Desalination and Water Treatment
www.deswater.com

doi: 10.5004/dwt.2022.28960

276 (2022) 28–39
November

Productivity modelling of an inclined stepped solar still for seawater desalination using boosting algorithms based on experimental data

Raniyah Wazirali[a], Mohammed Shadi S. Abujazar[b,c,*], Sohaib K.M. Abujayyab[d], Rami Ahmad[e], Suja Fatihah[f], A.E. Kabeel[g], Sakine Ugurlu Karaağaç[c], Salem S. Abu Amr[h], Motasem Y.D. Alazaiza[i], Mohammed J.K. Bashir[j], Ibrahim Y Sokar[k]

[a]*College of Computing and Informatics, Saudi Electronic University, Riyadh 11673, Saudi Arabia, email: r.wazirali@seu.edu.sa (R. Wazirali)*
[b]*Al-Aqsa Community Intermediate College, Al-Aqsa University, Gaza, Palestine, P.B.4051, email: shadiabujazar@gmail.com/ms.abujazar@alaqsa.edu.ps (M.S.S. Abujazar)*
[c]*Department of Environmental Engineering, Faculty of Engineering, Karabük University, Karabük 78050, Turkey, email: sakineugurlu@karabuk.edu.tr (S.U. Karaağaç)*
[d]*Fire safety engineering, International College for Engineering and Management, Muscat 112, Oman, email: sohaib@icem.edu.om (S.K.M. Abujayyab)*
[e]*College of Computer Information Technology, American University in the Emirates, 503000, United Arab Emirates, email: rami.alshwaiyat@aue.ae (R. Ahmad)*
[f]*Department of Civil Engineering, Faculty of Engineering and Built Environment, Universiti Kebangsaan Malaysia, 43600, Bangi, Selangor, Malaysia, email: fati@ukm.edu.my (S. Fatihah)*
[g]*Mechanical Power Engineering Department, Faculty of Engineering, Tanta University, Tanta, Egypt, email: kabeel6@hotmail.com (A.E. Kabeel)*
[h]*International College for Engineering and Management, 111 St, Seeb, Muscat, Oman, email: salemabuamro@karabuk.edu.tr (S.S. Abu Amr)*
[i]*Department of Civil and Environmental Engineering, College of Engineering, A'Sharqiyah University, 400 Ibra, Oman, email: my.azaiza@gmail.com (M.Y.D. Alazaiza)*
[j]*Department of Environmental Engineering, Faculty of Engineering and Green Technology, Universiti Tunku Abdul Rahman, 31900 Kampar, Perak, Malaysia, email: jkbashir@utar.edu.my (M.J.K. Bashir)*
[k]*Faculty of Computer Science and Information Technology, Gaza University, Gaza, Palestine, email: i.sokar@gu.edu.ps (I.Y. Sokar)*

Received 16 May 2022; Accepted 25 September 2022

ABSTRACT

Solar energy has recently become a viable option for desalinating seawater, primarily in arid regions. However, increasing the productivity of solar still by integrating experimental base and modelling methods is still subject to prediction errors; therefore, the main objective of this research is to postulate and test boosting algorithms for predicting the efficiency and productivity of the system. Five boosting regressors were deployed and evaluated: categorical boosting, adaptive boosting, extreme gradient boosting, gradient boosting machine, and gradient boosting machine (LightGBM). The proposed regressors are implemented based on the system's actual recorded dataset (consisting of 720 observations). The dataset consists of input variables, which are the wind speed (V), cloud cover, humidity, ambient temperature (T), solar radiation (SR), (T_{io}), (T_w), (T_v), and (T_t). Also, the output variable is represented by the productivity of the system. The dataset was separated into training (70%) and testing (30%) sets. In order to decrease regressors

* Corresponding author.

errors, hyperparameter optimization was employed. GradientBoosting approach provided the best prediction, with 95% R^2 accuracy and 39.57 root mean square error (RMSE) error. The LightGBM technique achieved 94% R^2 accuracy and 40.07 RMSE error in the testing dataset. The results reveal that GradientBoosting outperforms the cascaded forward neural network in predicting system productivity (CFNN).

Keywords: Solar desalination; Meteorological data; Boosting algorithms; Modelling; Productivity evaluation

1. Introduction

Apart from air, water is the most important resource that a human being needs for survival. According to projections, by 2025, eight hundred million people will live in regions with absolute water scarcity, and two-thirds of the global total may suffer from severe water stress and lack access to safe drinking water [1,2]. The problem has been exacerbated in recent decades by extreme climate change [3,4]: In many places, surface waters have evaporated, and humans, animals, and agricultural sectors are competing for the rest scarce resources [5]. The spread of waterborne diseases such as diarrhea, typhoid, and cholera is facilitated by the poor quality of dirty water [6].

Water covers roughly three-fourths of the globe [7] but is unevenly distributed. Only 3% of the surface water is fresh; the other 90% is in the ocean. Freshwater is found in glaciers at 69%, underground at 30%, lakes, rivers, and swamps at less than 1% [8–10]. Water scarcity occurs when there are insufficient water resources to meet current and projected demand from all sectors, whether due to a sharp drop in supply, an increase in demand, population increase or changes in consumer behavior, or institutional factors [11–13].

Reportedly, nearly one-fifth of the world's population lives in water shortages, while another one-quarter lives in areas of economic water scarcity. In the last century, water use has increased at a rate of more than double that of population growth. In comparison, the rapid increase in global population and acceleration in global economic activity leads to enhanced consumption of clean energy and finite natural resources, such as water [14]. Desalination techniques have become an extremely popular choice for new water supplies in coastlines regions, with various advanced desalination technologies such as reverse osmosis (RO) [15–17], multi-stage flashing (MSF) [18,19], thin film desalination [20,21], multi-effect evaporation electrodialysis [22,23], humidification dehumidification [24,25], and solar stills [1,26] have been the most widely used process for this purpose.

However, various advanced technology desalination techniques using fossil fuel or electrical energy derived from fossil fuel are used worldwide, for example, either the thermal or the membrane process [11,27,28]. Which directly affects global warming and has high economic costs. Nevertheless, the massive shortage of fossil fuel resources, crude oil, and energy resources is attributed to the increasing tendency to replace expensive energies with renewable ones [29]. However, for remote areas that lack fresh water, the land is available at a low cost and is blessed with abundant solar radiation, so solar energy is preferred as an alternative energy source. Solar still desalination is a sustainable tool for freshwater production with a cheap and simple method using sunshine to provide drinking water, and its environmentally safe outcome is the major attraction point to research [30,31].

Consequently, solar radiation may be a viable source of renewable energy for seawater desalination in sandy deserts and semi-arid provinces where fossil fuels are also scarce and expensive. [2]. Solar distiller, on the other hand, is one of the efficient environmentally systems used for small-scale applications and is described by the easiness of operation and low construction and maintenance costs for providing drinking water, — especially in arid and semiarid regions. Their safe environmental outcome is the main attraction for research [32,33].

Modeling methods were used more extensively than experimental methods because they have benefits, no operational expenses, less time consumption, and higher reliability. According to the publications, various researchers have produced numerous efforts to improve the productivities of solar stills by implementing differing theories and improvements using experimental methods [5,16,40,27,32,34–39], while others used modelling methods [2,41–45]. Meanwhile, modelling including mathematical, machine learning [5,16,36,37,40,46–49].

Sadeghi et al. [50] and Das & Debnath [51] has developed different machine learning models to predict the temperature of the solar system. The developed models achieved a 0.9 mean relative percentage error. Another study done by Das & Debnath [51] and Sadeghi et al. [52] implemented to predict the thermal characteristics of the solar collector by using artificial neural networks (ANNs). He found that the multi-layer perceptron (MLP) model made a more accurate prediction of the collector performance than other tested models. Sadeghi showed that the multivariate regression splines (MARS) method has highly promising accuracy in predicting thermal properties of solar systems compared with other statistical methods such as the M5Model tree (M5MT) [53–55]. In addition, gene-expression programming (GEP) and evolutionary polynomial regression (EPR) methods were implemented to estimate the thermal energy from the solar cell and found that the GEP method is reliable and trustworthy and can be conveniently employed to estimate varied factors of still solar systems [56,57]. However, the mathematical models' precision is doubtful, particularly in handling highly unpredictable SR [58].

