

# A Performance Neural Network Model for Conventional Solar Stills via Transfer Learning

Hashim H. Migaybil <sup>a,b</sup>, Bhushan Gopaluni <sup>b</sup>

<sup>a</sup> Department of Chemical and Materials Engineering, Faculty of Engineering, King Abdulaziz University, Jeddah 21589, Saudi Arabia

<sup>b</sup> Department of Chemical and Biological Engineering, Faculty of Applied Science, The University of British Columbia, Vancouver, BC V6T1Z3, Canada

<sup>a</sup> Corresponding author, Hashim H. Migaybil (e-mail: [hmigaybil@kau.edu.sa](mailto:hmigaybil@kau.edu.sa)).

---

## Abstract

Predictive solar-desalination models are becoming more widespread using ML and AI. However, forecasting solar still water productivity based on numerous designs still needs to be improved. Herein, we used transfer learning to create precise supervised predictive ANN regression models for water productivity (L/m<sup>2</sup>.day) predictions based on literature findings. Such observation datasets from single-basin solar stills were utilized to build the random initialization ANN model. The transfer learning method was applied to the latter model by taking the learned network (weights) for fine-tuning the hyperparameters from the earlier developed novel hybrid solar still known as the source (pre-trained) ANN model, to predict the target ANN model. Based on most minor statistical errors, the pre-trained model with 5-64-64-1 architecture and ReLU activation function was the most appropriate for water productivity prediction. All created ANN models were compared to the MLR model. The results revealed that the generated target ANN model outperformed the ANN RI and MLR with OI values of 0.872, 0.834, and 0.803, respectively, in all modeling stages. The target ANN model's accuracy and generalization were sufficient. The target ANN model had residuals of forecasted distillate values of around 1%. This work discusses the significance of transfer learning to generate accurate target ANN models for predicting freshwater outputs in single-slope solar stills, which can be integrated with established theoretically tuned parameters to enhance performance and maximize distillate water yields.

*Keywords:* Solar still, Artificial neural networks (ANNs), Transfer learning, Random initialization (RI), Multiple linear regression (MLR), Distillate water

---

## 1. Introduction

Globally, water scarcity is a significant challenge and entails thoroughly exploring all available possibilities for maintaining present water supplies while investigating new alternative sources. The abundance of saline water inspires substantial prospects for recovering vast amounts of water that demand treatment from this water source [1]. There is an urgent and severe necessity to acquire fresh water from the saline/brackish or polluted water on or within the earth owing to the stability of accessible fresh water, estimated to be about 1%. Its demand is increasing daily due to the rapidly expanding industry and population growth increase [2] [3].

Desalination processes are feasible and of interest for potable water consumption; thus, they raise distillate water production, resulting in lower water consumption. The overproduction of industrial wastewater has never been as high as it is now owing to the rising demand for consumer goods and technology from various industrial and water-related manufacturing processes [4]. Hence, finding alternative approaches to pricy and energy-intensive separation methods such as: (multi-stage flash, multiple effects, solvent extraction, membrane filters, and reverse osmosis via emerging solar distillation may, therefore, contribute to fulfilling the world's increasing need for pumping purified and industrial water to the desired location [5] [6].

According to Cuce et al. [7], solar distillation systems can be classified into two groups: (i) direct systems, which use directly absorbed solar radiation energy like in solar stills. In addition, (ii) indirect systems use energy converted from solar to thermal or photovoltaic systems, such as thermal and membrane systems. The intriguing water distillation technique known as "solar stills" was found to have promise as a less costly substitute for the high-priced membrane technology [8]. This technology had been discovered in the late 19th century and had just recently entered the market for manufacturing and delivering water. Leveraging an environmentally friendly solar desalination process, offered by solar intensity energy, is a breakthrough since no wastes/emissions are created. Also, no additional energy is required, potentially lowering the costs of employing solar stills for desalination [9] [10]. Water solar distillation replicates the natural manner of "rain" through water's recurring evaporation/ condensation process [11]. The distilled and treated water produced is typically potable and of valuable high-quality owing to the complete removal of inorganic and organic pollutants and total dissolved solids (TDS) [12].

In terms of solar distillation unit design aspect, direct passive solar stills are common practice due to their simplicity and capacity to improve water production in single slope solar stills via solar radiation absorption, basin thermal insulation, and energy storage medium [12] [13] [14]. Similarly, for the optimization of solar still freshwater productivity, some critical factors that must be taken into consideration comprise but are not limited to solar radiation, ambient temperature, wind speed, desired feed water flowrate (maximum permissible water levels inside the saline water tank), feed water temperature (the difference in water/glass temperature), air humidity, basin surface area, inclination angle for the top cover, and glass transparency for solar radiation [8]. The direct passive solar still's performance was shown to be significantly impacted by a number of these factors, including the solar incident, the ambient, water, and vapor temperatures, the brine depth (2-2.5 cm), and the wall insulating film thickness (1–5 cm). Previously, these parameters have been experimentally investigated, and mathematical models were used to correlate them to the performance with high prediction accuracy [15]. A total water productivity of 2,197.4 mL/week was found to be possible from the passive design with the sensible medium energy

storage, according to research by Cuce et al. [7], who developed a novel non-insulated solar distillation unit through a sensible medium for energy storage coupled with a passive supporter reflector. Yet, the thermal insulation and cooling for the aperture glazing and exterior structure, which would drive up rates of the condensation-evaporation process via decreasing heat losses and improving convective coefficients, were the only factors affecting water productivity. Furthermore, Wang et al. [16] discovered that for forecasting productivity in tubular solar stills, seawater temperature, basin temperature, and solar radiation were among the most fundamental indicators, with 40.87%, 32.43%, and 18.2 %, respectively. The single-slope solar still was shown to be more efficient than a pyramid-shaped still in terms of the design aspect, with a 30% greater output in winter due to the reduced escape of absorbed solar intensity energy from the substantial surface cover [17].

Due to the relatively low conversion rates and high installation costs, solar radiation employment for solar stills or power conversion still needs to be extensively industrialized. Supervised machine learning (ML) and cross-validation (CV) techniques should be the subject of more innovative research. Inventing unique technical technologies capable of in-depth assessing the patent outputs and literature datasets that are now available would aid in solar harvesting, material selection, and solar desalination application [18]. These newly developed algorithms can make it easier for large-scale solar stills to be commercialized and progress in solar-desalination technologies, which will benefit the creativities' community, business fields, and technical R&D facilities [19] [20].

Even though the inherent problem of solar still is its low water distillate productivity, it utilizes a sustainable and renewable energy source (pollution-free) to yield high-quality treated water. It is vital to select solar still systems wisely based on their productivity, which is considered a prime design goal to maintain design simplicity and operation feasibility. However, this is challenging due to various measurements and multiple heat transfer calculations [21]. There is a rising trend toward using ML/AI algorithms to build numerical models or complicated computational simulators capable of handling engineering problems. As a result, developing mathematical models is a valuable means and practical tool for estimating solar still productivity and evaluating system performance. These created mathematical models, embodied in ML models, play a primary role in optimizing and simulating solar still productivity, resulting in efficient and economical design [22].

Moreover, it would be easy to build strategies to harness this new purified water source after the solar desalination process to service numerous industrial sectors by predicting and knowing solar still productivity. In contrast, the classical mathematical models used to implement these computations remain complex. These algorithms necessitate a lengthy computational time, a large processing capacity, and the solution of multiple complicated equations [23].

In this work, we developed a framework to establish a mathematical model that predicts how productive single-slope solar still would be using artificial neural networks (ANNs). In addition, the performance of the developed random initialization ANN model was evaluated through a statistical comparison of the results of solar still productivity derived from the constructed ANN model (predicted results) and previous experimental findings (observed results). Regarding the suitability for forecasting solar still productivity, transfer learning was applied to the random

initialized ANN model. Fine-tuning via transfer learning allowed us to obtain the target ANN model by unfreezing the input layer nodes and freezing at least one hidden layer neuron per the learned weights. The learned networks are then transferred from the built pre-trained ANN model (novel solar still source model) to the random initialization ANN model for the conventional solar still model, ultimately yielding the target ANN model. The generated ANN latter model was then compared to the pre-trained ANN and multiple linear regression (MLR) models. Lastly, the ANN target model was contrasted with random initialization ANN and MLR models to compare models' accuracies in predicting water productivity in single-basin solar still. This research paper highlights multiple novelty and significant contributions:

- The study utilizes transfer learning to build accurate predictive ANN models for freshwater productivity in solar stills, which is a unique application of solar desalination systems. Employing a pre-trained source ANN model improves prediction accuracy.
- The study introduces a unique target ANN model with a specific 5-64-64-1 architecture. The model has been enhanced through transfer learning and is designed to improve the accuracy of water productivity predictions in solar stills, distinguishing it from conventional ANN models.
- The established target ANN model performs well in predicting purified water outputs, outperforming the random initialization ANN and MLR models developed in the study. This was evaluated using improved statistical performance metrics.
- The study demonstrates the practical effectiveness of the ANN model by confirming its generalization and capability to accurately predict distillate water production under various conditions.
- This research article provides valuable insights into solar desalination systems and suggests cost-effective, practical solutions for addressing significant industrial challenges. These discoveries form the basis for future advancements in solar desalination technologies and pave the way for potential practical applications, instilling a sense of hope in the reader about the future of the field.

The subsequent sections of this paper were organized as follows. Section 2 presents the essential literature review and discusses related work to identify the research gap. Section 3 introduces a supervised learning-based ANN methodology and framework for developing a pre-trained ANN model. Additionally, we investigate the application of transfer learning on the random initialization ANN model to generate the optimal target ANN model, describing the main building blocks, including network architecture, and shedding light on their criteria for evaluation. Moving on to section 4, we extensively discuss the implementation details for our proposed study with a performance analysis of the RI ANN compared to MLR models and an exploration of hyperparameter optimizations. Finally, this paper culminates with a discussion of closing remarks in section 5, emphasizing the potential prospects that lie ahead.

## **2. Literature Review**

Knowledge-based machine learning (ML) has emerged as a powerful tool in solar-desalination systems, offering rapid and accurate predictions and cost-effective designs by leveraging existing frameworks that have been proposed in the past. Using and implementing artificial neural network models in the solar still desalination process could attain desirable results that are difficult to obtain with conventional design equations. This might be attributed to the revolution in computer technology and innovative mathematical modeling techniques [24]. In a recent study, Murugan et al. [25] aimed to improve freshwater production using solar stills, focusing on performance in

diverse climatic conditions. They utilized machine learning models to enhance yield predictions, addressing the limitations of traditional forecasting methods. The research involved evaluating several machine learning models, including linear regression, decision trees, random forest, support vector machines, and multilayer perceptron, employing metrics such as Mean Absolute Error (MAE), cross-validation, and parameter tuning techniques. The decision tree model emerged as the most effective predictor, achieving MAEs of 5.43 and 5.74 through random and grid search methods, respectively. This approach significantly enhances the accuracy of solar still output predictions, thereby contributing to optimizing solar still designs and solar energy conversion mechanisms. The study concluded that machine learning, particularly the decision trees model, offers a robust solution for forecasting distillate outputs in single basin single slope solar stills, enabling more efficient and sustainable water production. Future research avenues involve refining the model, integrating additional data, and further exploring advanced techniques such as deep learning to improve predictive accuracy and performance across diverse solar still designs.

In a study by S. Nazari et al. [26], accurate models for predicting the energy efficiency, exergy efficiency, and water productivity of a single-slope solar still were developed using ANN and an ANN optimized with the Imperialist Competition Algorithm (ICA). These models were trained using empirical data, with various environmental and operational parameters such as time, ambient temperature, solar radiation, glass temperature, basin temperature, and water temperature considered. The neural network with five hidden neurons demonstrated the best performance. Implementing the ICA significantly improved ANN performance, reducing Mean Absolute Error (MAE) by 54.30%, 40.11%, and 53.35% for water productivity, energy efficiency, and exergy efficiency, respectively. Furthermore, the ANN-ICA achieved root mean square error (RMSE) values of 15.77, 1.37, and 0.29 for predicting these metrics. The results highlighted that the ICA-enhanced ANN provided superior predictions compared to traditional ANN models. This study emphasizes the potential of integrating neural networks with optimization algorithms to enhance predictive accuracy in solar still performance metrics.

