

Data-driven models of a solar field for membrane distillation: A comparison study

A. Bueso^a, J.D. Gil^b and G. Zaragoza^a

^aCIEMAT-Plataforma Solar de Almería, Ctra. de Senés, km. 4,5, Almería, 04200, Spain

^bCentro Mixto CIESOL, CI3, Ctra. Sacramento, s/n, Almería, 04120, Spain

ARTICLE INFO

Keywords:

Modelling, Flat-plate solar collector, Empirical model, Solar desalination

ABSTRACT


The pressing issue of water scarcity has led to increased research focussing on enhancing access to fresh water, with sustainable desalination emerging as a prominent solution. The use of solar energy is often proposed because of the geographical coincidence of high solar irradiance and water scarcity. However, the variability of the energy source in a stationary-designed process such as desalination must be addressed, and modelling solar desalination systems is crucial to understanding the dynamics and optimising the performance. Solar thermal energy is cheaper to store than photovoltaic energy, and powers advanced desalination technologies such as membrane distillation (MD) that can reach higher water recovery. This study investigates the application of data-driven modelling techniques to an innovative solar collector field providing heat for a MD system. The novelty of using mirrors in the solar field to boost the thermal power yielded renders the classical first-principles-based models presented in the literature invalid, as they cannot account for the nonlinear impact of mirrors in the solar field performance. This justifies the use of new data-driven techniques, and four modelling methodologies are compared, with the NARX artificial neural network that proves the most effective, with an R^2 value of 0.9741 and an RMSE value of 6.3151. The best model is validated by simulation of a solar MD plant.

1. Introduction

The global water shortage is a critical issue and is now recognised as a key obstacle facing society (Fritzmann, Löwenberg, Wintgens, and Melin, 2007; Shannon, Bohn, Elimelech, Georgiadis, Mariñas, and Mayes, 2008; Sharon and Reddy, 2015). Water shortage is a challenge that adversely affects daily life, industry, and, consequently, the economy. Climate change and the growing demand for fresh water further complicate the problem. Consequently, technology has focused on extracting water from the largest existing water source, i.e. seawater, through desalination, a process that removes the salt content from seawater Khawaji, Kutubkhanah, and Wie (2008); Al-Karaghoul and Kazmerski (2013).

Despite the abundance of seawater, advances in conventional desalination methods have persistent limitations related to high energy consumption, usually relying on fossil fuels, and substantial environmental impact, not only associated with subsequent greenhouse gas emissions, but also with the production of brine as a residue (Sharon and Reddy, 2015; Greenlee, Lawler, Freeman, Marrot, and Moulin, 2009; Ghaffour, Bundschuh, Mahmoudi, and Goosen, 2015). In an effort to reduce the dependence on fossil fuels, there is a current trend towards incorporating renewable energy sources in desalination (Ghaffour et al., 2015; Pugsley, Zacharopoulos, Mondol, and Smyth, 2016). The geographical coincidence of water scarcity and abundance of solar radiation supports the use of solar energy for desalination. However, there is a limitation due to the intermittency of solar radiation, given that desalination technologies are usually suited to stationary operation. Therefore, storage energy is required for a steady supply of clean water. The most widely implemented desalination technology is reverse osmosis (RO), powered by electricity, which can be generated through photovoltaic (PV) energy. However, storing electricity requires batteries, which are expensive and environmentally challenging. Thermal desalination technologies, such as Multi-Effect Distillation (MED) and Membrane Distillation (MD) (Qiblawey and Banat, 2008), can be powered with low temperature solar heat, which is easier and cheaper to store as hot water (Chorak, Palenzuela, Alarcón-Padilla, and Abdellah, 2017; Ayou, Zaragoza, and Coronas, 2021; Zhou, Zhang, Sun, and Su, 2018). Solar-powered MD, in particular, is gaining

*Corresponding author

 juandiego.gil@ual.es (J.D. Gil)

ORCID(s):

increasing interest because of its higher thermal efficiency and modularity, which gives it greater adaptability to small sizes and decentralised applications compared to MED. In addition, its good performance in brine concentration makes it a promising alternative to conventional technologies for brine valorisation in circular economy schemes related to zero liquid discharge desalination.