Researchers are looking for alternative techniques to overcome the limitations and throwback models for predicting the performance of solar still using climatic factors data, which would include the daily (T), (SR), cloud cover, (V), and wind patterns, as well as other operating parameters such as (T_{io}), (T_w), (T_v), and (T_t) [59–61]. In addition, the Boosting models showed reputable performance in other research fields [62–64]. According to the literature and the authors' knowledge, few studies have been conducted to evaluate the prediction performance of boosting models in this research area.

As a result, this research article aims to evaluate the prediction performance of the five new Boosting models in estimating the water productivity from the solar still system. This model's ability to anticipate efficiency while accounting for uncertainty highlights its distinctiveness. In this work, three months of experimental records were analysed. The proposed model's results were compared to the results of the cascaded forward neural network model to demonstrate its superiority (CFNN) [65].

2. Experimental set-up

(GMI) the company manufactures the entire system, including the solar stills. The instrumentation is divided into the saline water barrel and the solar still. The system was located at a 30° angle with the horizontal, as shown in Fig. 1; the main design details can be seen in Abujazar et al. [28].

The inclined stepped solar still has internal dimensions of L 1.8 m, W 1.2 m, and H 0.20 m and is made up of 28 trays with 0.6 m height and 1.2 m length. Copper sheets were used to make the trays. The trays were placed in a stainless-steel box insulated with sawdust; the sawdust layer was 6 cm thick from the system's sides and bottom. The solar still was designed with a smaller footprint and a more effective evaporation area [28].

The system has a lower footprint and a more significant positive evaporation area. The system receives seawater from a barrel in black colour, gravity fed, and well-ordered with a water level sensor to control the water level at 3cm inside the trays. When SR reaches the glass cover, evaporation occurs and precipitates in the glass cover.

2.1. Experimental procedure

The system was designed, assembled, and evaluated outside Engineering and Built Environment department, Universiti Kebangsaan Malaysia, Malaysia (Latitude 2.939671°N and Longitude 101.78784°E) during this study, which lasted 12 h each day, from 8:00 morning to 19:00 evening time, for twelve runs (at 5 d/run) across three months, from September 27th to December 23rd, 2016. Because the climate profile in a tropical climate is practically steady throughout the year, a little deviance of the key responses (SR, T and humidity) is predictable for the further time through the year [66]. Furthermore, various periods of desalination research have been carried out by scientists, Ismail [67] carried out a 6-d open-air experiment, while Ismail et al. [68] used a 14-d laboratory experiment. While Hanson et al. [37] integrated basin solar still with a sandy heat reservoir was tested in Iranian climatic conditions for 3 d.

During the experiments, various data were recorded. These parameters were SR, T_{io}, T_v, T_w, and T_b. The measurements are taken every hour, as well as the collected freshwater productivity. Other meteorological parameters, such as T, humidity, V, and cloud cover, were obtained from the "AccuWeather" climatological internet page. [69]. All experimental data were taken to assess the achievement of system Bangi, Selangor, Malaysia metrological condition.

The stepped solar still studied in this research has proven productivity almost reached 4.4 L/m^2·d, with highest production rate comparing with other research work produced (1.27, 1.37, 1.4, 1.65 L/m^2·d) [70], (2.87, 3.55, 3.93, 3.23 L/m^2·d) [71].

The productivity of the system stays high throughout the day, which refers to the high thermal capacity of the copper trays material inside the system

The system studied in this research has proven hugely successful in producing fresh drinking water in accordance with WHO standards and Malaysia's NDWQS. The inclined solar still was efficient in removing physicochemical and biological contaminations, where it produced distilled water free from 99.98% TDS, 99.7% TSS, 100% salinity, 99.98% electrical conductivity, 98.96% turbidity, 99.98% Cl^-, 99.98% Na^+, 99.94% Mg^{2+}, 99.98% SO_4^{2-}, 99.87% Ca^{2+}, and 99.94% K^+, 43.56% BOD, and 99.6% NH_3–N as shown in published articles [1,28,32].

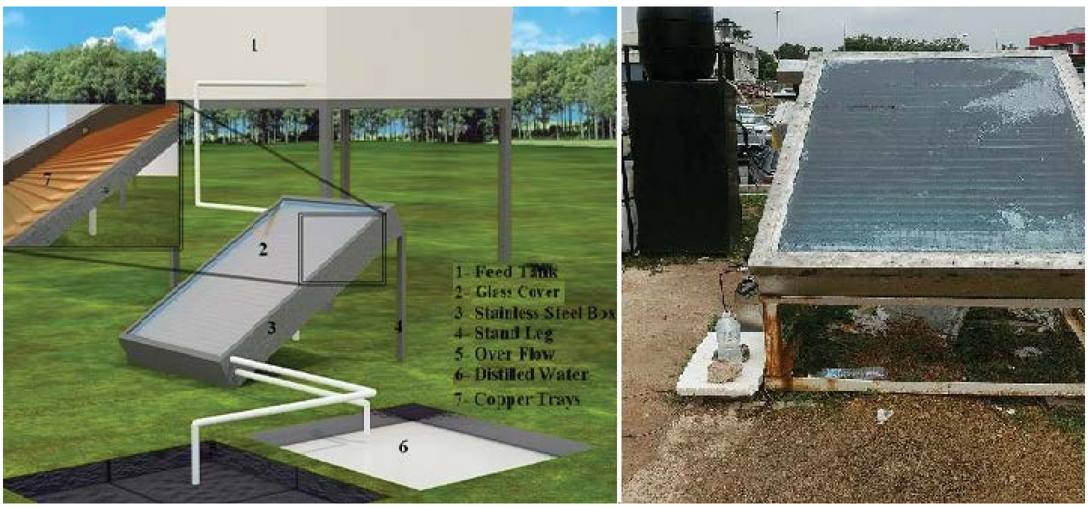

Fig. 1. Inclined system with the main components.

2.2. Measuring system

A small open space was opened at the corner of the system to install the thermocouples for monitoring temperature differences at various places by a Thermocouple Maltec-T device (Type-K). As indicated in Fig. 2, the sensors are put in various locations to monitor various temperature points. A device of Tenmars TM-750 sensor for measuring SR, and data are collected using a data recorder (OHKURA). Water flows through gravity forces, regulated by an electronic valve coupled with a water sensor.

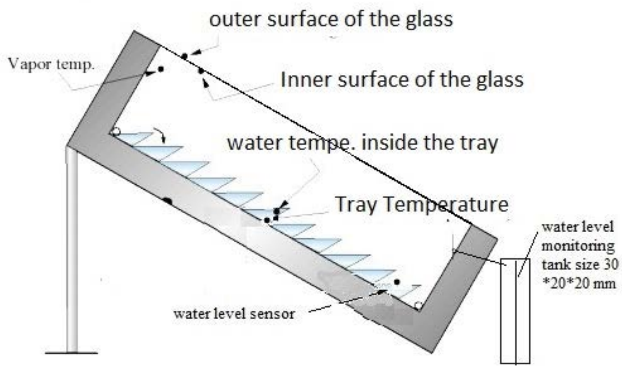

Fig. 2. Position of different points for measuring temperature variations.

3. Modelling of the system's productivity

Modelling, utilized to examine system behavior and optimize its components for improved performance, is significant during solar system design. This work employed five boosting algorithms to forecast the system productivity: categorical boosting (CatBoost), adaptive boosting (AdaBoost), extreme gradient boosting (XGBoost), gradient boosting, and LightGBM. Boosting algorithms are ensemble learning methods that encompass a family of methods [72].

3.1. Adaptive boosting

This is the first practical Boosting algorithm created by Freund and Schapire [73]. AdaBoost was created based on generating a robust model by combining many weak models [74]. AdaBoost creates an initial decision tree-based model with equally weighted samples for each leaf. The weak models are created iteratively until maximum accuracy is obtained, with fewer errors than the previous one [75].