Another advanced research conducted by H. A. Maddah [27], established precise supervised predictive machine learning models based on experimental results from the literature to estimate double slope solar still productivity and some key thermodynamic properties. Training collected datasets were created using previous inputs and outputs variables from different designed active or passive solar stills ( $\eta \sim 42\%$ ) for treating wastewater or brackish water with a TDS content of 45%. The difference in water and glass temperatures ( $T_w - T_g$ ) was shown to be semi-proportionally correlated with the water distillate with the most minor statistical errors. According to the proposed relationship, the highest distillate would be produced at 14:00 with an increase in both basin temperature ( $T_B$ ) and ( $T_w - T_g$ ). In double slope still designs, the regression models (FGSVM, SEGPR, and EBoT) had the lowest attained root mean square error (RMSE) of 138, suggesting their reliability in precisely forecasting the distillate quantities. The remarkable accuracy of the SEGPR-trained model ( $R^2 = 1$ ) and very low RMSE  $< 8$  demonstrated the model's capability to predict performance in comparable solar-desalination systems. Nevertheless, it was revealed that the stepwise linear regression (SLR) performed better predicting the  $T_B$  pattern than the FGSVM did predict the ( $T_w - T_g$ ) against the water distillate. For more precise predictions, H. A. Maddah [27] claimed that built models could be further optimized by utilizing inputs such as feed water flow rate, environmental conditions, and insulator properties. Such theoretical models would point

us in the right direction for tuning the parameters linked to the convective, evaporative, and radiative coefficients for boosting the distillate-water outputs in double-slope solar stills.

### **3. Methodology and Framework**

In this work, we developed a framework to establish a mathematical model that predicts how productive single-slope solar still would be using artificial neural networks (ANNs). Moreover, the performance of the developed RI ANN model was evaluated through statistical metrics and comparison of the solar still productivity results derived from the constructed ANN model (predicted results) and previous experimental findings (observed results). The transfer learning method was employed for the RI ANN model based on its suitability for forecasting water outputs. Fine-tuning via transfer learning allowed us to generate the target ANN model by unfreezing the input and first hidden layers neurons and freezing at least one hidden layer neuron inherent to the learned weights. Therefore, to get the target ANN model with a more accurate prediction of treated water, the optimal learned networks (intelligence) were obtained, frozen, and transferred from the developed pre-trained (source) ANN model to the RI ANN model. The source model was established from the novel solar still system datasets. Finally, the target ANN model was contrasted with RI ANN and MLR models to compare the models' accuracies in predicting water productivity in a single-basin solar still system.

#### *3.1 Developing source ANN model:*

The source datasets were collected from a novel hybrid solar still system designed and simulated using MATLAB/SIMULINK. The simulation software was used to accurately model system behavior with the aid of a graphical user interface (GUI) modeling tool [28]. This GUI application integrates a visual library into the Simulink browser, enhancing detailed and dynamic modeling capabilities. The initial datasets comprised operational variables of the solar still such as feed flow rate, meteorological conditions (sourced from NASA POWER) [29], water temperatures, salinity, and water distillates measured hourly from 9:00 to 18:00. These parameters were selected due to their critical impact on the efficiency and productivity of the solar still. The original data collected over several simulation runs was expanded by 7.3-fold using the MATLAB curve fitting toolbox through the signal builder tool to increase the dataset size from 100 to 730 data points for input parameters accompanied by distillate water output. This augmentation process involved adding controlled noise to the data to introduce non-linearity and prevent overfitting, which is essential for enhancing machine learning model training.

For the development of the ANN base model, the expanded dataset was divided into training (70%), testing (20%), and validation (10%) sets. Several iterations of random selection and partitioning were performed in multiple rounds to ensure the generalizability and robustness of the results. This method was iterated ten times to address potential biases from a single partitioning. Following each iteration of partitioning, we conducted separate training and evaluation of the models. The results obtained from these repeated iterations (runs) were averaged to evaluate the model's effectiveness comprehensively. This methodology facilitates attaining more reliable and comprehensive outcomes by considering the variations in the data splits. In the training phase, the ANN is trained to predict an output, specifically distilled water in the novel solar still system.

During the testing phase, the ANN model is used to forecast a single output value ( $P_{std}$ ) to examine the validity and reliability of the machine learning models and to determine whether the training

should stop or continue. The ANN is validated during the validation phase to ensure its performance on new cases [22]. A supervised machine learning technique, such as the ANN model, was chosen and compared with a multiple linear regression (MLR) model on labeled datasets in the training, testing, and validation phases to compare and select the best modeling stage for superior prediction outcomes. ANN and MLR models were constructed using a hold-out strategy of 32 random states, controlling the shuffling process in the train-test-split. The source ANN model was developed using Collab notebook/Python software and the widely used feed-forward/back-propagation algorithm [30].

Solar stills undergo energy gains and losses influenced by selected weather input variables [23]. The recorded daily operational feed water volume was included as an input into the ANN model. The feed water flow rate was considered in the neural network analysis as it is related to the constant factor of the basin area. When operating solar stills, users will likely utilize the feed water volume to determine the water quantity needed for flushing and operation [23]. The chosen input variables—daily solar radiation ( $I_s$ ), ambient temperature ( $T_{amb}$ ), wind speed ( $V_w$ ), relative humidity (RH), and saline water feed flow rate ( $P_{st}$ )—were identified as primary influencers of solar still performance. The dependent variable, water productivity ( $P_{std}$ ), was then correlated with these inputs that emphasize our machine learning analysis.

It is worth noting that the input parameters for our model were chosen based on their substantial influence on the performance of solar stills and their capability to forecast freshwater productivity precisely. Multiple factors justified the omission of the inlet water temperature. Initially, the inlet temperature of the basin was consistently maintained, leading to minimal fluctuations in its level. The initial investigation found that this parameter did not significantly improve the model's ability to predict water productivity accurately. Including it would have extended the model's complexity without yielding substantial benefits. Therefore, our attention was directed toward the five most influential operational variables to optimize the model and enhance its efficiency and interpretability.

Critical decisions in the model training process included selecting 92 iterations based on a batch size of 8 samples, determined to be optimal through hyperparameter tuning. The Adam optimizer's regularization strength and learning rate were set at 0.001 and 0.0003, respectively, to ensure robust training while minimizing overfitting. Early stopping with a patience value of 20 was employed to halt training if the validation loss did not improve, thus preserving model generalization. A single epoch occurs when an entire dataset is only once passed through the neural network forward and backward. This detailed approach to dataset preparation and model training highlights the systematic methodology for developing a reliable and accurate ANN model for predicting solar still performance. Table 1 demonstrates the collected input and output parameters data sources for both solar stills configuration systems.

Table 1: Inputs and outputs parameters data sources

<b>Solar Still Design Configurations</b>	<b>Input Parameters</b>	<b>Data Source</b>	<b>Output Parameters</b>	<b>Data Source</b>
Novel solar still system (Pretrained model)	Operational Variable: Saline water feed flowrate, L/day	Literature/ experimental datasets [28].		

Conventional solar still system (Random Initialization model)	Meteorological Conditions: Solar radiation, MJ/m <sup>2</sup> .day Ambient temperature, °C Wind speed, m/s Relative humidity, %	NASA Power [29].	Solar still productivity Water distillate (L/day)	Literature/experimental datasets [23].
--	---	------------------	--	--

### 3.2 Criteria of assessment for random initialization ANN model:

For evaluating the long-term performance of the conventional single-slope solar still, 365 numerical datasets were manually extracted from preceding experimental results found in the literature and were conducted by Santos [23]. The compiled datasets initially involved direct measurements of feed water volumes and daily average distillate productions between February 2006 and August 2007 using a rectangular basin still with an area of 0.976 m<sup>2</sup> [31]. They then were tabulated with corresponding daily recorded weather data as input variables, including solar incidents retrieved from NASA power webs [29]. The random initialization ANN model was built with feed water volumes and daily average weather data as inputs and the total freshwater production as the output variable ( $P_{std}$ , L/m<sup>2</sup>.day).

The earlier study [23] intended to predict single-basin solar still distillate production comprising some water physical properties of the feed and distillate. For instance, density and specific heat capacity are functions of feed salinity and initial water temperature since these factors were critical for assessing the water safety criteria for safe human consumption [32].

Likewise, as applied to the source model, the trained models used to forecast distillate water in the conventional solar still were employed in the testing analysis for both ANN and MLR models. The training datasets (70% of the curated data) comprised six matrices of [256 x 1], with each matrix representing an input parameter or water productivity as an output. This defined matrix size is also comparable to that a one [256 x 6] matrix was curated considering the inputs as mentioned earlier and the distillate water output. The same inputs were used in the testing datasets, extracted as 20% of the curated data to predict the known freshwater productivity for verifying trained model accuracy. Lastly, the 10% curated data were utilized to check the performance of new datasets and get a model that can generalize better on unseen or future datasets.

It is typical to evaluate the predictive performance for the developed ANN and MLR model using a variety of statistical metrics, such as coefficient of determination ( $R^2$ ), root mean square error (RMSE), mean square error (MSE), mean absolute error (MAE), the overall index of model performance (OI), and residual, which were calculated using their mathematical computations in eqns. (1-7), respectively [33].

$$RMSE = \sqrt{\frac{\sum_{i=0}^n (x_{o,i} - x_{p,i})^2}{n}} \quad (1)$$

$$R^2 = \frac{(\sum_{i=0}^n (x_{o,i} - \bar{x}_o)(x_{p,i} - \bar{x}_p))^2}{\sum_{i=0}^n (x_{o,i} - \bar{x}_o)^2 \times \sum_{i=0}^n (x_{p,i} - \bar{x}_p)^2} \quad (2)$$

$$MAE = \frac{\sum_{i=0}^n |x_{o,i} - x_{p,i}|}{n} \quad (3)$$

$$ME = 1 - \frac{\sum_{i=0}^n (x_{o,i} - x_{p,i})^2}{\sum_{i=0}^n (x_{o,i} - \bar{x}_o)^2} \quad (4)$$

$$OI = \frac{1}{2} \left( 1 - \left( \frac{RMSE}{x_{max} - x_{min}} \right) + ME \right) \quad (5)$$

$$CRM = \frac{(\sum_{i=0}^n x_{p,i} - \sum_{i=0}^n x_{o,i})}{\sum_{i=0}^n x_{o,i}} \quad (6)$$

$$Residual = x_{o,i} - x_{p,i} \quad (7)$$

Knowing that the observed value is denoted by  $x_{o,i}$ .  $x_{p,i}$  is symbolized as the predicted value by the ML models.  $\bar{x}_o$  is the averaged observed or mathematically obtained values;  $\bar{x}_p$  refers to the predicted values or theoretically estimated from averaging; the minimum and maximum observed values denoted by  $x_{min}$  and  $x_{max}$  respectively, and  $n$  indicates the number of observations or dataset size. The lower  $RMSE$  and  $MAE$  values determine the more accurate prediction results, and their values range from 0 to  $\infty$ . On the other hand, the higher  $R^2$  value demonstrates the more remarkable similarities between the trendlines of both observed and predicted samples (an identical pattern is reached when  $R^2 = 1$ ) [34].

A perfect fit between observed and forecasted results is generated when  $ME$  value is 1. In contrast, a zero value implies that a mean value would have produced the same level of accuracy [35] [36]. The latter metric could produce a negative value  $-\infty$  if the mean observed value  $\bar{x}_o$  is a better predictor than the simulated value, which indicates unacceptable performance. The  $OI$  value of 1 expresses an optimal fit between observed and predicted outputs. The  $CRM$  values are near  $\pm 1$ . Better model accuracy is achieved when the  $CRM$  value is close to zero. For ideal data modeling,  $RMSE$ ,  $MAE$  and  $CRM$  measures, which were obtained from the different trained models, need to be closer to zero, but  $R^2$ ,  $ME$  and  $OI$  values should converge to 1 as much as possible [37].