Solar thermal energy is generated through the use of solar thermal collectors, which absorb radiation from the Sun to heat a working fluid (Diez, Navas-Gracia, Martínez-Rodríguez, Correa-Guimaraes, and Chico-Santamarta, 2019). So far, the implementation of theoretical models has facilitated the identification of fundamental design parameters (Deng, Xu, and Yang, 2015), the prediction of the behaviour of the system (Kicsiny, 2016), the application of control strategies (Gil, Roca, Zaragoza, and Berenguel, 2018), and the conduct of economic viability studies that allow comparisons between the various technologies studied (Pasamontes, Álvarez, Guzman, Berenguel, and Camacho, 2013). There are several academic studies employing theoretical models of a flat-plate solar field to assess MD systems and design control strategies. For example, a control system for a pilot MD installation using the theoretical model of a flat-plate solar collector (Gil et al., 2018). In another instance, a theoretical model was used to demonstrate the viability of solar energy systems in MD-based pilot scale desalination systems (Andrés-Mañas, Roca, Ruiz-Aguirre, Acién, Gil, and Zaragoza, 2020). This was achieved by testing the prediction of the system with experimental data and simulating the behaviour of the system for one year.

However, numerous solar field systems currently analysed in literature in the context of solar MD processes are hybrid structures, which makes it difficult to formulate theoretical models because of their complexity. Some illustrations of the systems are presented in (Kabeel, Abdelgaied, and El-Said, 2017; Shafieian and Khiadani, 2019) or Zarzoum, Zhani, Bacha, and Koschikowski (2019). Another example is the one installed in the Plataforma Solar de Almería (PSA, www.psa.es), where mirrors were added to solar flat-plate collectors to reflect some of the sun's rays on the collector's surface (see Figure 1). This results in an increase of the equivalent area of the solar field without changing its size, as the power received by the collectors increases. In these cases, the First Principles (FP) model used by (Gil et al., 2018; Andrés-Mañas et al., 2020), and other authors in the literature is inadequate for the system, as it cannot account for the non-linear impact of mirrors in the solar field performance. Additionally, the FP model presents the issue that it only functions when there is mass flow through the solar collector, which can compromise simulation studies.

The aforementioned complexities need to be addressed and can be solved by developing empirical models that can be used when relationships are uncertain, mathematical descriptions are challenging, and non-linearities are present (Palit and Popovic, 2006). In this context, few examples can be found in the literature. In Porrazzo, Cipollina, Galluzzo, and Micale (2013), the objective was to develop an optimal control strategy for a solar powered membrane distillation using an Artificial Neural Network (ANN), which was in charge of analysing the performance of the process under different operating conditions. Based on the developed model, a control system was implemented that optimised distillate production under varying operating conditions. Behnam, Zargar, Shafieian, Razmjou, and Khiadani (2023) compared the predictive capacity of two different ANN models in a solar-powered Direct Contact Membrane Distillation (DCMD) system. The results show the good fit obtained with ANNs, especially with multilayer perceptron models. Diez et al. (2019) developed an empirical method to forecast the temperature at the outlet of a flat-plate solar collector, using solar irradiance, ambient temperature, inlet temperature, and fluid flow as input. The results obtained involve an ANN configured with four neurones as input, five neurones in a hidden layer, and one neurone as the model's outlet. Farkas and Geczy-Vig (2003) employed an ANN structure as an empirical model for a flat-plate solar collector, which was effectively trained for three different variations of solar collectors. However, despite the success of the strategies reviewed above, there are some shortcomings remaining in the literature.

1. The lack of comparative studies of data-based models for these types of applications. It can be seen how in most cases ANN models are chosen without analyzing different data-based modeling methodologies that deal with nonlinearities under different principles.
2. The lack of large datasets for the development of these models, since in most cases data sets limited to a few days of operation are used, which do not reflect a complete operating day of the system. Missing the transient behavior of the system when the pumps begin to operate.
3. The lack of data-driven modelling approaches applied to hybrid solar collector field structures in the context of MD systems.