3.2. Categorical boosting

It is a gradient boosting decision tree (GBDT) approach that can handle categorical data to reduce overfitting. CatBoost was presented by Dorogush, Ershov, and Gulin (2018). While CatBoost performs well with categorical features, the efficiency of the model increases in the absence of categoric features [76,77].

3.3. Gradient boosting machine

Gradient boosting machine (GBM) was proposed by Friedman in 2001. GBM, an ensemble algorithm, where many decision trees are trained sequentially. GBM is an iterative ensemble procedure used in supervised machine learning tasks such as classification and regression [78,79]. GBM can be used for regression analysis if the target is continuous data and for classification if the target is categorical data. The model generates binary trees to improve the performance of the previous one by eliminating errors [80]. GBM has significantly succeeded in various machine learning and data mining problems [81].

3.4. Extreme gradient boosting

XGBoost was described as a scalable end-to-end tree boosting system by Chen and Guestrin in 2016. XGBoost expresses an efficient implementation of gradient boosting principles [82]. XGBoost supervises learning challenges such as classification and regression [83]. The most significant factor behind the success of XGBoost, is the practical usage of computing resources and processing speed [74].

3.5. LightGBM

LightGBM is an open-source algorithm based on the decision tree algorithm developed by Microsoft [84]. LightGBM is used for ranking, classification, and regression problems. LightGBM includes gradient-based one-side sampling (GOSS) and exclusive feature bundling (EFB) techniques to deal with large numbers of data samples and large numbers of features [85]. LightGBM uses histogram-based algorithms to reduce memory consumption and significantly speed up the training process. Unlike other algorithms, LightGBM stands out with its leaf-wise growth strategy instead of checking the previous ones for each new leaf. LightGBM can process big data with higher efficiency and lower false error rates [86,87]. In several studies, it has been shown that LightGBM has a significantly better performance and much better accuracy compared to other methods [88,89].

3.6. Proposed boosting algorithms model

This research effort employed based on metrological data of Malaysia to highlight the potential benefit of these 5 Boosting suggested models Fig. 3 presents the location of the study area to highlight the potential benefit of these 5 Boosting suggested models. The dataset was created using meteorological data and displaying Malaysian weather hourly from 7 a.m. to 8 p.m. from September 27th to December 23rd, 2016.

The established dataset includes the nine explanatory factors: (V), cloud cover, humidity, (T), (SR), (T_{io}), (T_v), (T_w), and (T_t). In addition, the dependent factor was the productivity of solar still. The Box and Whisker plots for the productivity factor and explanatory factors are represented in Figs. 5–7, exhibiting the first, second (median), and third quartiles with minimum and maximum values. Figs. 4–6 display a considerable fluctuation for the chosen factors during the day's hours, either in the morning, noon, or evening. The higher variations in the productivity of solar still may be attributed to the high fluctuations in all selected meteorological factors, especially in the morning and noon.

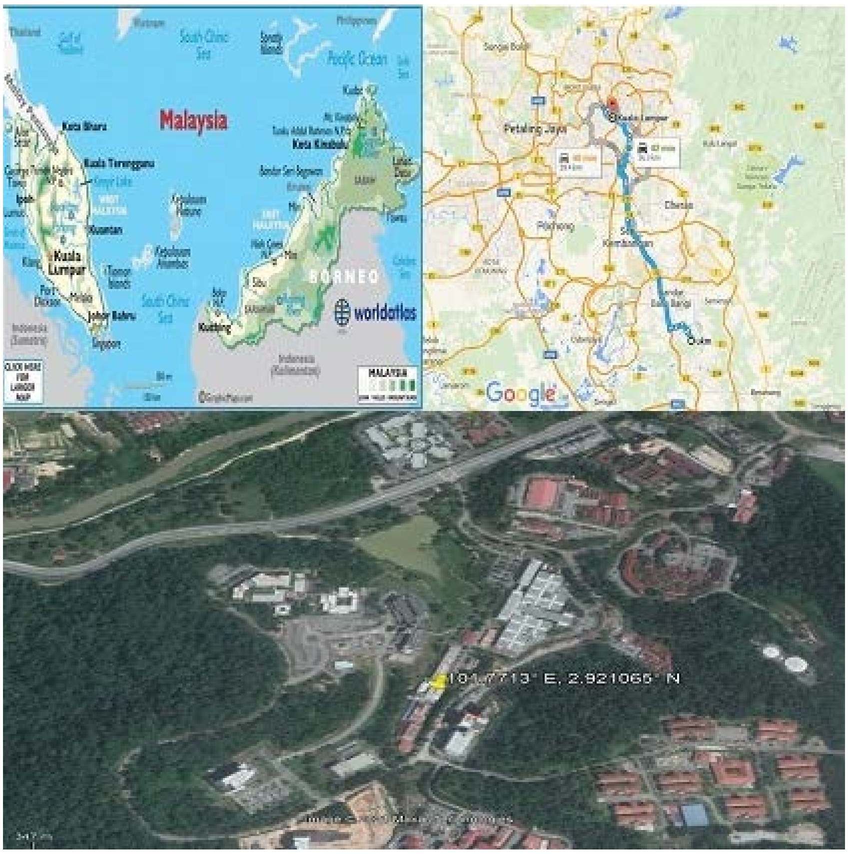

Fig. 3. Location of the case study.

In contrast, the data shows autocorrelation between the factors and productivity, which means the productivity values change similarly with some factors such as ambient temperature *T*, and solar radiation SR.

The formerly developed dataset, which included nine explanatory factors and one dependent variable with 720 samples, was used to develop the prediction models based on the 5 machine learning methods. The Google Collab platform was utilized to develop the five models. The methods were developed as follows:

3.6.1. Dataset pre-processing

Both explanatory and dependent variable datasets were split into two sets, 70% for training (504 samples) and 30% for testing (216 samples). Due to the collected dataset being measured at fixed time intervals, the "TimeSeriesSplit" function from Sklearn library is employed to divide the data instead of the classic data split methods.

3.6.2. Model developments

Different boosted trees were implemented to fit the models. The models' performance is enhanced by applying hyperparameters optimization. In addition, an optimizer applied on the number of boosted trees (from 1 to 100) to check the performance of each boosting method. The number of boosted trees were CatBoost (35), AdaBoost (50), XGBoost (35), GradientBoosting (100), LightGBM (30). The early stopping was enabled to prevent the training processing from the complexity.

3.6.3. Model testing

Both training and testing dataset were used to predict productivity using the 5 methods; then, the estimated data obtained are compared to the measured data. Various evaluation metrics such as root mean square error (RMSE), R^2, mean absolute error (MAE), median absolute error (MedAE), and mean squared error (MSE) were calculated to compare the methods.

After employing 5 boosting algorithms, the best models accuracies are shown in Table 1 and Fig. 7. The table presents the errors and accuracies of training and testing datasets for the 5 Boosting methods. GradientBoosting and LightGBM methods achieved the best model. GradientBoosting showed 95% R^2 accuracy and 39.57 RMSE error in the testing dataset. LightGBM algorithm demonstrated 94% R^2 accuracy and 40.07 RMSE error in the testing dataset. The results of GradientBoosting method indicate that this method performed very well in the application of modelling the productivity of solar still. XGBoost method showed the worst performance among the models, with 91% R^2 accuracy and 49.62 RMSE error over the testing dataset. Although the AdaBoost algorithm showed a relatively good R^2 accuracy of 93%, their RMSE error (44.28) is less compared with GradientBoosting regressor (39.57).