Fig. 1 demonstrates a flow chart illustrating the selection and development of the optimal ANN source and target models for precisely predicting distillate water productivity,  $P_{std}$ . Based on the statistical metrics presented in Table 2, the optimum ANN architecture was chosen through a trial-and-error technique by determining whether the  $R^2$  was greater than 0.75 and predicting the distillate water ( $P_{std}$ ) from the testing datasets. Selected models that only satisfied the previous criteria were retained for further analysis to compare their predicted response patterns and residuals (from the testing datasets) to the actual experimental results (from the training datasets).

The ideal chosen architecture of the source ANN model represented in trained network weights was employed to be repurposed (reused) in the random initialization ANN model and trained on new datasets and functions based on minimum measures of errors. Hence, as illustrated in Fig. 2, it was attempted to fine-tune the latter model by training only the first hidden layer, leaving the other frozen, and developing a predictive mathematical algorithm to obtain the target model from patterns that had been trained.

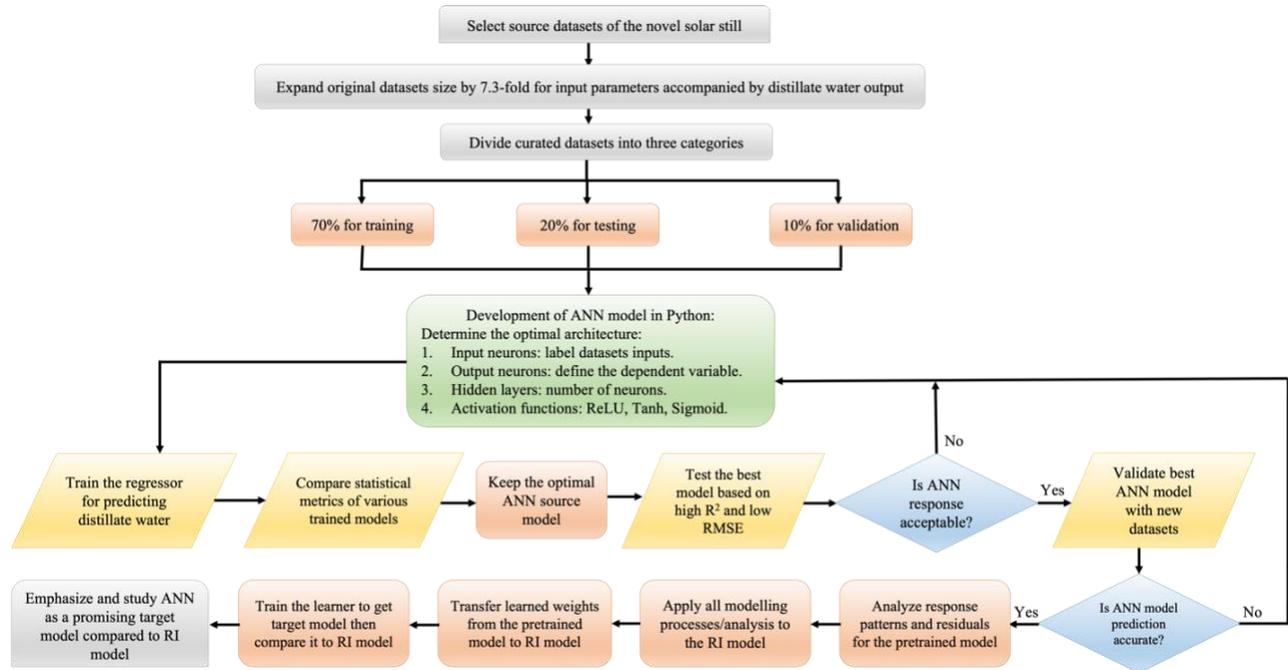


Figure 1: Flowchart for selecting and developing optimal pre-trained ANN and target models for accurately predicting distillate water productivity.

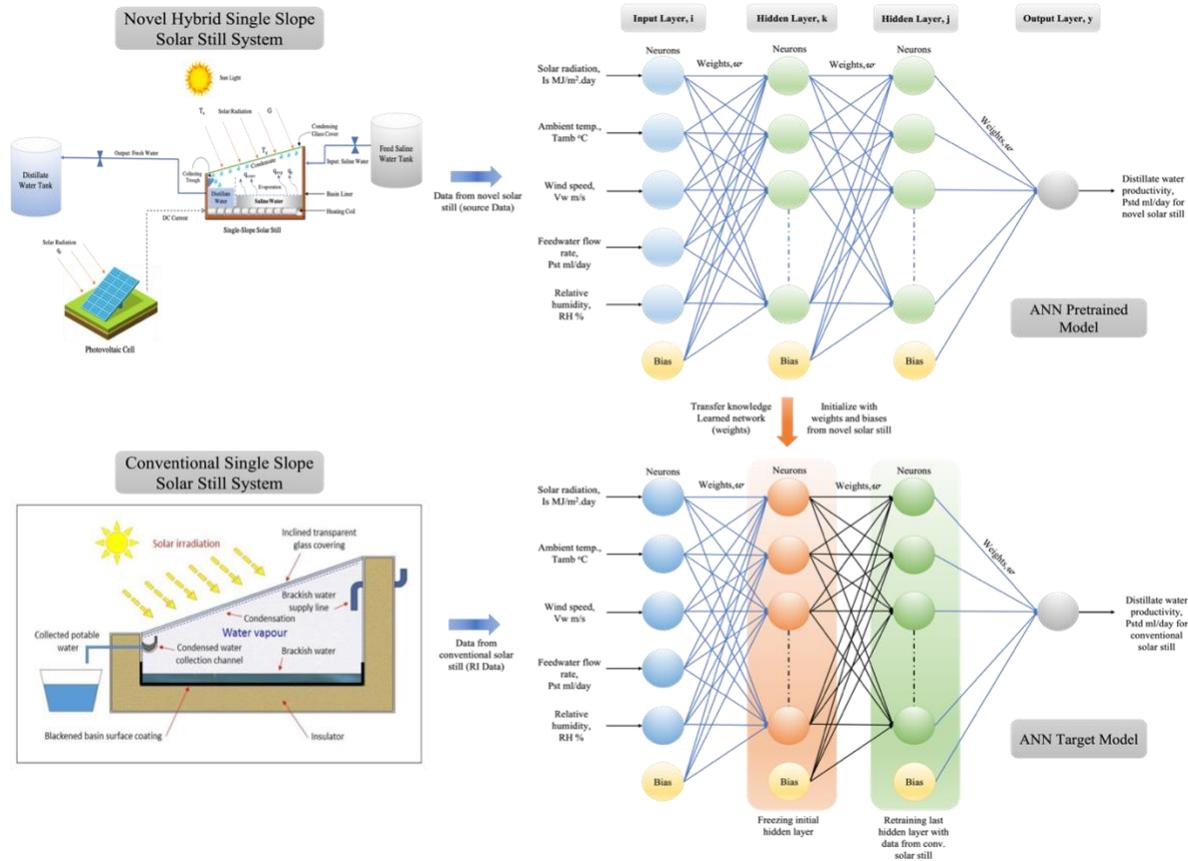


Figure 2: Schematic diagram for applying transfer learning technique on the conventional solar still system to accurately predict distillate water outputs  $P_{std}$ ,  $L/m^2 \cdot day$ .

These algorithms compute solar still freshwater productivity based on experimental findings from a previous study using ANN models [23]. This machine learning technique is identified as a transfer learning method, which was implemented in the conventional solar still base model to generate the final target model. random initialization ANN model was created as a first step in the conventional single slope solar still without applying the transfer learning method. Then, it was compared to the developed optimal target ANN model. In this research, we focused on developing an ANN regression model to forecast continuous numerical variables associated with water productivity in solar stills. Unlike binary classification problems that generally utilize hard-limit transfer functions [0,1], our ANN regression model employs non-linear transfer functions such as ReLU to capture the complex relationships within continuous data. By integrating these non-linear transfer functions into our feedforward neural network, we can more effectively model and predict continuous values than using linear or hard-limit functions.

It is important to note that graphical and statistical comparisons were used to evaluate the performance of all developed ANN and MLR models. Finally, it was aimed to study and emphasize whether the target ANN model generated by learned (frozen) weights would provide better prediction than the random initialization model based on their accuracies and minimum statistical errors. As illustrated in Fig. 3, repurposing the pre-trained (source) model was applied to the random initialization model to predict target output data (distillates) in the future using the trained

network model (transferred knowledge) and to evaluate overall long-term system performance in single slope solar still.

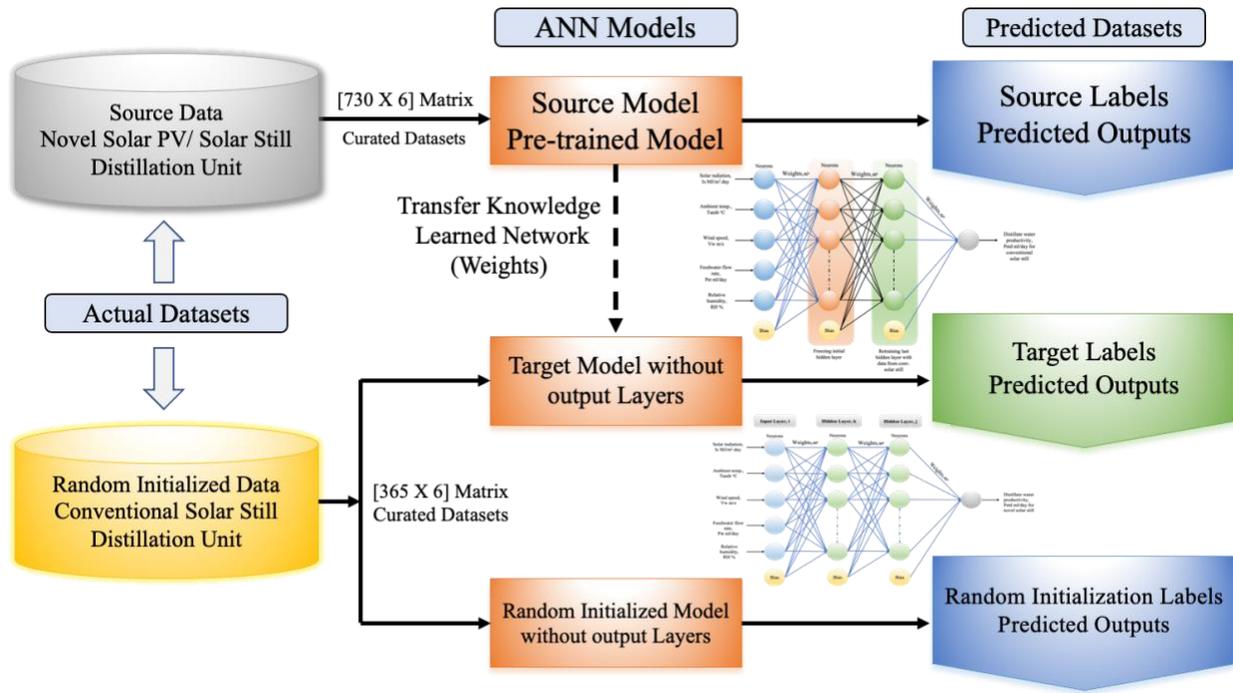


Figure 3: A schematic block flow diagram for applying the transfer learning method on the conventional solar still system known as the random initialization (RI) ANN model to predict distillate water ( $P_{std}$ ,  $L/m^2.day$ ) outputs more accurately. The optimal source (pre-trained) ANN model was developed to produce the best-learned weights, freeze, and transfer them to the RI ANN network to generate further the target ANN model with the minimum statistical errors.

Table 2: Presenting the RI ANN model methodology in a pseudocode form.