This study aims to improve the state-of-the-art by developing a data-driven model capable of simulating the performance of flat plate solar systems, taking into account the effects of reflectors, variable ambient conditions, and

the operation of the solar field circulation pump. This model can be applied as a powerful tool to develop effective control / operating strategies and to facilitate the evaluation of MD systems in a simulated environment. The procedure presented in this work involves the comparison of various data-driven modelling strategies to determine the model that best fits the data acquired over 25 days with a sampling interval of one second, taking into account a full day's data, unlike existing models in the literature, which use fewer data points and only when the pumps are active. By evaluating a full 24-hour period and considering the operation state of the pumps, a model can be created to explore various strategies to operate this type of system. In general, the original and innovative nature of this investigation can be summarised as follows:

- Data-driven models have been developed to accurately predict the behaviour of solar fields used in MD units. Factor such as pump status and the influence of the mirror element in the outlet temperature of the solar field have been taken into account, which add significant progress in relation to previous studies in literature.
- Linear and nonlinear models were tested to assess the validity of these kind of strategies for the application at hand.
- A large amount of data, recorded over a period of 25 days with a sampling time of one second considering data from an entire day, were used to create the solar field model. The acquired data enable the model to forecast the transient state that happens when the pump initiating the fluid flow through the solar field is activated.
- The performance of the most effective data-driven model is demonstrated in a simulation environment, in which it is used as a heat source for a MD process, similar to that installed in PSA.

The paper is structured as follows: Section 2 details the experimental data and facilities used, while Section 3 provides an in-depth explanation of the modelling methodology used. Section 4 presents a summary and discussion of the primary outcomes achieved. Finally, Section 5 includes the conclusions of the research and a discussion of the work in progress.

2. Experimental facility and data collected

This study examines a solar field based on flat-plate collectors located at the PSA experimental plant for solar membrane distillation. This solar field (see Figure 1) comprises four files of five solar thermal collectors each, totalling an area of 44.4 m². The panels were inclined at an angle of 45° and operated at an internal hydraulic pressure of 2-2.5 bar. Furthermore, the solar field utilises mirrors to redirect certain reflected sun rays towards the collectors, resulting in an increase in the equivalent area of the field with the additional power received. The mirrors adjust their position on the basis of solar time, ranging from 120° to -6°. They have a mirror surface of 34.85 m² and there are a total of 15 mirrors. A motor (GEARED MOTOR model 24.15.300-REV. A1) and four roller bearings allow angular movement.



Figure 1: The PSA setup includes the solar thermal field, the thermal storage tank, the MD unit, and the heat exchanger.

The solar field that is being studied provides the necessary heat to operate a pilot plant for solar membrane distillation. The experimental installation comprises a thermal storage tank, a heat exchanger, and an MD unit, which

includes a cooling circuit to refrigerate the feed solution. The storage tank, as shown in Figure 1, is a 1.5 m³ tank enveloped by a thick layer of insulation. It is connected to the solar field for heat storage during times of excess energy and for use when solar energy is unavailable. The MD unit, illustrated in Figure 1, uses spiral wound multi-wound MD modules from the Dutch firm Aquastill BV. These operate in Vacuum Air Gap Membrane Distillation (V-AGMD) mode, comprising a total membrane surface of 26 m², with 12 envelopes of 2.7 m² each for increased energy efficiency. Lastly, the experimental plant uses a plate heat exchanger (Sondex BV PHE type S14A-IG16-13-TL-LIQUID), which is shown in Figure 1, with an effective heat exchange surface of 1.65 m².

As mentioned, in this solar field an experimental campaign of 25 days was performed. The recorded variables include the solar field inlet temperature ($T_{s,in}$), the solar field outlet temperature ($T_{s,out}$), the ambient temperature (T_a), irradiance (I), and solar field flow rate (m_s). The recorded data collected from the solar field are presented in Figure 2. Note that from a modelling/control point of view, the solar field can be described as a multi-input, single-output (MISO) system as presented in Figure 3. In this representation, the input of the model includes m_s , which represents the controllable input, $T_{s,in}$, T_a , and I , which represent the disturbance inputs, and $T_{s,out}$, which represents the solar field output.