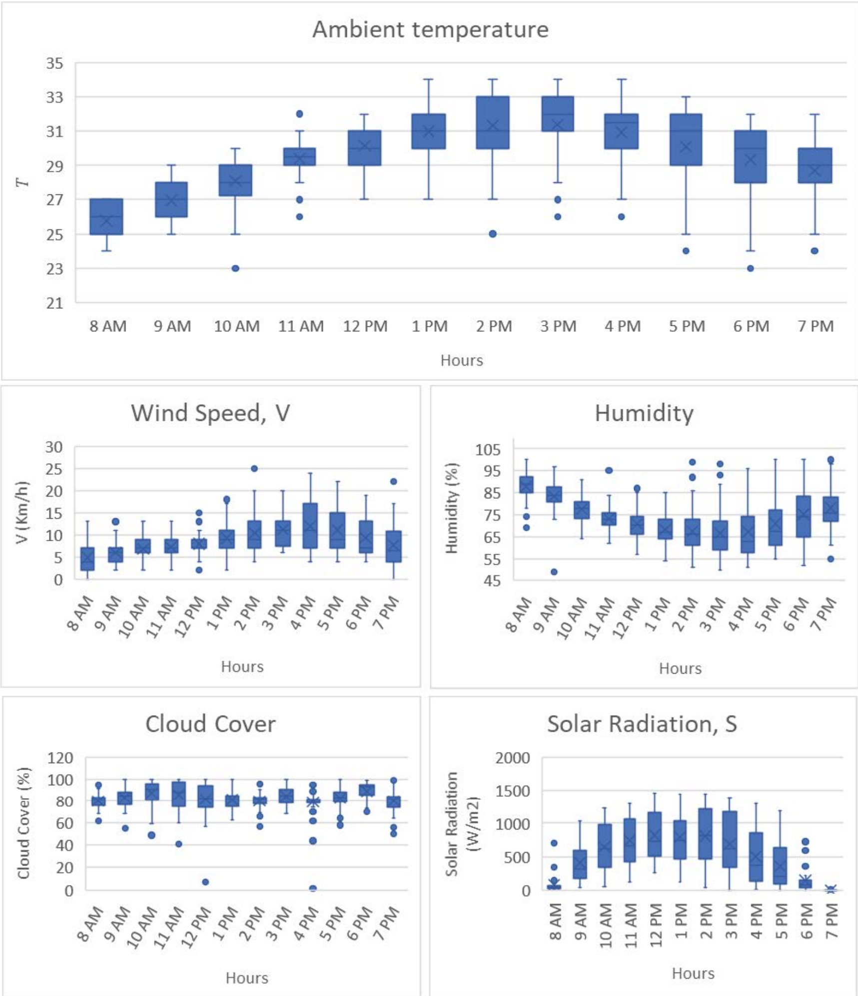

Fig. 4. Explanatory factors: wind speed, cloud cover, humidity, ambient temperature, solar radiation.

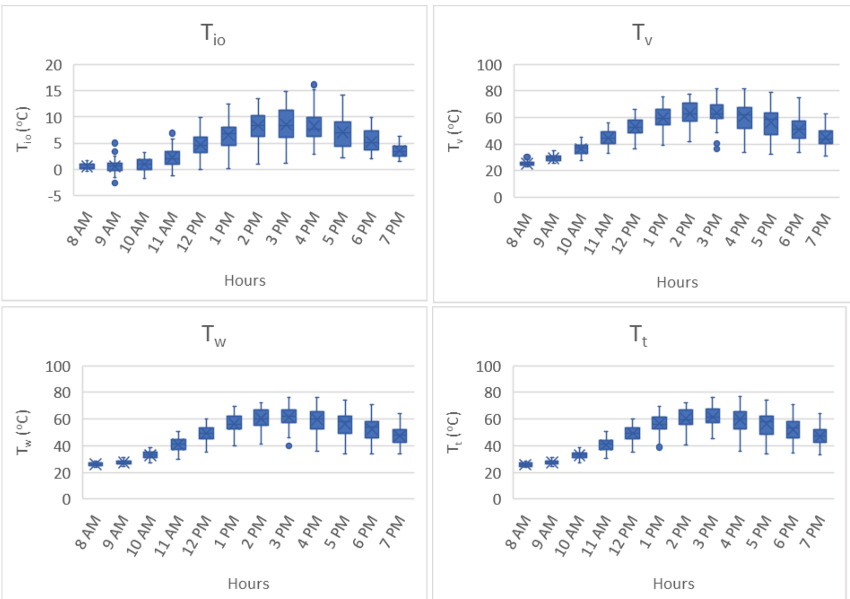

Fig. 5. Explanatory factors: T_{io}, T_v, T_w and T_t.

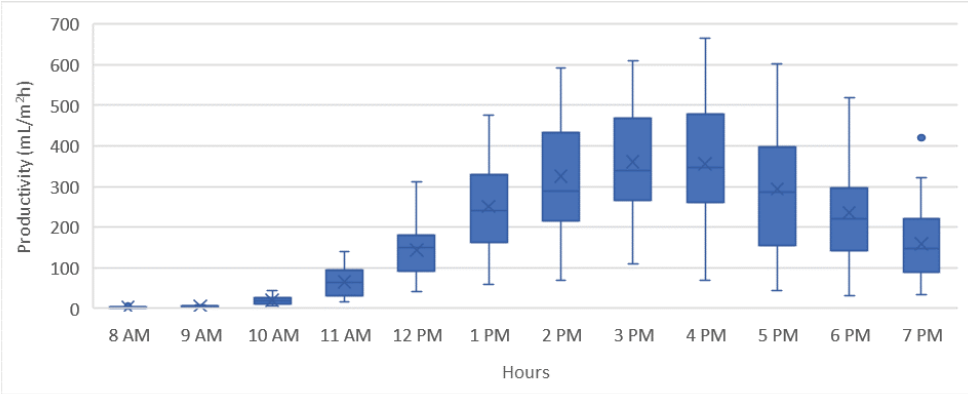

Fig. 6. Dependent factor is the productivity of solar still.

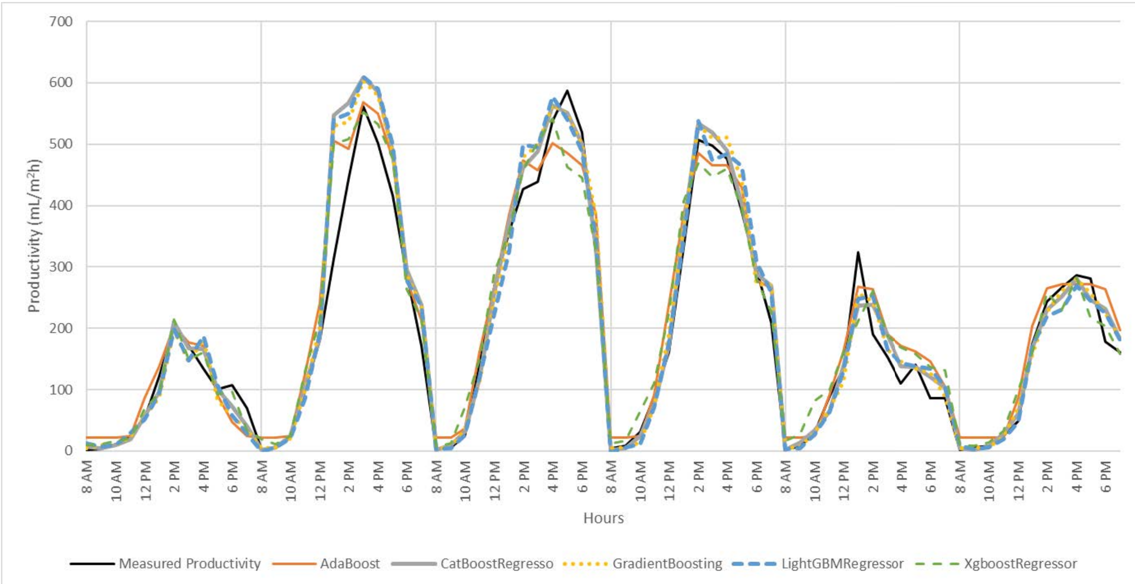

Fig. 7. Compression between measured and predicted solar still productivity by five boosting models based on the testing datasets.

Table 1
Statistical errors of CatBoost, AdaBoost, XGBoost, gradient boosting, and LightGBM regressors in the prediction of the yield of solar still

Methods	Dataset	CatBoost	AdaBoost	XGBoost	GradientBoosting	LightGBM
RMSE	Training	6.55	34.07	27.35	16.13	13.59
	Testing	41.51	44.28	49.62	39.57	40.07
R^2	Training	1.00	0.96	0.97	0.99	0.99
	Testing	0.94	0.93	0.91	0.95	0.94
MAE	Training	4.77	27.42	18.05	10.92	8.58
	Testing	25.61	32.06	32.69	25.84	26.14
MedAE	Training	3.40	20.30	10.70	6.95	4.98
	Testing	13.88	19.12	18.64	17.18	16.23
MSE	Training	42.85	1,160.67	748.23	260.17	184.67
	Testing	1723.40	1,960.87	2,461.63	1,565.82	1,605.60
MAPE	Training	0.14	1.06	0.42	0.20	0.14
	Testing	0.27	1.09	0.66	0.32	0.30

Bold values identifies the lowest model errors.