S/N	Step	Pseudocode Explanation
1	<b>Input Datasets</b> (70% training, 20% testing, 10% validation)	X: Input features dataset (2D array).
		Y: Distillate water variable dataset (1D array).
		X_train, X_test, X_val: Split X into training, testing, and validation datasets (2D arrays).
		y_train, y_test, y_val: Split y as an output variable into training, testing, and validation datasets (1D arrays).
2	<b>Data Normalization</b>	Import necessary libraries (numpy, sklearn.model).
		Define MinMaxScaler objects for input and target variables.
		Fit MinMaxScaler on training datasets.
		Normalize the training, validation, and test data using the fitted scalers (Scaler Fit. transform).
3	<b>ANN Model Creation</b>	Import necessary libraries (numpy, keras, sklearn.model and tensorflow).
		Create a sequential neural network model.

	<b>First Hidden Layer</b>	Create the first hidden layer with 16 units, input dimension 5, using ReLU activation function, and normal weight initialization.
	<b>Second Hidden Layer</b>	Create the second hidden layer with 32 units, using ReLU activation function, and normal weight initialization.
	<b>Output Layer</b>	Create the output layer with 1 unit, representing a single fully connected node, using normal weight initialization.
	<b>Compile the model</b>	Compile the RI ANN model using mean squared error as the loss function and Adam optimizer with a learning rate of $3e-4$ .
4	<b>Fitting the ANN to the Training Dataset</b>	Train the model on the training data ( $X_{train}$ , $y_{train}$ ) with a batch size of 20, 500 epochs, and validation data ( $X_{val}$ , $y_{val}$ ).
5	<b>Early Stopping</b>	Apply early stopping during training to monitor the validation loss and restore the best weights when patience exceeds 20 epochs.
6	<b>Hyperparameter Optimization</b>	Define a set of hyperparameter values for tuning (see Table 3).
		Use techniques such as grid search or random search to find the optimal hyperparameters with a maximum evaluation of 100 trials.
		Train the model using the optimal hyperparameter values.
7	<b>Statistical Metrics</b>	Train the model on the training datasets.
		Evaluate the model on the validation and test datasets.
		Calculate statistical metrics such as Root Mean Squared Error (RMSE), R-squared, OI, etc.
		Print or store the metrics results for analysis and comparison.

#### 4. Results and Discussion

This chapter explains the development of an optimal pre-trained (source) ANN model architecture to get the best-learned network (weights). The ideal trained weights will then be repurposed (reused) in the random initialization (RI) ANN model to generate the target ANN model and predict daily water outputs. This part of the study will present a statistical comparison and performance analysis between the created random initialization ANN, the target ANN model, and the MLR model.

##### 4.1 Optimization of pre-trained ANN model architecture:

The ANN and MLR regression models used to estimate distillate water outputs for both solar still configuration systems were covered in this study. The statistics of the pre-trained ANN model's performance with various hidden layer neuron/node numbers and activation functions are displayed in Table 3. Based on the network hyper-parameters / gradient descent optimization algorithm shown in Table 3, the optimum ANN architecture, 5-64-64-1, for the novel solar still was chosen through a rigorous optimization process using a trial-and-error method with a maximum evaluation of 100 trials aiming to balance between the model's complexity and predictive accuracy. It was observed that the best neurons for the source ANN model's hidden layer were selected through trial and error (examined from 16 to 128 neurons) with a batch size ranging from 8 to 64. Each hidden layer's activation function varied and switched between sigmoid, ReLU, and Tanh. Despite the seemingly large number of parameters, our approach validated the model's generalization capabilities. Moreover, the empirical results confirmed that the chosen architecture

5-64-64-1 was optimal for the given data that minimized validation loss and enhanced predictive accuracy, and the use of techniques such as regularization and early stopping effectively mitigated overfitting. Although a higher number of neurons, this configuration was the most effective at capturing our dataset's complexities and non-linear relationships. Additionally, the number of neurons in the hidden layers plays a crucial role in the model's capability to effectively learn from and generalize the data. While fewer neurons may simplify the model and reduce computational demands, they can also limit the model's ability to generalize and represent the underlying data patterns precisely. In our scenario, the architecture with 64 neurons in each hidden layer balanced model complexity and predictive performance, ensuring robust and accurate predictions.

Certain literature presents empirical formulas to choose the number of neurons in neural network models. However, our study examined the number of neurons, ranging from 16 to 128, for forecasting distillate water productivity in solar stills using an ANN model. Instead of solely relying on existing empirical methods and guidelines, empirical assessment was carried out to optimize the architecture of the ANN model. This decision has been made after considering various factors. While empirical formulas provide a useful framework, they may not always generate the best configuration for a specific dataset. Fine-tuning the architecture of the ANN model through empirical testing can ensure that it fits the dataset's characteristics, leading to improved performance. The range of neurons from 16 to 128 has been selected to balance computational efficiency and model complexity, making it a widely preferred option. Fine-tuning the ANN within this range notably enhanced prediction accuracy and reliability as well as boosted learning and convergence efficiency.

Several gradient descent methods optimizers, such as RMSprop, Adagrad, Adam, and SGD, have also been employed and tuned in the optimization process. Regularization strength, namely lambda, was also assessed between 0.001 and 0.1 to calibrate ANN models and to reduce the adjusted loss function; thus, avoid overfitting or underfitting (see Table 3). Trials were initiated using 16 neurons in the input hidden layer as the starting point for optimizing the ANN base model architecture. It was noted that fixing the number of neurons to 64 resulted in a clear and remarkable improvement in the ANN model performance, particularly for the ReLU function. This transfer function provided the best network performance since  $R^2$ , RMSE, MAE, ME, and OI values for the testing datasets were 0.996, 0.098, 0.074, 0.993, and 0.991, respectively.

Figure 4 demonstrates the ReLU activation function graph and compares the denoising results of ReLU and PReLU activation functions in the pre-trained ANN model. It was observed that both transfer functions were very close in shape in terms of denoising capabilities, and the mean squared error, MSE values (ReLU = 0.0111 and PReLU = 0.0102) were similar, with the only difference being a slight slope for negative values in the PReLU curve. This behavior suggests that both activation functions performed similarly in fitting the training data and generalizing it to the validation data.

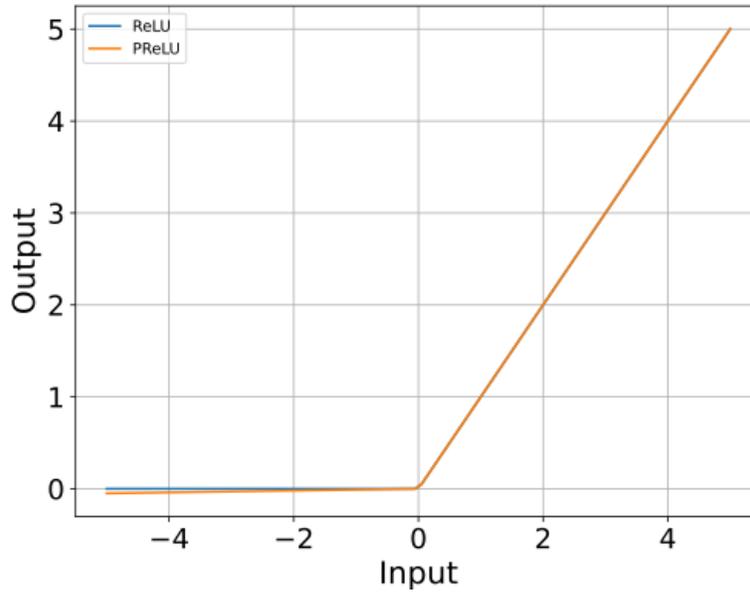


Figure 4: ReLU and PReLU activation functions graph.

The optimization strategy primarily motivates the decision to utilize the ReLU activation function in our ANN model during training. By normalizing the input data to a range from -1 to 1, the ReLU function is optimized, leading to improved performance and model convergence speed. While the sigmoid transfer function is constrained to the interval  $[0,1]$ , its efficacy was evaluated alongside other functions, including ReLU and Tanh, during the hyper-parameter optimization method. This examination helped us identify our regression task's optimal activation function. Ultimately, the ReLU function emerged as the most suitable choice for the dense layers of our ANN model due to its outstanding performance observed during training. It is essential to understand that in hyperparameter optimization, the output range of the activation functions does not necessarily have to match the normalization range of the input data. Irrespective of the activation functions being evaluated, it is advantageous to normalize the data in the training phase to a range from -1 to 1.

In many practical applications, this slight difference may be insignificant. Moreover, ReLU is more computationally efficient than PReLU because it does not require additional learnable parameters, making it advantageous for resource-constrained environments. Its simplicity and robustness contribute to ReLU's viability as a reasonable option. It can be emphasized that the choice of activation is merely one component of the entire modeling pipeline, in which other factors, such as network architecture and hyperparameter tuning, also contribute to the overall performance. As shown in Table 3, the produced hyper-parameters optimization trials also revealed that having 16, 32, and 128 neurons with 4, 4, and 5 hidden layers, respectively, would harm ANN model performance with a more excellent loss value than one. Also, the ReLU function was better than Tanh during most trials, having the smallest lambda (weight decay) value of 0.001 in the prediction of  $P_{std}$ . Another hyper-parameter considered in the training process is the number of epochs, which specifies how often the learning algorithm will run through the entire training dataset. In addition, the learning rate values for the Adam optimizer and momentum were adjusted to maintain a constant value of 0.0003 and 0.99, respectively.

Having too many epochs can result in an overfit model while having too few can lead to an underfitting of the training dataset. Consequently, an early stopping strategy, supported by a

callback function, was considered while building the model that enables us to define an arbitrarily large number of training epochs (e.g., 500 epochs) and terminate training whenever the model performance on a holdout validation datasets halt improving. This technique can be accomplished by passing the validation dataset to the fit function during model training and allowing us to identify the performance indicator to monitor to stop training. The loss on the validation dataset will then be made available under the term "validation loss" with a minimum value. Stopping training at the first sign of no further improvement is usually not the most excellent option. The early stopping event while training the datasets may occur because the model may reach a plateau of little progress or perhaps slightly deteriorate before improving significantly. We may adjust for this possible behavior by adding a delay to the trigger based on how many epochs we want to observe no improvement. Adding delay to the trigger while fitting the ANN to the training set can be achieved by specifying the "patience" parameter with a value of 20 [38].

Table 3: Hyperparameters optimization for source (pre-trained) ANN model in sequence from minimum to maximum loss.

<b>Trial number</b>	<b>Hidden layer size</b>	<b>No. hidden layers</b>	<b>Activation function</b>	<b>Batch size</b>	<b>Optimizer</b>	<b>Lambda (Weight decay)</b>	<b>Loss</b>
<b>1</b>	<b>64</b>	<b>2</b>	<b>ReLU</b>	<b>8</b>	<b>Adam</b>	<b>0.001</b>	<b>0.000621</b>
2	32	2	ReLU	8	Adam	0.001	0.00063
3	32	2	ReLU	8	Adam	0.001	0.000723
4	128	2	ReLU	16	Adam	0.001	0.000727
5	32	2	ReLU	16	Adam	0.001	0.000734
...	...	...	...	...	...	...	...
96	32	2	ReLU	64	Adagrad	0.005	0.17612
97	64	5	ReLU	64	Adagrad	0.001	0.255997
98	32	4	ReLU	32	Adagrad	0.05	1.080505
99	16	4	Tanh	32	Adagrad	0.1	1.750207
100	128	5	Tanh	32	Adagrad	0.01	1.939331

Visualizing the performance of any ANN model is a simple approach to making sense of the data coming out of the model and creating a knowledgeable choice about the adjustments that need to be applied to the hyper-parameters that influence the ANN model. It is vital to comprehend how to visualize data using the Matplotlib library and combine it with the ANN model to make an enlightened decision and enhance the ANN model. One of the main challenges of any ANN model is generalizing the model so that it can predict good results for new data and data that has already been trained.

Figure 5 illustrates the average contribution of each input variable to the output within the optimal ANN model architecture. These values demonstrate the comparative significance of each input parameter in the training process of the developed RI ANN model.