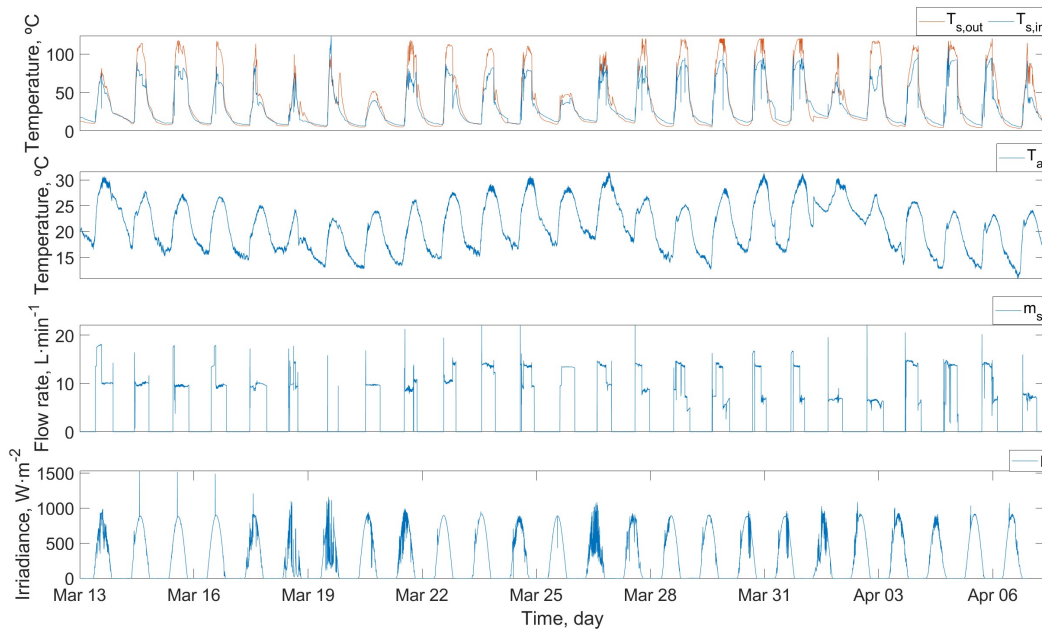


Figure 2: The picture shows the data collected from the solar field.

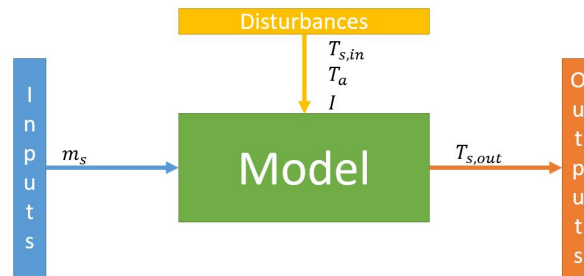


Figure 3: Diagram depicting the input provided to the model, the model's output, and the disturbances that influence the output.

3. Modelling methodologies

In previous studies, a lumped-parameter FP model was used mainly to successfully simulate the dynamics of the solar field (Andrés-Mañas et al., 2020; Ampuño, Roca, Berenguel, Gil, Pérez, and Normey-Rico, 2018; Pataro, Gil, Guzmán, Berenguel, and Lemos, 2023). However, and as pointed out in the Introduction section, since this model does not include the mirrors within the solar field and has not been validated when there is no feed flow through the collectors, the use of data-driven models has been considered in this work. For this purpose, various structures of data-driven models have been assessed. The idea is to compare linear structures, the behaviour that the system would have at a specific operating point, to non-linear structures which treat non-linearity with different principles, with the aim of identifying the most appropriate methodology for modelling these systems. Thus, the range of models tested is described in the following.