In general, the GradientBoosting method is more accurate in estimating the system's yield than the CatBoost, AdaBoost, XGBoost, and LightGBM methods.

A scatter plot of the testing datasets utilizing the 5 regressors is shown in Fig. 8. The scatter plot shows the relationship between the measured estimated and productivity based on 5 regressors. The dots distribution shows a strong positive and linear relationship between the measured estimated and productivity. Most of the dots appear around the diagonal line, representing the perfect prediction of still productivity. Although the scatter plots contain outliers, these values are almost fixed between the five methods. The fixed position of outliers could be attributed to the model error margin or the measurement process during the experiments.

Fig. 9 demonstrates the influence of the 9 independent factors on the productivity prediction of solar still based on the CatBoost, AdaBoost, XGBoost, GradientBoosting, and LightGBM regressors. All models show that the solar radiation factor is the most influencing factor. In addition, three

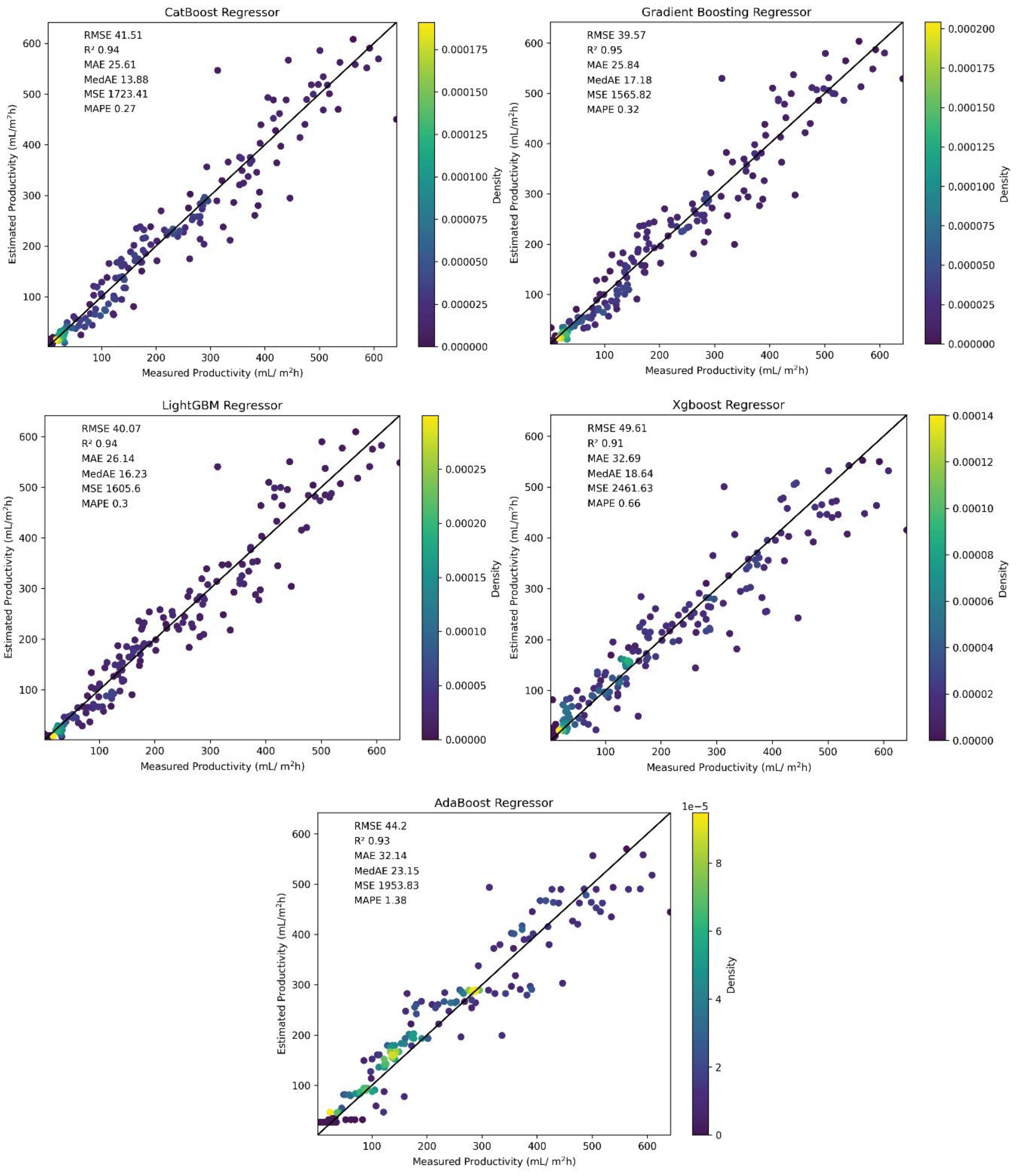

Fig. 8. Scatter plots show the differences between the measured and estimated productivity for the five models: CatBoost, AdaBoost, XGBoost, gradient boosting, and LightGBM.

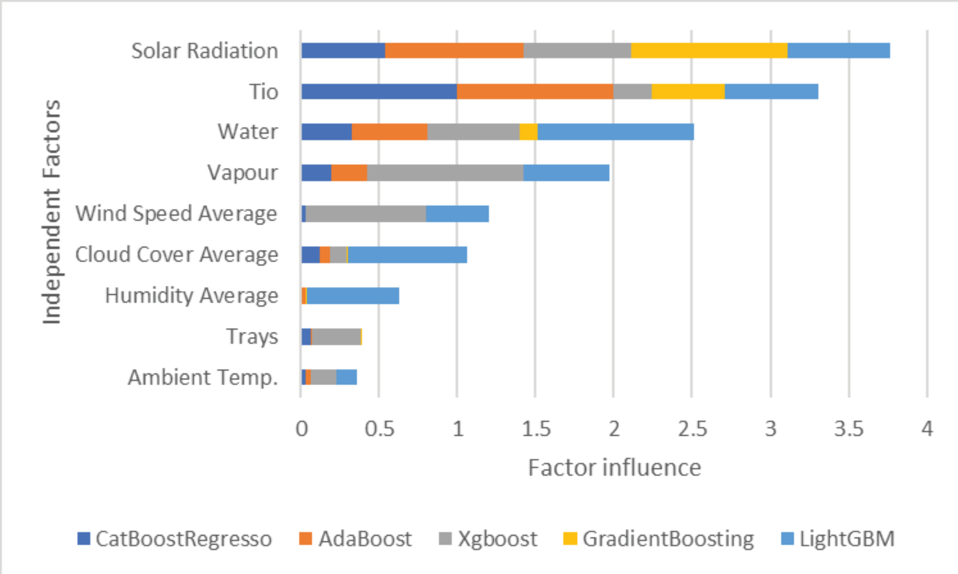

Fig. 9. Influence of the independent variables on different prediction models for the prediction of still productivity.

factors: T_{io}, T_v, and T_w showed a very significant influence on the dependent variables. In addition, the ambient temperature, humidity, and Tray's water temperature factors showed less influence. The wind speed and cloud cover variables were less important.

According to the former research works, Abujazar et al. [65], have implemented cascaded forward neural network (CFNN) as a machine learning method to predict the yield of the solar system. The CFNN method demonstrated 41.01 RMSE error, while our models using GradientBoosting regressor showed 39.570 RMSE error and high accuracy in predicting the solar still productivity.