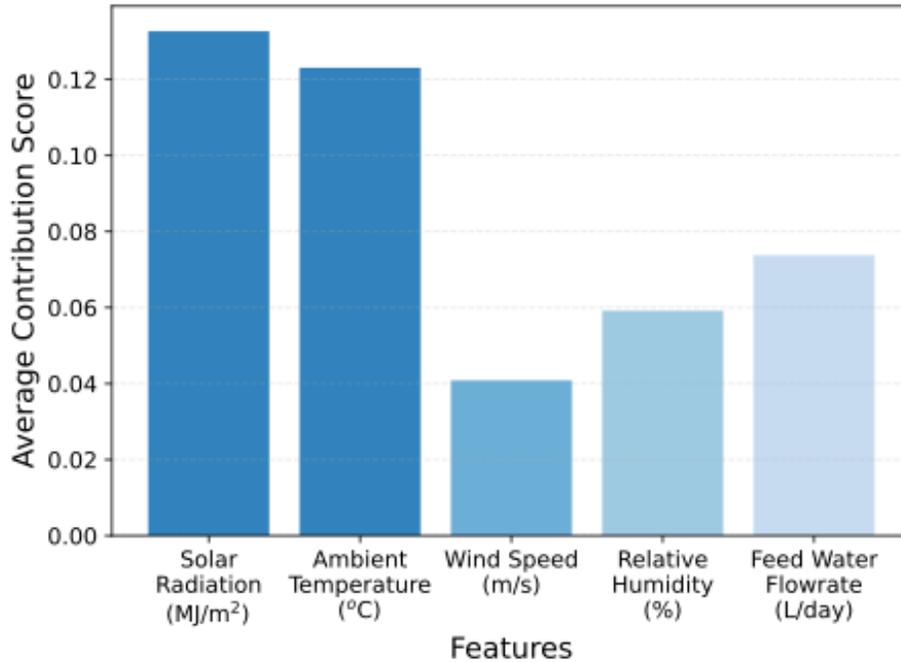


Figure 5: Average contribution values of input node variables on distillate water outputs  $P_{std}$ , L/m<sup>2</sup>.day.

Moreover, this analysis of contributions can provide insights into the relative importance of variables by identifying those with the highest contribution values relative to other inputs. Based on the conducted analyses, it was observed that the variable with the least significant impact is wind speed, with a contribution score of 0.04. On the other hand, relative humidity and feed water flow rate exhibit moderate effects on the predicted distillate water, with contribution scores of 0.061 and 0.075, respectively. The variables that indicate the most significant contributions to the distillate water output are solar radiation (0.135) and ambient temperature (0.122). These parameters are typically the most influential factors impacting the distillation process, primarily due to their involvement in the transfer of energy through radiant and convective mechanisms within the solar still system and their role as energy sources. The findings of this study are consistent with prior research [32,33], which indicated that solar radiation and ambient temperature significantly affect the quality of treated water.

Visualizing the validation loss vs. training loss over the number of epochs (385, 235) for both the pre-trained and RI models, respectively, might help to decide whether the model has been appropriately trained. An ideal fit, as plotted in Fig. 6, or a model that does neither overfit nor underfit, is specified when both the validation and training losses fall and stabilize at specific points. This approach is critical to avoid overfitting the ANN model and underfitting it to the point that it begins memorizing the training data, thus reducing its predictive accuracy [39].

Based on the learning curve analysis, the ANN model's performance appears balanced and does not exhibit clear signs of high variance, high bias, or overfitting.

1- Loss reduction: Both training and validation losses decrease as the number of epochs increases, indicating that the model is effectively learning from the data and improving its predictions.

2- Minimum plateau: Both training and validation losses reaching a minimum plateau suggest that the model has converged to a point where further training does not significantly reduce the losses. This behavior indicates that the model is not overfitting the training data.

3- Convergence of curves: Both curves meet and become steady as the number of epochs increases, indicating that the model's performance on both the validation and training sets has stabilized. This means the model has found a balance between fitting the training data and generalizing it to unseen data.

4- Absence of high variance or overfitting: Since the curves are steady and meet, there is no evident gap between training and validation losses, which suggests that overfitting is not a significant concern. Overfitting typically indicates a growing gap between training and validation losses as the model overlearns the training data.

5- Absence of underfitting or high bias: There is no indication of high bias either, which would manifest as both training and validation losses remaining high and not decreasing. High bias implies that the ANN model should be more complex to capture the underlying patterns in the datasets.

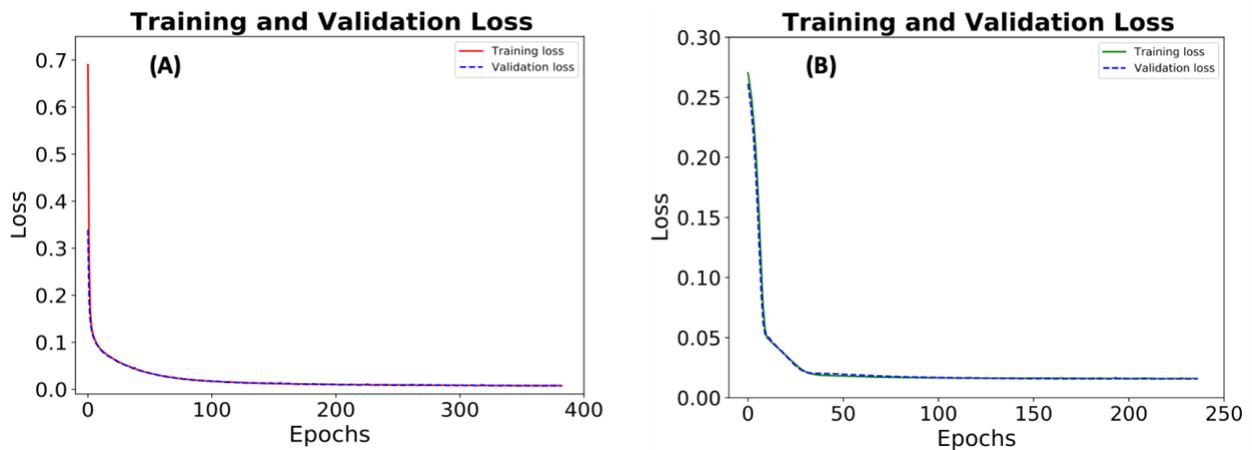


Figure 6: Training loss vs. validation loss over the number of epochs: (A) for the source (pre-trained) ANN model and (B) for the RI ANN model, showing both models are approaching the zero loss and being optimized by the desired number of Epochs.

In light of the optimal water distillates prediction and lowest error (the minimum RMSE, CRM, and MAE; and the largest  $R^2$ , ME, and OI), the most appropriate ANN model was 5-64-64-1 with a ReLU activation function (bolded in Table 3). Fig. 7 schematically illustrates the ideal generated source ANN model architecture 5-64-64-1.

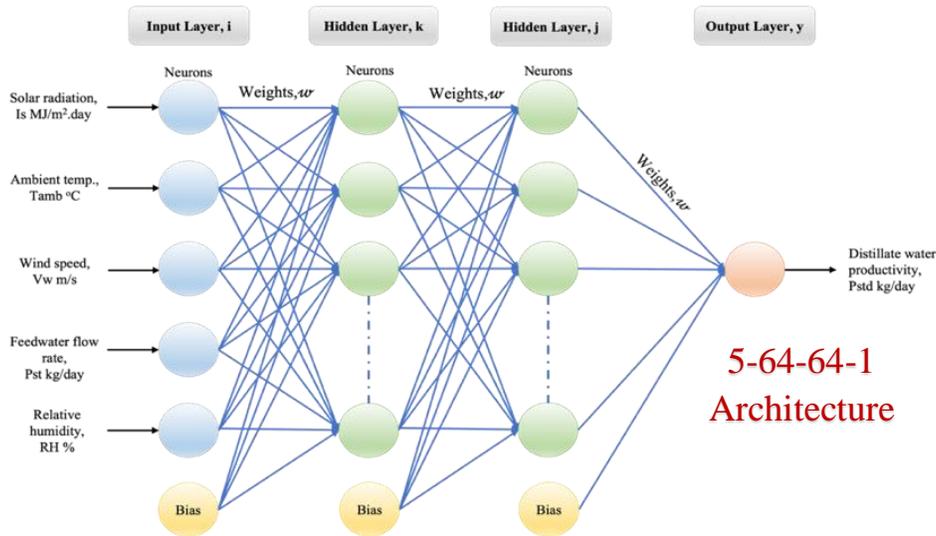


Figure 7: Optimal architecture of the developed source ANN model to predict distillate water.

#### 4.2 Performance analysis of ANN RI and comparison with MLR:

Fig. 8 compares the observed results, ANN, and MLR predictions for the training, testing, and validation datasets. The ANN prediction values for random initialization models were closer to the observed values than the MLR prediction values. The proportionate agreement between predicted and actual values evidences the predictive power and reliability of the built ANN model for predicting distillate water productivity  $P_{std}$ , L/m<sup>2</sup>.day. Fig. 8 (A) demonstrates that the values predicted by the developed random initialization (RI) ANN model fit precisely with the observed values during training. As shown in Table 3, some statistical analyses were employed to evaluate the overall performance of the ANN and MLR models. As revealed in Fig. 8 (B), most of the values predicted by the ANN models were equally and tightly distributed around the 1:1 plotted line. The coefficient of determination,  $R^2$  (0.80) score for the RI ANN model, was partly close to one. The RI ANN model had an RMSE value of less than one (specifically, 0.743 L/m<sup>2</sup>.day), as presented in Table 3. Model efficiency, ME, and an overall index of model performance, OI-related values were 0.80 and 0.832, respectively, which draw near one. The coefficient of residual mass, CRM, was extremely near zero at - 0.0024 L/m<sup>2</sup>.day, but the mean absolute error, MAE, was somewhat diverged from zero at 0.607 L/m<sup>2</sup>.day. Furthermore, Fig. 8 (B) illustrates that multiple points acquired by applying the MLR model during the training stage were positioned above and below the 1:1 slope line. Table 3 also presents the MLR statistical performance metrics.

The created RI ANN model relatively outperformed the MLR's accuracy during the training phase, with an  $R^2$  value of 4% higher. The RMSE and MAE values were 0.5 and 0.43 times higher than their ANN model values. The OI value for the MLR model is reduced by roughly 1.8% compared to the value for the ANN model, and the ME value for the MLR model (0.771) was slightly far from one. The relative errors of the forecasted  $P_{std}$  values using the RI ANN and MLR models are shown in Fig. 8 (C).

Residuals were often in the proximity of  $\pm 2\%$  throughout the training process for the ANN model. In contrast, as stated in Fig. 8 (C), the MLR model's residuals were more significant than those using the RI ANN model, indicating that the developed RI ANN model was adequately accurate.

Moreover, Fig. 8 (D) also indicates how the RI ANN and MLR models were employed to investigate the link between the observed and predicted distillate values during the testing process. As was the case during the training phase, the built ANN model outperformed the MLR model regarding the agreement between observed and predicted values. Most of the freshwater output values for the ANN model, as stated in Fig. 8 (E), followed the 1:1 line during the testing process. As a result, the predicted and observed values were relatively close.

Table 3 shows the statistical parameters  $R^2$ , ME, OI, RMSE, CRM, and MAE used to examine the agreement between the actual and predicted values by the ANN and MLR models using the testing datasets. During the testing process, the  $R^2$ , ME, and OI values were near to one using the established RI ANN model, demonstrating high agreement between the observed and predicted variables using the developed RI ANN model.

On the other hand, the MLR model utilizing the testing datasets had  $R^2$ , ME, and OI values of 2.3, 2, and 3.1 %, respectively, and hence was less accurate than the constructed RI ANN model. It can be seen from Table 3 that the values of RMSE, CRM, and MAE are high for the distillate water predicted by the MLR model and low for the  $P_{std}$  forecasted by the RI ANN model. The created RI ANN model's RMSE (0.721 L/m<sup>2</sup>.day) was more than double that of the MLR model (0.750 L/m<sup>2</sup>.day). In comparison to the CRM value for the MLR model, the CRM value for the RI ANN model was closer to zero. Similarly, the CRM value for the MLR model was one time greater than that of the generated RI ANN model. The MAE value (0.602 L/m<sup>2</sup>.day) for the developed RI ANN model was about 1% lower than those from the MLR model (0.615 L/m<sup>2</sup>.day).