- **Autoregressive with Extra Input (ARX).** The ARX model was selected to evaluate the fit of a linear model to the solar field system. In this model, the output of a system is computed as a linear function of its previous values and external or exogenous inputs.
- **Hammerstein-Wiener (HW).** The HW structure was assessed due to its ability to describe the dynamics of a system using a transfer function, which represents a linear model. Additionally, this structure can incorporate nonlinearities by utilizing nonlinear functions of the inputs and outputs of the linear system.
- **Nonlinear Autoregressive with Extra Input (NARX).** The NARX models are based on the principles of ARX models. However, they incorporate a non-linear component that is applied to the time series data.
- **NARX-ANN.** This kind of model consists of a dynamic recurrent neural network, which incorporates feedback connections with multiple layers. Leveraging its memory capacity, this model utilizes past values of predicted or actual time series.

The aforementioned models have their own differences and enable the evaluation of various strategies from diverse perspectives. These encompass time-series-orientated models, both linear and nonlinear in nature, models founded on fundamental components like transfer functions, and models purely characterised by non-linear attributes, such as neural networks. The mathematical formulation of each model will be described in the following sections.

3.1. ARX model

The ARX model, used in time series analysis, comprises two key elements: autoregressive (AR) terms and exogenous (X) terms (Mathworks, 2023). The autoregressive component delineates the connection between a variable and its historical values, while the exogenous component elucidates the impact of external factors on the variable of interest. The mathematical formulation of this model, considering a SISO structure for the sake of visualisation, is given by:

$$y(t) = -a_1 \cdot y(t-1) - \dots - a_{na} \cdot y(t-na) + b_1 \cdot u(t-1-t_d) + \dots + b_{nb} \cdot u(t-nb-t_d) + e(t), \quad (1)$$

where $y(t)$ represents the output at time t , na stands for the number of coefficients associated with the autoregressive part, nb denotes the coefficients that depend on external inputs, $y(t-1), \dots, y(t-na)$ refers to the previous outputs on which the current output is based. Furthermore, $u(t-1-t_d), \dots, u(t-nb-t_d)$ represents the previous and delayed inputs on which the current output depends, where t_d stands for the delay, while $e(t)$ represents the white noise. A more concise rewording of the difference equation is provided below:

$$A(z^{-1}) \cdot y(t) = B(z^{-1}) \cdot u(t-t_d) + e(t), \quad (2)$$

where z^{-1} represents the back shift operator, and $A(z^{-1})$ and $B(z^{-1})$ are defined by the following expressions:

$$A(z^{-1}) = 1 + a_1 \cdot z^{-1} + \dots + a_{na} \cdot z^{-na}, \quad (3)$$

$$B(z^{-1}) = b_0 + b_1 \cdot z^{-1} + \dots + b_{nb} \cdot z^{-nb}. \quad (4)$$

Note that Equation (2) can be easily expanded to a MISO system, which is the case at hand in this work:

$$A(z^{-1}) \cdot y(t) = B_1(z^{-1}) \cdot u_1(t-t_{d_1}) + B_2(z^{-1}) \cdot u_2(t-t_{d_2}) + \dots + B_{nu}(z^{-1}) \cdot u_{nu}(t-t_{d_{nu}}), \quad (5)$$

where the sub-index 1, 2, ..., nu denote the number of inputs, being B_1, \dots, B_{nu} , the polynomials associated to the respective input u_1, \dots, u_{nu} . Note that these inputs can represent both the inherent inputs of the process and disturbances.

3.2. HW model

HW models describe dynamic systems using one or two non-linear static blocks in series with a linear block, which is a discrete transfer function that represents the dynamic element of the model (Karaca, Özen, and Kasnakoğlu, 2016). The structure of an HW model can be seen in Figure 4. In this structure, f is a non-linear function that transforms input data $u(t)$ into $w(t) = f(u(t))$, A/B represents a linear transfer function, following the definition in Equation 3 and 4. This function transforms $w(t)$ into $x(t) = (A/B) \cdot w(t)$, and h is a non-linear function that maps the output of the linear block $x(t)$ to the output of the system $y(t)$ into $y(t) = h(x(t))$. The nonlinear terms of the HW model, i.e., functions f and h can be determined using various non-linear functions.