4. Conclusion

In this work, five boosting strategies were developed, and their prediction performance was evaluated: categorical boosting (CatBoost), adaptive boosting (AdaBoost), extreme gradient boosting (XGBoost), gradient boosting machine (GBM), and gradient boosting machine (LightGBM). The solar still productivity from an inclined stepped solar still system was projected using boosting approaches. The recorded dataset was hyper-parameter adjusted using the hold-out validation method, with 70% of the dataset allocated for training and 30% for testing. Cloud cover V, humidity, T, SR, T_{io}, T_v, T_w, and T_t were the input variables for the five models. Six accuracy or statistical error measures were used to evaluate the constructed regressors: RMSE, R^2, MAE, MedAE, MSE, and mean absolute percentage error (MAPE). According to the results, the best prediction was achieved using the GradientBoosting regressor, which demonstrated 95% R^2 accuracy and 39.57 RMSE error. The LightGBM regressor had 94% R^2 accuracy in the testing dataset and a 40.07 RMSE error.

In general, the boosting strategies produced better results and lower error rates while developing productivity prediction models for solar still. This investigation shows that the GradientBoosting approach worked well in predicting solar still yield.

Abbreviations

RMSE	—	Root mean square error
R^2	—	Coefficient of determination
MSE	—	Mean squared error
MedAE	—	Median absolute error
MAPE	—	Mean absolute percentage error
MAE	—	Mean absolute error

Symbols

V	—	Wind velocity, m/s
SR	—	Solar radiation, W/m^2
T_w	—	Water temperature, °C
T_g	—	Glass temperature, °C
T_b	—	Basin Temperature, °C
T_a	—	Ambient temperature, °C
T_{io}	—	Glass inner and outer cover temperatures, °C
T_v	—	Vapour temperature, °C

Acknowledgment

The authors thank the numerous individuals and organizations that generously supported the study, particularly Universiti Kebangsaan Malaysia through its GUP-2016-020 grant.

References

[1] M.S.S. Abujazar, S. Fatihah, A.E. Kabeel, S. Sharil, S.S. Abu Amr, Evaluation quality of desalinated water derived from inclined copper-stepped solar still, Desal. Water Treat., 131 (2018) 83–95.
[2] A.F. Mashaly, A.A. Alazba, A.M. Al-Awaadh, M.A. Mattar, Predictive model for assessing and optimizing solar still performance using artificial neural network under hyper arid environment, Sol. Energy, 118 (2015) 41–58.

[3] B. Leveque, J.B. Burnet, S. Dorner, F. Bichai, Impact of climate change on the vulnerability of drinking water intakes in a northern region, Sustainable Cities Soc., 66 (2021) 102656, doi: 10.1016/j.scs.2020.102656.
[4] S.S. Ray, R.K. Verma, A. Singh, S. Myung, Y.-I. Park, I.-C. Kim, H.K. Lee, Y.-N. Kwon, Exploration of time series model for predictive evaluation of long-term performance of membrane distillation desalination, Process Saf. Environ. Prot., 160 (2022) 1–12, doi: 10.1016/j.psep.2022.01.058.
[5] S.W. Sharshir, G. Peng, L. Wu, N. Yang, F.A. Essa, A.H. Elsheikh, S.I.T. Mohamed, A.E. Kabeel, Enhancing the solar still performance using nanofluids and glass cover cooling: experimental study, Appl. Therm. Eng., 113 (2017) 684–693.
[6] K. Rhoden, J. Alonso, M. Carmona, M. Pham, A.N. Barnes, Twenty years of waterborne and related disease reports in Florida, USA, One Health, 13 (2021) 100294, doi: 10.1016/j.onehlt.2021.100294.
[7] S. Nazari, M. Bahiraei, H. Moayedi, H. Safarzadeh, A proper model to predict energy efficiency, exergy efficiency, and water productivity of a solar still via optimized neural network, J. Cleaner Prod., 277 (2020) 123232, doi: 10.1016/j.jclepro.2020.123232.
[8] S. Khanmohammadi, S. Khanjani, Experimental study to improve the performance of solar still desalination by hydrophobic condensation surface using cold plasma technology, Sustainable Energy Technol. Assess., 45 (2021) 101129, doi: 10.1016/j.seta.2021.101129.
[9] A.D. Khawaji, I.K. Kutubkhanah, J.M. Wie, Advances in seawater desalination technologies, Desalination, 221 (2008) 47–69.
[10] S.K. Suraparaju, R. Dhanusuraman, S.K. Natarajan, Performance evaluation of single slope solar still with novel pond fibres, Process Saf. Environ. Prot., 154 (2021) 142–154.
[11] O. Bait, Exergy, environ–economic and economic analyses of a tubular solar water heater assisted solar still, J. Cleaner Prod., 212 (2019) 630–646.
[12] A.E. Kabeel, R. Sathyamurthy, A.M. Manokar, S.W. Sharshir, F.A. Essa, A.H. Elshiekh, Experimental study on tubular solar still using graphene oxide nano particles in phase change material (NPCM's) for fresh water production, J. Energy Storage, 28 (2020) 101204, doi: 10.1016/j.est.2020.101204.
[13] M. Mukherjee, S. Roy, K. Bhowmick, S. Majumdar, I. Prihatiningtyas, B. Van der Bruggen, P. Mondal, Development of high performance pervaporation desalination membranes: a brief review, Process Saf. Environ. Prot., 159 (2022) 1092–1104.
[14] WWAP, The United Nations World Water Development Report 2015: Water For a Sustainable World, Paris, 2016. Available at: http://www.unesco.org/new/en/natural-sciences/environment/water/wwap/wwdr/2015-water-for-a-sustainable-world/
[15] L. Mu, L. Chen, L. Lin, Y.H. Park, H. Wang, P. Xu, K. Kota, S. Kuravi, An overview of solar still enhancement approaches for increased freshwater production rates from a thermal process perspective, Renewable Sustainable Energy Rev., 150 (2021) 111458, doi: 10.1016/j.rser.2021.111458.
[16] A.E. Kabeel, S.A. El-Agouz, R. Sathyamurthy, T. Arunkumar, Augmenting the productivity of solar still using jute cloth knitted with sand heat energy storage, Desalination, 443 (2018) 122–129.
[17] N. Najid, S. Fellaou, S. Kouzbour, B. Gourich, A. Ruiz-García, Energy and environmental issues of seawater reverse osmosis desalination considering boron rejection: a comprehensive review and a case study of exergy analysis, Process Saf. Environ. Prot., 156 (2021) 373–390.
[18] R. Lokk, S.M. Alsadaie, I.M. Mujtaba, Dynamic simulation of once-through multistage flash (MSF-OT) desalination process: effect of seawater temperature on the fouling mechanism in the heat exchangers, Comput. Chem. Eng., 155 (2021) 107515, doi: 10.1016/j.compchemeng.2021.107515.
[19] H. Lv, Y. Wang, L. Wu, Y. Hu, Numerical simulation and optimization of the flash chamber for multi-stage flash seawater desalination, Desalination, 465 (2019) 69–78.
[20] A. Darmawan, L. Karlina, I. Khairunnisak, R.E. Saputra, C. Azmiyawati, Y. Astuti, A.P. Noorita, Hydrophobic silica thin film derived from dimethyldimethoxysilane-tetraethylorthosilicate for desalination, Thin Solid Films, 734 (2021) 138865, doi: 10.1016/j.tsf.2021.138865.
[21] H. You, X. Zhang, D. Zhu, C. Yang, P. Chammingkwan, T. Taniike, Advantages of polydopamine coating in the design of ZIF-8-filled thin-film nanocomposite (TFN) membranes for desalination, Colloids Surf., A, 629 (2021) 127492, doi: 10.1016/j.colsurfa.2021.127492.
[22] B.S. Al-Anzi, A. Al-Rashidi, L. Abraham, J. Fernandes, A. Al-Sheikh, A. Alhazza, Brine management from desalination plants for salt production utilizing high current density electrodialysis-evaporator hybrid system: a case study in Kuwait, Desalination, 498 (2021) 114760, doi: 10.1016/j.desal.2020.114760.
[23] G. Zheng, J. Jiang, X. Wang, W. Li, J. Liu, G. Fu, L. Lin, Nanofiber membranes by multi-jet electrospinning arranged as arc-array with sheath gas for electrodialysis applications, Mater. Des., 189 (2020) 108504, doi: 10.1016/j.matdes.2020.108504.
[24] X. Huang, T. Ke, Y. Li, X. Ling, Experimental investigation and optimization of total energy consumption in humidification-dehumidification system, Energy Procedia, 158 (2019) 3488–3493.
[25] M.M. Farid, S. Parekh, J.R. Selman, S. Al-Hallaj, Solar desalination with a humidification-dehumidification cycle: mathematical modeling of the unit, Desalination, 151 (2003) 153–164.
[26] S.M. Parsa, A. Rahbar, M.H. Koleini, S. Aberoumand, M. Afrand, M. Amidpour, A renewable energy-driven thermoelectric-utilized solar still with external condenser loaded by silver/nanofluid for simultaneously water disinfection and desalination, Desalination, 480 (2020) 114354, doi: 10.1016/j.desal.2020.114354.
[27] S. Sethi, S. Walker, J. Drewes, P. Xu, Existing & emerging concentrate minimization & disposal practices for membrane systems, Florida Water Resour. J., 38 (2006) 40–45.
[28] M.S.S. Abujazar, S. Fatihah, A.E. Kabeel, Seawater desalination using inclined stepped solar still with copper trays in a wet tropical climate, Desalination, 423 (2017) 141–148.
[29] A. Cipollina, E. Tzen, V. Subiela, M. Papapetrou, J. Koschikowski, R. Schwantes, Renewable energy desalination: performance analysis and operating data of existing RES desalination plants, Desal. Water Treat., 55 (2015) 3126–3146.
[30] A. El-Bahi, D. Inan, Analysis of a parallel double glass solar still with separate condenser, Renewable Energy, 17 (1999) 509–521.
[31] G.M. Ayoub, L. Malaeb, Economic feasibility of a solar still desalination system with enhanced productivity, Desalination, 335 (2014) 27–32.
[32] M.S.S. Abujazar, S. Fatihah, E.R. Lotfy, A.E. Kabeel, S. Sharil, Performance evaluation of inclined copper-stepped solar still in a wet tropical climate, Desalination, 425 (2018) 94–103.
[33] H. Sharon, Energy, exergy, environmental benefits and economic aspects of novel hybrid solar still for sustainable water distillation, Process Saf. Environ. Prot., 150 (2021) 1–21, doi: 10.1016/j.psep.2021.04.003.
[34] F.A. Essa, A.S. Abdullah, Z.M. Omara, Improving the performance of tubular solar still using rotating drum – experimental and theoretical investigation, Process Saf. Environ. Prot., 148 (2021) 579–589.
[35] A.A. AL-Karaghouli, W.E. Alnaser, Experimental comparative study of the performances of single and double basin solar-stills, Appl. Energy, 77 (2004) 317–325.
[36] B.A. Akash, M.S. Mohsen, W. Nayfeh, Experimental study of the basin type solar still under local climate conditions, Energy Convers. Manage., 41 (2000) 883–890.
[37] F.F. Tabrizi, A.Z. Sharak, Experimental study of an integrated basin solar still with a sandy heat reservoir, Desalination, 253 (2010) 195–199.
[38] O.O. Badran, Experimental study of the enhancement parameters on a single slope solar still productivity, Desalination, 209 (2007) 136–143.