The residuals during the testing process are plotted as illustrated in Fig. 8 (F). With relative errors (90% of the values) largely falling between -1% and +1 %, the ANN RI model was closer to zero and can be utilized to predict  $P_{std}$ . Nevertheless, using the MLR model, the residual values that fall within  $\pm 1\%$  imply less than 90% of the total error values during the testing process.

Regarding the validation datasets, Fig. 8 (G) compares the observed  $P_{std}$  outputs, including ANN RI, forecasted results, and MLR prediction results. During the validation process for the established RI ANN model, Fig. 8 (H) demonstrates that the distillate water typically followed a 1:1 line, suggesting a great match between actual and predicted values. As was the case during the training and testing processes, the constructed RI ANN model outperformed the MLR model based on the agreement between observed and predicted data.  $R^2$  was close to 0.80 according to the ANN RI model prediction result.

Table 3 displays that while the  $R^2$ , ME, and OI values were comparatively close to one, the RMSE and MAE values were relatively low. The CRM value (-0.001) was also very near zero. The generated RI ANN model exhibited a higher  $R^2$  value than the MLR model, which was 4.7% less accurate. About 1.02 times as much RMSE was present in the MLR model (0.837 L/m<sup>2</sup>.day) compared to the ANN RI model.

Furthermore, the developed RI ANN model had a ME value of approximately 5% more accurately than the MLR model. The MLR model's OI value (0.816) was not very close to one compared to the RI ANN model, with a value of 0.841. The MAE value of 0.728 for the MLR model rose by 97.02% from that of the RI ANN model. The validation datasets for both RI ANN and the MLR

model's relative errors of the forecasted  $P_{std}$  values are shown in Fig. 8 (I), with an average relative error of  $\pm 1.5\%$  for the MLR model when utilizing the validation datasets. This figure shows the discrepancies between the findings of the two models. It highlights the slight underestimation of the predicted distillate water. As evidence of the insignificant underestimation in the predicted distillates, the comparable value for the built RI ANN model was lower at 0.5% for the validation datasets.

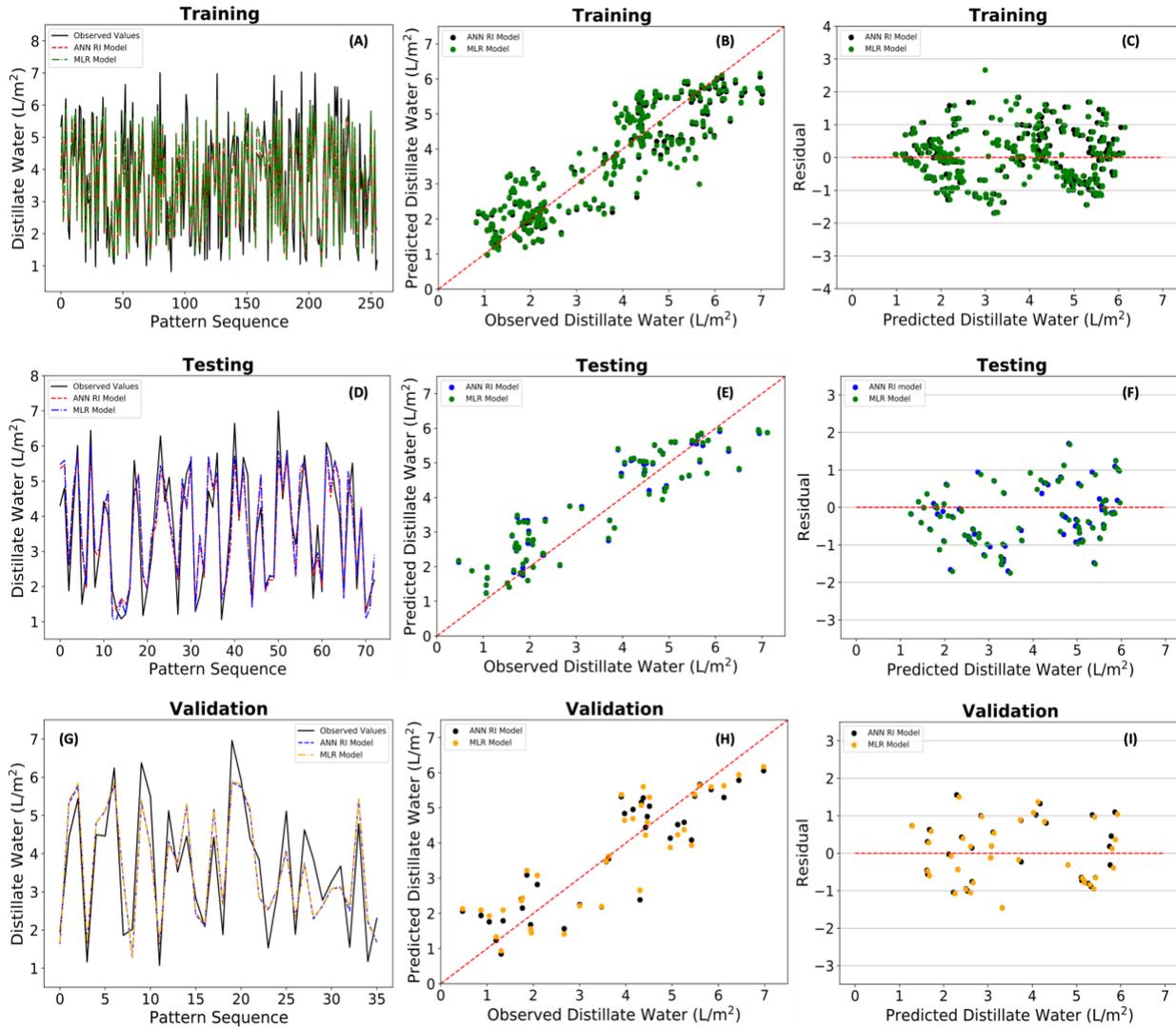


Figure 8: Comparison between observed vs. predicted distillate water values using RI ANN and MLR models during modeling stages with their acquired residuals: (A-C) from training datasets; (D-F) from testing datasets; (G-I) from validation datasets.

It is worth noting that when the plots illustrated in Fig. 9 (A-I) were examined for the training, testing, and validation datasets using the transfer learning method, the predicted distillate water results using the developed target ANN model appeared too close to the observed values. The target ANN model was generated by repurposing the transferred learned weights from the source ANN model. In corresponding to the built target ANN model, the results obtained using the RI ANN and MLR models revealed a more significant deviation from the actual values for all

modeling processes. Based on the preceding, it was possible to conclude that the generated target ANN model was precisely trained and presented consistency in predicting freshwater productivity.

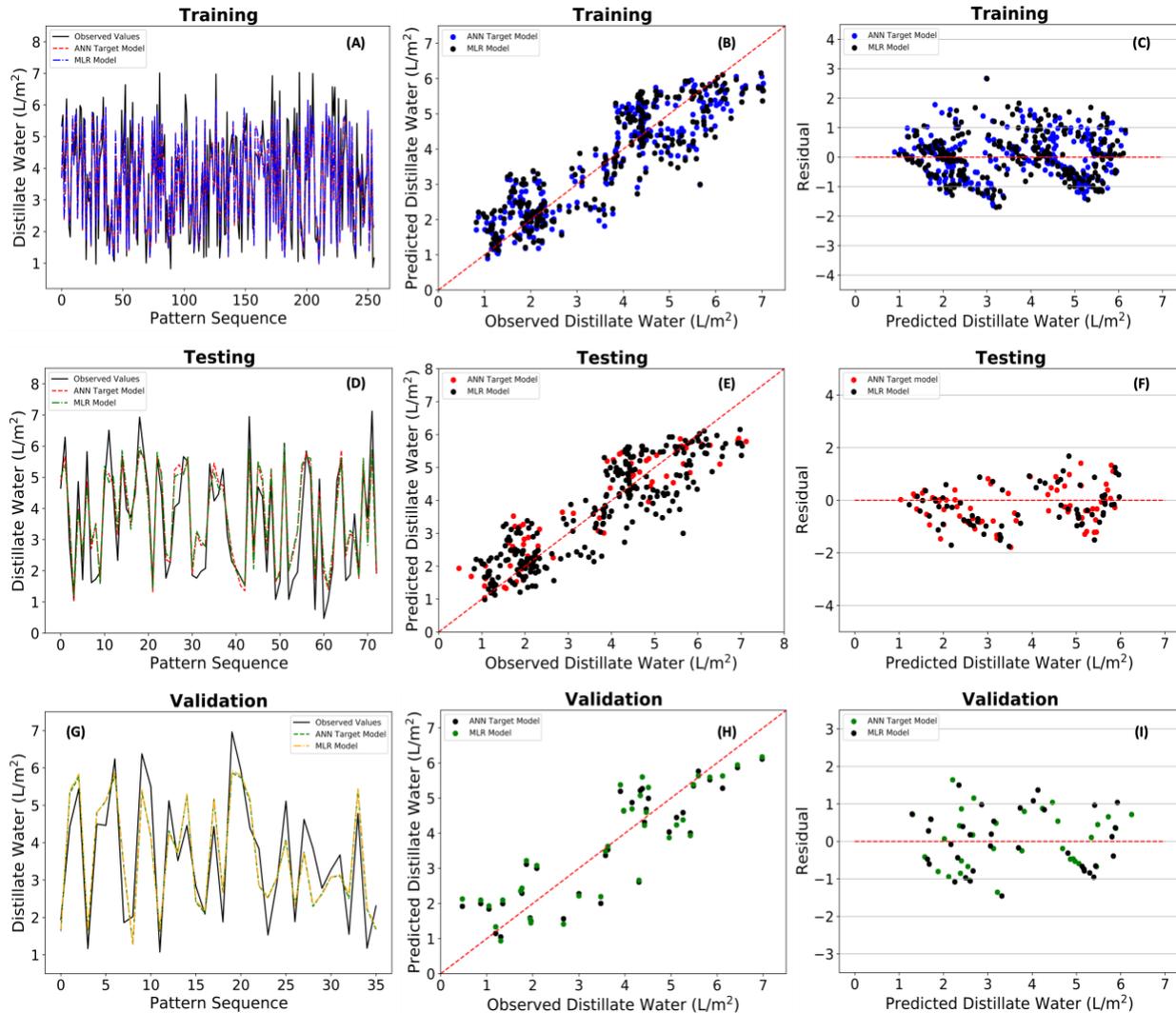


Figure 9: Comparison between observed vs. predicted distillate water values using target ANN and MLR models during modeling stages with their acquired residuals: (A-C) from training datasets; (D-F) from testing datasets; (G-I) from validation datasets.

Likewise, it provided better predictions using the same statistical performance indicators, as demonstrated in Table 4, with 0.832 ME and RMSE value of 0.618, than the RI ANN model, which scored a higher value with an increment of 0.103. These findings strengthen the reliability of the developed ANN models, particularly for ANN as a promising target model.

Moreover, it was verified that the developed latter model effectively forecasted  $P_{std}$  without requiring further experimental work, which consumes more time and involves higher expenses. To sum it up, the target ANN model performed much better than both RI ANN and MLR models during all modeling phases, namely training, testing, and validation sets. The evaluation criteria were in terms of OI with a unique value (0.872) relatively close to 90% compared to RI ANN and

MLR models, which attained less overall performance with reduced values of 0.834 and 0.803, respectively.

It was observed that the prime criteria for assessing the performance of the developed ANN and MLR models were during testing datasets based on statistical analysis results, as presented in Fig. 10. Moreover, these outcomes for treated water predictions were comparatively in conformity with the studies of [40].

