
124. Carter, C.; Liang, S. Evaluation of ten machine learning methods for estimating terrestrial evapotranspiration from remote sensing. *Int. J. Appl. Earth Obs. Geoinf.* **2019**, *78*, 86–92. [\[CrossRef\]](#)
125. Knipper, K.; Hogue, T.; Scott, R.; Franz, K. Evapotranspiration estimates derived using multi-platform remote sensing in a semiarid region. *Remote Sens.* **2017**, *9*, 184. [\[CrossRef\]](#)
126. Knipper, K.R.; Kustas, W.P.; Anderson, M.C.; Alfieri, J.G.; Prueger, J.H.; Hain, C.R.; Gao, F.; Yang, Y.; McKee, L.G.; Nieto, H.; et al. Evapotranspiration estimates derived using thermal-based satellite remote sensing and data fusion for irrigation management in California vineyards. *Irrig. Sci.* **2018**, *37*, 431–449. [\[CrossRef\]](#)
127. Feng, G.; Masek, J.; Schwaller, M.; Hall, F. On the blending of the Landsat and MODIS surface reflectance: Predicting daily Landsat surface reflectance. *IEEE Trans. Geosci. Remote Sens.* **2006**, *44*, 2207–2218. [\[CrossRef\]](#)
128. Zhu, X.; Chen, J.; Gao, F.; Chen, X.; Masek, J.G. An enhanced spatial and temporal adaptive reflectance fusion model for complex heterogeneous regions. *Remote Sens. Environ.* **2010**, *114*, 2610–2623. [\[CrossRef\]](#)
129. Hilker, T.; Wulder, M.A.; Coops, N.C.; Seitz, N.; White, J.C.; Gao, F.; Masek, J.G.; Stenhouse, G. Generation of dense time series synthetic Landsat data through data blending with MODIS using a spatial and temporal adaptive reflectance fusion model. *Remote Sens. Environ.* **2009**, *113*, 1988–1999. [\[CrossRef\]](#)
130. Cammalleri, C.; Anderson, M.C.; Gao, F.; Hain, C.R.; Kustas, W.P. A data fusion approach for mapping daily evapotranspiration at field scale. *Water Resour. Res.* **2013**, *49*, 4672–4686. [\[CrossRef\]](#)
131. Cammalleri, C.; Anderson, M.C.; Gao, F.; Hain, C.R.; Kustas, W.P. Mapping daily evapotranspiration at field scales over rainfed and irrigated agricultural areas using remote sensing data fusion. *Agric. For. Meteorol.* **2014**, *186*, 1–11. [\[CrossRef\]](#)
132. Semmens, K.A.; Anderson, M.C.; Kustas, W.P.; Gao, F.; Alfieri, J.G.; McKee, L.; Prueger, J.H.; Hain, C.R.; Cammalleri, C.; Yang, Y.; et al. Monitoring daily evapotranspiration over two California vineyards using Landsat 8 in a multi-sensor data fusion approach. *Remote Sens. Environ.* **2016**, *185*, 155–170. [\[CrossRef\]](#)
133. Ma, Y.; Liu, S.; Song, L.; Xu, Z.; Liu, Y.; Xu, T.; Zhu, Z. Estimation of daily evapotranspiration and irrigation water efficiency at a Landsat-like scale for an arid irrigation area using multi-source remote sensing data. *Remote Sens. Environ.* **2018**, *216*, 715–734. [\[CrossRef\]](#)
134. Alavi, N.; Warland, J.S.; Berg, A.A. Filling gaps in evapotranspiration measurements for water budget studies: Evaluation of a Kalman filtering approach. *Agric. For. Meteorol.* **2006**, *141*, 57–66. [\[CrossRef\]](#)
135. Peters-Lidard, C.D.; Kumar, S.V.; Mocko, D.M.; Tian, Y. Estimating evapotranspiration with land data assimilation systems. *Hydrol. Process.* **2011**, *25*, 3979–3992. [\[CrossRef\]](#)
136. Yin, J.; Zhan, C.; Ye, W. An Experimental Study on Evapotranspiration Data Assimilation Based on the Hydrological Model. *Water Resour. Manag.* **2016**, *30*, 5263–5279. [\[CrossRef\]](#)
137. Partal, T. Modelling evapotranspiration using discrete wavelet transform and neural networks. *Hydrol. Process.* **2009**, *23*, 3545–3555. [\[CrossRef\]](#)
138. Partal, T. Comparison of wavelet based hybrid models for daily evapotranspiration estimation using meteorological data. *KSCE J. Civ. Eng.* **2015**, *20*, 2050–2058. [\[CrossRef\]](#)
139. Adamala, S.; Raghuwanshi, N.S.; Mishra, A.; Singh, R. Generalized wavelet neural networks for evapotranspiration modeling in India. *ISH J. Hydraul. Eng.* **2017**, *25*, 119–131. [\[CrossRef\]](#)
140. Adamala, S. Temperature based generalized wavelet-neural network models to estimate evapotranspiration in India. *Inf. Process. Agric.* **2018**, *5*, 149–155. [\[CrossRef\]](#)
141. Patil, A.P.; Deka, P.C. Performance evaluation of hybrid Wavelet-ANN and Wavelet-ANFIS models for estimating evapotranspiration in arid regions of India. *Neural Comput. Appl.* **2015**, *28*, 275–285. [\[CrossRef\]](#)

142. Cobaner, M. Reference evapotranspiration based on Class A pan evaporation via wavelet regression technique. *Irrig. Sci.* **2011**, *31*, 119–134. [[CrossRef](#)]
143. Adarsh, S.; Sanah, S.; Murshida, K.K.; Nooramol, P. Scale dependent prediction of reference evapotranspiration based on Multi-Variate Empirical mode decomposition. *Ain Shams Eng. J.* **2018**, *9*, 1839–1848. [[CrossRef](#)]
144. Misaghian, N.; Shamsirband, S.; Petković, D.; Gocic, M.; Mohammadi, K. Predicting the reference evapotranspiration based on tensor decomposition. *Theor. Appl. Climatol.* **2016**, *130*, 1099–1109. [[CrossRef](#)]



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).