[39] A.E. Kabeel, M. Abdelgaied, A. Eisa, Enhancing the performance of single basin solar still using high thermal conductivity sensible storage materials, J. Cleaner Prod., 183 (2018) 20–25.
[40] F. Ketabchi, S. Gorjian, S. Sabzehparvar, Z. Shadram, M.S. Ghoreishi, H. Rahimzadeh, Experimental performance evaluation of a modified solar still integrated with a cooling system and external flat-plate reflectors, Sol. Energy, 187 (2019) 137–146.
[41] V. Velmurugan, S. Pandiarajan, P. Guruparan, L.H. Subramanian, C.D. Prabaharan, K. Srithar, Integrated performance of stepped and single basin solar stills with mini solar pond, Desalination, 249 (2009) 902–909.
[42] D. Kumar, P. Kumar, Mathematical modeling of conventional solar still coupled with solar air heater, IJISET – Int. J. Innovative Sci. Eng. Technol., 1 (2014) 379–385.
[43] A.F. Mashaly, A.A. Alazba, ANFIS modeling and sensitivity analysis for estimating solar still productivity using measured operational and meteorological parameters, Water Sci. Technol. Water Supply, 18 (2018) 1437–1448.
[44] S. Shoeibi, N. Rahbar, A. Abedini Esfahlani, H. Kargarsharifabad, Improving the thermoelectric solar still performance by using nanofluids– experimental study, thermodynamic modeling and energy matrices analysis, Sustainable Energy Technol. Assess., 47 (2021) 101339, doi: 10.1016/j.seta.2021.101339.
[45] M. Keshtkar, M. Eslami, K. Jafarpur, Effect of design parameters on performance of passive basin solar stills considering instantaneous ambient conditions: a transient CFD modeling, Sol. Energy, 201 (2020) 884–907.
[46] M. Feng, Y. Tao, A Mathematical Model for the Performance of a Horizontal Convective Solar Still, ASME 2005 Summer Heat Transfer Conference collocated with the ASME 2005 Pacific Rim Technical Conference and Exhibition on Integration and Packaging of MEMS, NEMS, and Electronic Systems, 2015, pp. 1–9.
[47] Y.A.F. El-Samadony, A.E. Kabeel, Theoretical estimation of the optimum glass cover water film cooling parameters combinations of a stepped solar still, Energy, 68 (2014) 744–750.
[48] M. Mohanraj, S. Jayaraj, C. Muraleedharan, Applications of artificial neural networks for refrigeration, air-conditioning and heat pump systems – a review, Renewable Sustainable Energy Rev., 16 (2012) 1340–1358.
[49] A.H. Elsheikh, V.P. Katekar, O.L. Muskens, S.S. Deshmukh, M.A. Elaziz, S.M. Dabour, Utilization of LSTM neural network for water production forecasting of a stepped solar still with a corrugated absorber plate, Process Saf. Environ. Prot., 148 (2021) 273–282.
[50] G. Sadeghi, A.L. Pisello, S. Nazari, M. Jowzi, F. Shama, Empirical data-driven multi-layer perceptron and radial basis function techniques in predicting the performance of nanofluid-based modified tubular solar collectors, J. Cleaner Prod., 295 (2021) 126409, doi: 10.1016/j.jclepro.2021.126409.
[51] P. Das, A. Debnath, Reactive orange 12 dye adsorption onto magnetically separable $CaFe_2O_4$ nanoparticles synthesized by simple chemical route: kinetic, isotherm and neural network modeling, Water Pract. Technol., 16 (2021, doi: 10.2166/wpt.2021.064.
[52] G. Sadeghi, S. Nazari, M. Ameri, F. Shama, Energy and exergy evaluation of the evacuated tube solar collector using Cu_2O/water nanofluid utilizing ANN methods, Sustainable Energy Technol. Assess., 37 (2020) 100578, doi: 10.1016/j.seta.2019.100578.
[53] G. Sadeghi, M. Najafzadeh, M. Ameri, M. Jowzi, A case study on copper-oxide nanofluid in a back pipe vacuum tube solar collector accompanied by data mining techniques, Case Stud. Therm. Eng., 32 (2022) 101842, doi: 10.1016/j.csite.2022.101842.
[54] G. Sadeghi, M. Najafzadeh, M. Ameri, Thermal characteristics of evacuated tube solar collectors with coil inside: an experimental study and evolutionary algorithms, Renewable Energy, 151 (2020), doi: 10.1016/j.renene.2019.11.050.
[55] A. Debnath, M. Majumder, M. Pal, N.S. Das, K.K. Chattopadhyay, B. Saha, Enhanced Adsorption of hexavalent chromium onto magnetic calcium ferrite nanoparticles: kinetic, isotherm, and neural network modeling, J. Dispersion Sci. Technol., 37 (2016) 1141100, doi: 10.1080/01932691.2016.1141100.
[56] G. Sadeghi, M. Najafzadeh, H. Safarzadeh, Utilizing gene-expression programming in modelling the thermal performance of evacuated tube solar collectors, J. Energy Storage, 30 (2020) 101546, doi: 10.1016/j.est.2020.101546.
[57] M. Bhowmik, K. Deb, A. Debnath, B. Saha, Mixed phase Fe_2O_3/Mn_3O_4 magnetic nanocomposite for enhanced adsorption of methyl orange dye: neural network modeling and response surface methodology optimization, Appl. Organomet. Chem., 32 (2018), doi: 10.1002/aoc.4186.
[58] R. Eke, H. Demircan, Performance analysis of a multi crystalline Si photovoltaic module under Mugla climatic conditions in Turkey, Energy Convers. Manage., 65 (2013) 580–586.
[59] R. Ata, Artificial neural networks applications in wind energy systems: a review, Renewable Sustainable Energy Rev., 49 (2015) 534–562.
[60] N.I. Santos, A.M. Said, D.E. James, N.H. Venkatesh, Modeling solar still production using local weather data and artificial neural networks, Renewable Energy, 40 (2012) 71–79.
[61] M.A. Hamdan, R.A. Haj Khalil, E.A.M. Abdelhafez, Comparison of neural network models in the estimation of the performance of solar still under Jordanian climate, J. Clean Energy Technol., 1 (2014) 238–242.
[62] R. Barzegar, A. Asghari Moghaddam, J. Adamowski, B. Ozga-Zielinski, Multi-step water quality forecasting using a boosting ensemble multi-wavelet extreme learning machine model, Stochastic Environ. Res. Risk Assess., 32 (2018) 799–813.
[63] A. Mosavi, F.S. Hosseini, B. Choubin, Ensemble Boosting and Bagging Based Machine Learning Models for Groundwater Potential Prediction Content courtesy of Springer Nature, Terms of Use Apply, Rights Reserved, Content Courtesy of Springer Nature, Terms of Use Apply, Rights Reserved, 2021, pp. 23–37.
[64] D.H. Nguyen, X. Hien Le, J.Y. Heo, D.H. Bae, Development of an extreme gradient boosting model integrated with evolutionary algorithms for hourly water level prediction, IEEE Access, 9 (2021) 125853–125867.
[65] M.S.S. Abujazar, S. Fatihah, I.A. Ibrahim, A.E. Kabeel, S. Sharil, Productivity modelling of a developed inclined stepped solar still system based on actual performance and using a cascaded forward neural network model, J. Cleaner Prod., 170 (2017) 147–159.
[66] T. Khatib, A. Mohamed, K. Sopian, M. Mahmoud, Solar energy prediction for Malaysia using artificial neural networks, Int. J. Photoenergy, 2012 (2012) 419504, doi: 10.1155/2012/419504.
[67] B.I. Ismail, Design and performance of a transportable hemispherical solar still, Renewable Energy, 34 (2009) 145–150.
[68] A. Hanson, W. Zachritz, K. Stevens, L. Mimbela, R. Polka, L. Cisneros, Distillate water quality of a single-basin solar still: laboratory and field studies, Sol. Energy, 76 (2004) 635–645.
[69] AccuWeather, Malaysia Weather, 2016. Available at: https://www.accuweather.com/en/my/malaysia-weather (Accessed December 23, 2016).
[70] V. Velmurugan, K.J. Naveen Kumar, T. Noorul Haq, K. Srithar, Performance analysis in stepped solar still for effluent desalination, Energy, 34 (2009) 1179–1186.
[71] R.S. Hansen, C.S. Narayanan, K.K. Murugavel, Performance analysis on inclined solar still with different new wick materials and wire mesh, Desalination, 358 (2015) 1–8.
[72] D. Opitz, R. Maclin, Popular ensemble methods: an empirical study, J. Artif. Intell. Res., 11 (1999) 169–198.
[73] Y. Freund, R.E. Schapire, A desicion-theoretic generalization of on-line learning and an application to boosting BT – computational learning theory, Comput. Learn Theory, 904 (2005) 23–37.
[74] S. Jhaveri, I. Khedkar, Y. Kantharia, S. Jaswal, Success Prediction Using Random Forest, CatBoost, XGBoost and AdaBoost for Kickstarter Campaigns, 2019 3rd International Conference on Computing Methodologies and Communication (ICCMC), IEEE, Erode, India, 2019, pp. 1170–1173. Available at: https://doi.org/10.1109/ICCMC.2019.8819828.