It is worth mentioning that the current study involved a comparative analysis of the performance of ANN and MLR models. This comparison is important for several reasons. Firstly, it serves as a foundational benchmark to assess the improvements offered by ANNs. While ANNs are known for their advanced predictive capabilities, particularly in capturing non-linear data relationships, MLR provides a simpler, more traditional model for comparison. Secondly, the analysis demonstrates the advantages of using more complex models like ANN in practical scenarios, such as solar desalination. By showcasing the superior performance of ANN over MLR, the study highlights the effectiveness of ANNs in addressing the inherent complexities of the dataset. Lastly, the inclusion of MLR in the study enhances the reliability and robustness of the findings, ensuring that the observed improvements with ANN are substantive and well-supported. This comprehensive approach contributes to a more nuanced understanding of the value of ANNs in predictive modeling tasks related to solar desalination applications.

H. A. Maddah et al. [3] developed precise supervised machine-learning regressions to generate trained models based on experimental outcomes. The suggested methodology aims to enhance the precision of prediction models by employing dimensional analysis and expanding datasets through in-between randomization. The constructed models were utilized to forecast the thermal efficiency ( $\eta_{th}$ ) when substituting polystyrene with an alternative wall insulation material. The water glass difference temperature ( $T_w - T_g$ ) and evaporative heat transfer coefficients ( $h_{ewg}$ ) were correlated with the solar still outputs using the stepwise linear regression (SLR) method, which yielded minimal statistical errors ( $R^2 = 1$ ) and a root mean square error (RMSE) value  $< 0.016$ .

Another comparable state-of-the-art study by M. Safa et al. [40] was focused on investigating irrigated and dryland wheat fields in Canterbury. The research involved comprehensive data collection methods. The study aimed to identify several direct and indirect parameters contributing to developing an ANN model for forecasting energy consumption in wheat production. The ultimate model could forecast energy consumption by considering various factors such as the size of the crop area, farmers' social considerations, and energy inputs. This model revealed a predictive accuracy for energy usage, with a margin of error of  $\pm 12\%$  (equivalent to  $\pm 2900$  MJ/ha).

Moreover, a comparative analysis between the ANN model and the MLR model showed that the ANN model has superior predictive capabilities for energy consumption compared to the MLR model, specifically within the designated training and validation datasets. The ultimate ANN model can forecast energy consumption by considering various factors such as farm conditions, farmers' social considerations, and energy inputs. This model exhibited a margin of error of  $\pm 2900$  MJ/ha when predicting energy usage. Comparatively, the ANN model demonstrates superior predictive accuracy in estimating energy consumption compared to the MLR model.

Figure 11 displays an illustrative residual histogram of the ANN models employed in the context of the conventional solar still. The study's findings indicate that the ANN modeling can yield test results wherein approximately 80% and 85% of the predictions fall within a 20% to 25% deviation from the actual value for the RI ANN and target ANN models, respectively. The average standard deviations for the relative errors of the optimal ANN predictions for distillate water were 0.46 and 0.50 for the RI ANN and target ANN models, respectively. The residual plots consistently exhibited a minor rightward skew, suggesting a slightly elevated occurrence of overestimations.

Table 4: Statistical parameters for assessing the performance of ANN and MLR models during training, testing, and validation stages.

Statistical Parameters	Testing Models			Training Models			Validation Models		
	ANN-RI	ANN-TL	MLR	ANN-RI	ANN-TL	MLR	ANN-RI	ANN-TL	MLR
<b>R<sup>2</sup></b>	0.803	0.835	0.786	0.801	0.831	0.761	0.797	0.816	0.748
<b>RMSE</b>	0.721	0.618	0.750	0.743	0.622	0.783	0.822	0.708	0.837
<b>MAE</b>	0.602	0.523	0.615	0.607	0.545	0.649	0.685	0.575	0.728
<b>CRM</b>	-0.035	-0.001	-0.034	-0.0024	-0.002	-1.228	-0.001	-0.0026	-0.068
<b>ME</b>	0.801	0.832	0.782	0.798	0.831	0.771	0.795	0.813	0.746
<b>OI</b>	0.834	0.872	0.803	0.832	0.871	0.814	0.841	0.851	0.816

Note: R<sup>2</sup>: coefficient of determination, RMSE: root mean square error, MAE: mean absolute error, ME: model efficiency, OI: overall index of model performance.

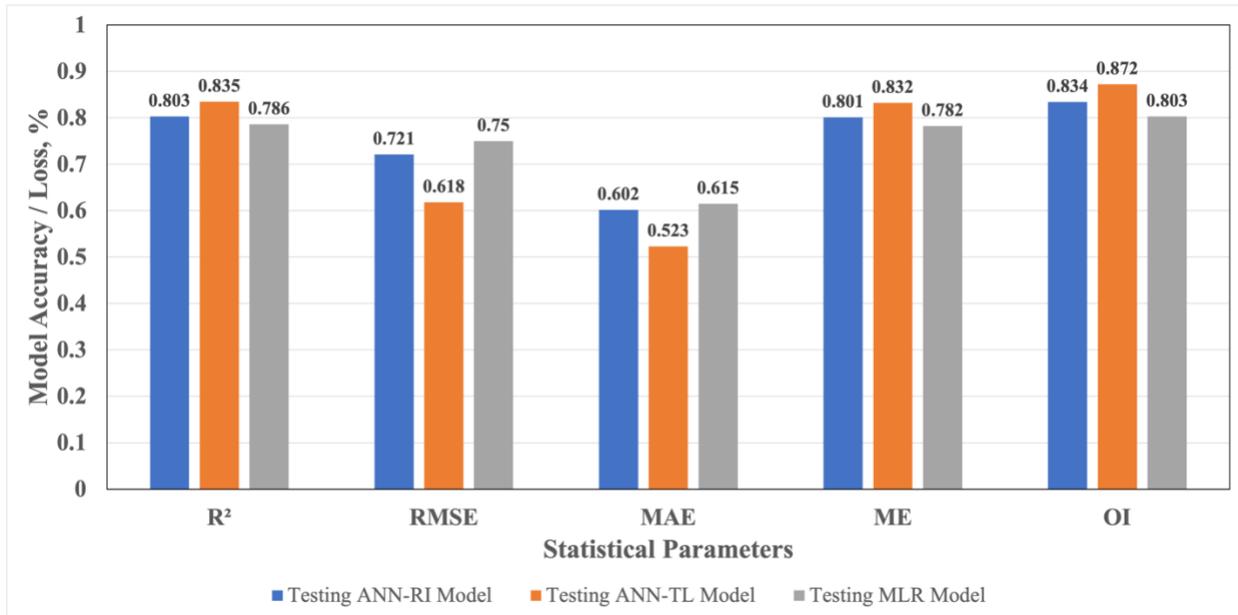


Figure 10: Evaluation of ANN and MLR models' performance during the testing modeling process.

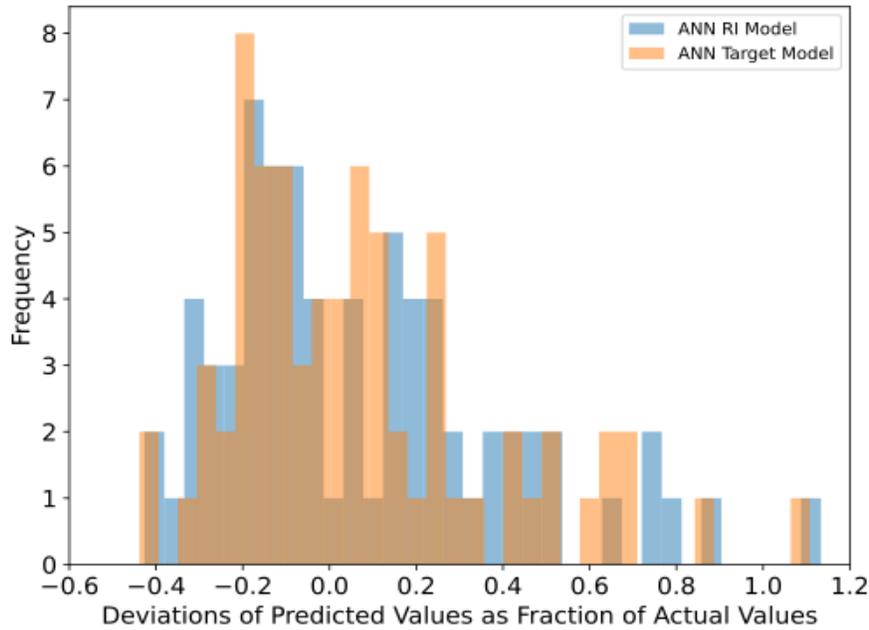


Figure 11: A histogram comparison of deviations between predicted and actual values for the RI ANN model vs. the target ANN model.

## 5. Conclusion

This research paves the way towards creating novel statistical tools for developing precisely accurate predictive supervised ANN models for freshwater productivity in conventional single slope solar stills using the transfer learning method. Due to the significant capital costs associated with solar distillation, which include the cost of land and equipment, the use of solar still technology for saltwater water desalination would demand an optimization process. Therefore, for the entire process to successfully optimize expenses and enhance water production, precise forecasting of the solar still's predicted productivity remains essential. An ANN model was presented in this study to forecast the water productivity ( $P_{std}$ , L/m<sup>2</sup>.day) of solar still desalinating saline water. The primary objective of this work was to apply the transfer learning technique to the RI ANN model for conventional solar still. This method can be achieved by employing the best-learned weights acquired from the developed source ANN model to ultimately generate the optimum ANN model that has the potential to outperform the training ANN model from scratch. The generated ANN was studied and emphasized as a promising target model compared to the created RI and MLR models. The pre-trained ANN model was established from which learned weights (intelligence) were transferred. Several neural network architectures were trained with varying numbers of neurons in the hidden layers. Their hyper-parameters were tuned using the trial-and-error technique to determine which source ANN model offered the best performance and ensure the selection of the best-learned network. The chosen ideal pre-trained ANN model had a 5-64-64-1 architecture with a ReLU transfer function. A feed-forward back-propagation algorithm was utilized to build ANN models, and their applicability for  $P_{std}$  prediction was investigated.

The transfer learning method predominantly enhances the predictive performance of neural networks, which enables the model to begin with a more informed state rather than learning from scratch. This approach is achieved by leveraging a pre-trained model to perform similar tasks. This

method more effectively captures complex patterns and relationships in the selected variables. Furthermore, transfer learning frequently results in improved generalization and faster convergence, particularly when the training datasets are limited. Conversely, the inherent distinctions in the capabilities of the transfer learning-enhanced ANN models over MLR can be attributed to their superiority. The linear nature of MLR renders it incapable of capturing non-linear relationships and interactions between input variables, which neural networks, particularly those with transfer learning, can efficiently model. Therefore, ANN can make more precise predictions in complex scenarios like solar desalination systems.

It was concluded that when employing the transfer learning method, the target ANN model outperformed the RI ANN and MLR models in all phases of the modeling process, including training, testing, and validation, based on the highest recorded OI of 0.872 and lowest RMSE of 0.618 respectively. These findings suggest that the optimal target ANN model may be employed effectively in diverse solar still configuration systems, which desalinates brackish water, to forecast distillate water productions after the solar desalination process. The results also confirmed the developed ANN model's accuracy and generalization capability. Finally, a major contribution of this research implicates an evaluation of the ANN modeling techniques during the saltwater solar desalination process, which adds a new perception and guides us towards tuning key parameters correlated to some design aspects, such as solar still surface area and glass cover area, heat transfer energies, including evaporative, radiative, and convective coefficients, to maximize purified water productions in various configurations of solar still desalination systems.