[75] F. Anggraeni, D. Adytia, A.W. Ramadhan, Forecasting of Wave Height Time Series Using AdaBoost and XGBoost, Case Study in Pangandaran, Indonesia, 2021 International Conference on Data Science and Its Applications (ICoDSA), IEEE, Bandung, Indonesia, 2021, pp. 97–101. Available at: https://doi.org/10.1109/ICoDSA53588.2021.9617524.
[76] R. Punmiya, S. Choe, Energy theft detection using gradient boosting theft detector with feature engineering-based preprocessing, IEEE Trans. Smart Grid, 10 (2019) 2326–2329.
[77] S. Lee, T.P. Vo, H.T. Thai, J. Lee, V. Patel, Strength prediction of concrete-filled steel tubular columns using categorical gradient boosting algorithm, Eng. Struct., 238 (2021) 112109, doi: 10.1016/j.engstruct.2021.112109.
[78] H. Lu, S.P. Karimireddy, N. Ponomareva, V. Mirrokni, Accelerating Gradient Boosting Machine, Int. Conf. Arti Cial Intell. Stat. (AISTATS), 2020, pp. 1–10. Available at: http://arxiv.org/abs/1903.08708.
[79] E.K. Sahin, Comparative analysis of gradient boosting algorithms for landslide susceptibility mapping, Geocarto Int., 37 (2022) 2441–2465.
[80] R. Costache, Q.B. Pham, M. Avand, N.T. Thuy Linh, M. Vojtek, J. Vojteková, S. Lee, D.N. Khoi, P.T. Thao Nhi, T.D. Dung, Novel hybrid models between bivariate statistics, artificial neural networks and boosting algorithms for flood susceptibility assessment, J. Environ. Manage., 265 (2020) 110485, doi: 10.1016/j.jenvman.2020.110485.
[81] A. Natekin, A. Knoll, Gradient boosting machines, a tutorial, Front. Neurorobot., 7 (2013) 1–21.
[82] A. Ibrahem Ahmed Osman, A. Najah Ahmed, M.F. Chow, Y. Feng Huang, A. El-Shafie, extreme gradient boosting (XGBoost) model to predict the groundwater levels in Selangor Malaysia, Ain Shams Eng. J., 12 (2021) 1545–1556.
[83] J. Cao, Z. Zhang, J. Du, L. Zhang, Y. Song, G. Sun, Multi-geohazards susceptibility mapping based on machine learning—a case study in Jiuzhaigou, China, Nat. Hazards, 102 (2020) 851–871.
[84] G. Ke, Q. Meng, T. Finley, T. Wang, W. Chen, W. Ma, Q. Ye, T.Y. Liu, LightGBM: a highly efficient gradient boosting decision tree, 31st Conference on Neural Information Processing Systems (NIPS 2017), Long Beach, CA, USA, 2017, pp. 3147–3155.
[85] E.K. Sahin, Assessing the predictive capability of ensemble tree methods for landslide susceptibility mapping using XGBoost, gradient boosting machine, and random forest, SN Appl. Sci., 2 (2020) 1–17.
[86] E. Al Daoud, Comparison between XGBoost, LightGBM and CatBoost using a home credit dataset, World Acad. Sci. Eng. Technol., Int. J. Comput. Inf. Eng., 13 (2019) 6–10.
[87] M. Tang, Q. Zhao, S.X. Ding, H. Wu, L. Li, W. Long, B. Huang, An improved LightGBM algorithm for online fault detection of wind turbine gearboxes, Energies, 13 (2020) 13040807, doi: 10.3390/en13040807.
[88] A. Haithm, A.Y. Saleh, A. Odabaş, Comparison of gradient boosting decision tree algorithms for CPU performance, J. Inst. Sci. Technol., 37 (2021) 157–168.
[89] Y. Wang, T. Wang, Application of improved LightGBM model in blood glucose prediction, Appl. Sci., 10 (2020), doi: 10.3390/app10093227.