### Nomenclature

$P_{std}$	—	Solar still productivity, L/m <sup>2</sup> .day
$\eta_{st}$	—	Solar still efficiency, %
$A_{st}$	—	Solar still area, m <sup>2</sup>
$T_w$	—	Basin water temperature, °C
$T_g$	—	Glass Cover Temperature, °C
$P_{st}$	—	Saline water feed flowrate, L/day

### List of Abbreviations

ML	—	Machine Learning
AI	—	Artificial Intelligence
ANN	—	Artificial Neural Network
MLR	—	Multiple Linear Regression
GUI	—	Graphical User Interface
KSA	—	Kingdom of Saudi Arabia
RO	—	Reverse Osmosis
TL	—	Transfer Learning
RI	—	Random Initialization
$R^2$	—	Coefficient of determination
RMSE	—	Root mean square error
MAE	—	Mean absolute error
ME	—	Model Efficiency

OI — Overall index of model Performance

### List of Subscripts

e — Evaporative  
 b-w — Basin to water  
 w-g — Water to glass  
 ft — Feed tank  
 amb — Ambient

### Appendix

#### Mathematical Symbols Definition

Symbol Name		Definition/ Meaning
$Y_y$	—	Output variable identified as distillate water.
$w_{yj}$		Weights between 2 <sup>nd</sup> hidden and output layers.
$x_i$		Input variables.
$w_{jk}$		Weights between the 1 <sup>st</sup> and 2 <sup>nd</sup> hidden layers.
$w_{ki}$		Weights between the input and first hidden layers
$m$		Number of neurons in the 1 <sup>st</sup> hidden layer.
$p$		Number of neurons in the 2 <sup>nd</sup> hidden layer.
$n$		Number of neurons in the input layer
$b_k$		Bias values of the neurons in the 1 <sup>st</sup> hidden layer.
$b_j$		Bias values of the neurons in the 2 <sup>nd</sup> hidden layer.
$b_y$		Bias values of the neurons in the output layer.
$F$		Activation function.
$F_i$		Hidden units/ neurons in the hidden layers.
$w_i$		Weight values that are attached to each input.
$y$		Predicted parameter for the MLR model formula.
$\beta$		The intercept of the MLR model formula.
$\beta_i$		The coefficients of the predictors.
$x_i$		The predictors of the MLR model formula.
$X_0$	—	Original values of input and output variables.
$X_n$	—	Normalized value.
$X_{min}$	—	Minimum of input and output variables.
$X_{max}$	—	Maximum of input and output variables.
$x_{o,i}$		Observed input and output values.
$x_{p,i}$		Predicted value by the machine learning models.
$\bar{x}_o$		Averaged observed/ mathematically obtained values.
$\bar{x}_p$		Predicted values/ theoretically estimated from averaging.
$n$		Number of observations or dataset size.

## Acknowledgment

The authors, H. H. Migaybil and B. Gopaluni express their gratitude for the invaluable support and guidance from the UBC Data Analytics and Intelligent Systems Lab (DAIS) team and the Chemical and Biological Engineering Department at the University of British Columbia. Additionally, they acknowledge the generous support and funding provided by the esteemed Deanship of Scientific Research (DSR) at King Abdulaziz University.

## Conflicts of Interest

There are no conflicts of interest to declare by the authors considering the publication of this paper.

## Data Availability

The data used in this study to support the findings are available upon request from the corresponding authors.

## References

- [1] H. F. Sorour, M. H., El Defrawy, N. M. H., & Shaalan, "Treatment of agricultural drainage water via lagoon/reverse osmosis system," *Desalination*, vol. 152, no. (1–3), pp. 359–366, 2003.
- [2] G. N. Tiwari, H. N. Singh, and R. Tripathi, "Present status of solar distillation," *Sol. Energy*, vol. 75, no. 5, pp. 367–373, 2003.
- [3] H. A. Maddah, M. Bassyouni, M. H. Abdel-Aziz, M. S. Zoromba, and A. F. Al-Hossainy, "Performance estimation of a mini-passive solar still via machine learning," *Renewable Energy*, vol. 162, pp. 489–503, 2020.
- [4] H. A. Maddah and A. M. Chogle, "Applicability of low pressure membranes for wastewater treatment with cost study analyses," *Membr. Water Treat.*, vol. 6, no. 6, pp. 477–488, 2015.
- [5] H. Maddah and A. Chogle, "Biofouling in reverse osmosis: phenomena, monitoring, controlling and remediation," *Appl. Water Sci.*, vol. 7, no. 6, pp. 2637–2651, 2017.
- [6] K. Sampathkumar, T. V. Arjunan, P. Pitchandi, and P. Senthilkumar, "Active solar distillation-A detailed review," *Renew. Sustain. Energy Rev.*, vol. 14, no. 6, pp. 1503–1526, 2010.
- [7] P. M. Cuce, E. Cuce, and A. Tonyali, "Performance analysis of a novel solar desalination system – Part 1: The unit with sensible energy storage with thermal insulation and cooling system," *Sustain. Energy Technol. Assessments*, vol. 37, p. 100566, 2020.
- [8] Z. Kumar, D., Himanshu, P., & Ahmad, "Performance analysis of single slope solar still," *Int J Mech Robot Res*, vol. 3(3), pp. 66–72, 2013.
- [9] T. Arunkumar, K. Vinothkumar, A. Ahsan, R. Jayaprakash, and S. Kumar, "Experimental Study on Various Solar Still Designs," *ISRN Renew. Energy*, pp. 1–10, 2012.
- [10] Hongfei Zheng, 2017) Traditional Solar Desalination Units. *Solar Energy Desalination*

- Technology, pp.259-321.
- [11] S. Kalita, P., Dewan, A., & Borah, "A review on recent developments in solar distillation units," *Sadhana*, vol. 41(2), pp. 203–223, 2016.
  - [12] N. Saxena, A., & Deval, "A high rated solar water distillation unit for solar homes," *J. Eng.* 2016, pp. 1–8, 2016.
  - [13] Kumar, S., & Tiwari, A. (2010). Design, fabrication and performance of a hybrid photovoltaic/thermal (PV/T) active solar still. *Energy Conversion and Management*, 51(6), 1219-1229.
  - [14] Mohaisen, H. S., Esfahani, J. A., & Ayani, M. B. (2021). Improvement in the performance and cost of passive solar stills using a finned-wall/built-in condenser: An experimental study. *Renewable Energy*, 168, 170-180.
  - [15] M. M. Badran, O. O., & Abu-Khader, "Evaluating thermal performance of a single slope solar still," *Heat mass Transf.*, vol. 43(10), pp. 985–995, 2007.
  - [16] N. Wang, Y., Kandeal, A. W., Swidan, A., Sharshir, S. W., Abdelaziz, G. B., Halim, M. A., ... & Yang, "Prediction of tubular solar still performance by machine learning integrated with Bayesian optimization algorithm," *Appl. Therm. Eng.*, vol. 184, p. 116233, 2021.
  - [17] A. Fath, H. E., El-Samanoudy, M., Fahmy, K., & Hassabou, "Thermal-economic analysis and comparison between pyramid-shaped and single-slope solar still configurations," *Desalination*, vol. 159(1), pp. 69–79, 2003.
  - [18] S. K. B. H.A. Maddah, V. Berry, "Biomolecular photosensitizers for dye-sensitized solar cells: recent developments and critical insights," *Renew. Sustain. Energy Rev.*, vol. 121, p. 109678, 2020.
  - [19] L. Trappey, A. J., Chen, P. P., Trappey, C. V., & Ma, "A machine learning approach for solar power technology review and patent evolution analysis," *Appl. Sci.*, vol. 9(7), p. 1478, 2019.
  - [20] M. M. Shokrieh, H. Rajabpour-Shirazi, M. Heidari-Rarani, and M. Haghpanahi, "Simulation of mode I delamination propagation in multidirectional composites with R-curve effects using VCCT method," *Comput. Mater. Sci.*, vol. 65, pp. 66–73, 2012, doi: 10.1016/j.commatsci.2012.06.025.
  - [21] H. Mahmoudi, O. Abdellah, and N. Ghaffour, "Capacity building strategies and policy for desalination using renewable energies in Algeria," *Renewable and Sustainable Energy Reviews*. 2009, doi: 10.1016/j.rser.2008.02.001.
  - [22] A. A. Mashaly, A. F., & Alazba, "Neural network approach for predicting solar still production using agricultural drainage as a feedwater source," *Desalin. Water Treat.*, vol. 57(59), pp. 28646–28660, 2016.
  - [23] N. I. Santos, A. M. Said, D. E. James, and N. H. Venkatesh, "Modeling solar still production using local weather data and artificial neural networks," *Renew. Energy*, vol. 40, no. 1, pp. 71–79, 2012, doi: 10.1016/j.renene.2011.09.018.
  - [24] E. A. M. Hamdan, M. A., Khalil, H. R., & Abdelhafez, "Comparison of neural network models in the estimation of the performance of solar still under jordanian climate," *J. Clean Energy Technol.*, vol. 1(3), pp. 238–242, 2013.
  - [25] D. K. Murugan, Z. Said, H. Panchal, N. K. Gupta, S. Subramani, A. Kumar, and K. K. Sadasivuni, "Machine learning approaches for real-time forecasting of solar still distillate output," *Environmental Challenges*, vol. 13, p. 100779, 2023.
  - [26] S. Nazari, M. Bahiraei, H. Moayedi, and H. Safarzadeh, "A proper model to predict

- energy efficiency, exergy efficiency, and water productivity of a solar still via optimized neural network," *Journal of Cleaner Production*, vol. 277, p. 123232, 2020.
- [27] H. A. Maddah, "Predictive supervised machine learning models for double-slope solar stills," *Desalin. WATER Treat.*, vol. 244, pp. 1–11, 2021.
- [28] H. H. Migaybil, H. A. Maddah, and B. Gopaluni, "Design and Simulation of a Novel Solar Photovoltaic System Assisted a Single-Slope Solar Still Distillation Unit," *Can. J. Chem. Eng.*, vol. 101, no. 6, pp. 3059-3073, 2023.
- [29] NASA, "NASA Prediction Of Worldwide Energy Resource (POWER)," <https://Power.Larc.Nasa.Gov/Data-Access-Viewer/>, 2020. <https://power.larc.nasa.gov>.
- [30] R. J. W. D.E. Rumelhart, G.E. Hinton, "Learning Internal Representations by Error Propagation," in *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*, 1986, p. vol. 1, Chapter 8.
- [31] N. Holur Venkatesh, "Performance evaluation of single and double-basin solar stills in Las Vegas, Nevada," University of Nevada, Las Vegas, United States e Nevada.
- [32] O. Burn, S., Hoang, M., Zarzo, D., Olewniak, F., Campos, E., Bolto, B., & Barron, "A review of the opportunities for desalination in agriculture. Desalination," *Desalin. Tech.*, vol. 364, pp. 2–16, 2015.
- [33] B. K. Rahman, M. M., & Bala, "Modelling of jute production using artificial neural networks," *Biosyst. Eng.*, vol. 105(3), pp. 350–356, 2010.
- [34] A. Zangeneh, M., Omid, M., & Akram, "A comparative study between parametric and artificial neural networks approaches for economical assessment of potato production in Iran," *Spanish J. Agric. Res.*, vol. 9(3), pp. 661–671, 2011.
- [35] Tang, J., Xia, H., Aljerf, L., Wang, D., & Ukaogo, P. O. (2022). Prediction of dioxin emission from municipal solid waste incineration based on expansion, interpolation, and selection for small samples. *Journal of Environmental Chemical Engineering*, 10(5), 108314.
- [36] Xia, H., Tang, J., Aljerf, L., Wang, T., Gao, B., Xu, Q., ... & Ukaogo, P. (2023). Assessment of PCDD/Fs formation and emission characteristics at a municipal solid waste incinerator for one year. *Science of The Total Environment*, 883, 163705.
- [37] M. T. Alazba, A. A., Mattar, M. A., ElNesr, M. N., & Amin, "Field assessment of friction head loss and friction correction factor equations," *J. Irrig. Drain. Eng.*, vol. 138(2), pp. 166–176, 2012.
- [38] Jason Brownlee, "Use Early Stopping to Halt the Training of Neural Networks At the Right Time," December 10, vol. 9, pp. 1–33, 2018, [Online]. Available: <https://machinelearningmastery.com/how-to-stop-training-deep-neural-networks-at-the-right-time-using-early-stopping/>.
- [39] W. Lee, *Python® Machine Learning*. Packt publishing ltd., 2019.
- [40] M. Safa and S. Samarasinghe, "Determination and modelling of energy consumption in wheat production using neural networks: A case study in Canterbury province, New Zealand," *Energy*, vol. 36, no. 8, pp. 5140–5147, 2011.