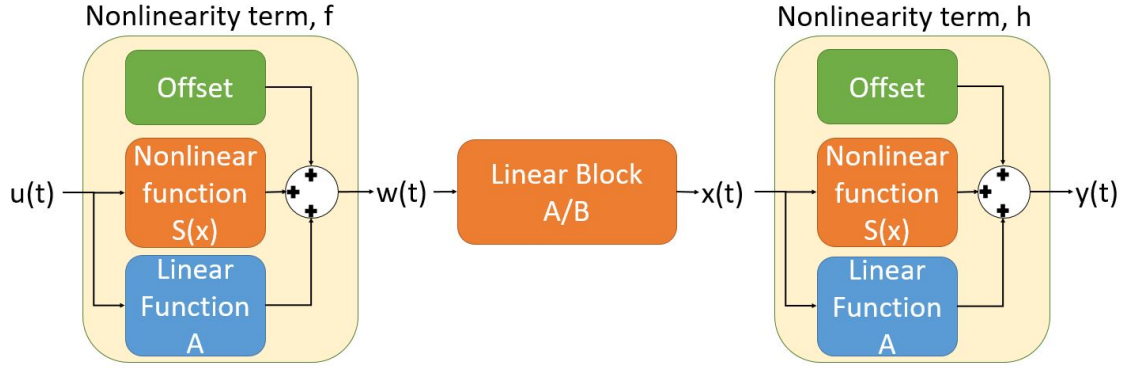


Figure 4: Structure of HW model.

In this study, a Sigmoid function was used as a non-linear function and is described by the following expression (Koroglu, Sergeant, and Umurkan, 2010; Mathworks, 2023a):

$$S(X) = \sum_{i=1}^n s_i \cdot g(X^T \cdot Q \cdot d_i + c_i), \quad (6)$$

where $S(X)$ is equal to the sum of the dilated and translated sigmoid functions. Note that this variable generically represents the output of the non-linear block, which may correspond to $w(t)$ or $y(t)$ depending on the representation in Figure 4. Q represents a projection matrix $m - by - q$ where m is greater than or equal to q . The scalar weights, called the output coefficients, are denoted by s_i , and the vectors $q - by - 1$, called the dilation coefficients, are denoted by d_i . Scalars, called translations, are denoted by c_i . The unit function of the sigmoid network, commonly known as the sigmoid function, is mathematically expressed as $g(z) = \frac{1}{e^{-z} + 1}$. Here, the variable z takes the form of a scalar $X^T \cdot Q \cdot d_i + c_i$.

3.3. NARX model

The NARX model extends the linear ARX model to the non-linear case (Mathworks, 2023b). The structure of these models uses flexible non-linear functions, including wavelet and sigmoid networks, to represent intricate non-linear phenomena (Fernández-Ares, Mora, Odeh, Garcia-Sanchez, and Arenas, 2017).

The main features of these models lie in the fact that, instead of reflecting a direct weighted sum of the regressors, as in the ARX model (which represents a linear mapping), the non-linear ARX model offers a flexible non-linear mapping function L . In this study, L is given by a Sigmoid function, as presented in the previous section. Thus, the mathematical model of an NARX model for a SISO case can be formulated as follows:

$$y(t) = -a_1 \cdot y(t-1) - \dots - a_{na} \cdot y(t-na) + b_1 \cdot u(t-1-t_d) + \dots + b_{nb} \cdot u(t-nb-t_d) + L(y(t-1), y(t-2), \dots, y(t-n_a), u(t-t_d), u(t-1-t_d), \dots, u(t-n_b-t_d)) + offset, \quad (7)$$

where the same nomenclature for the input and output of the process as in the ARX case has been used, and *offset* represents a bias vector. The structure can be summarised as presented in Figure 5. Note that this structure can be easily expanded to the MISO case, as presented in the ARX methodology.

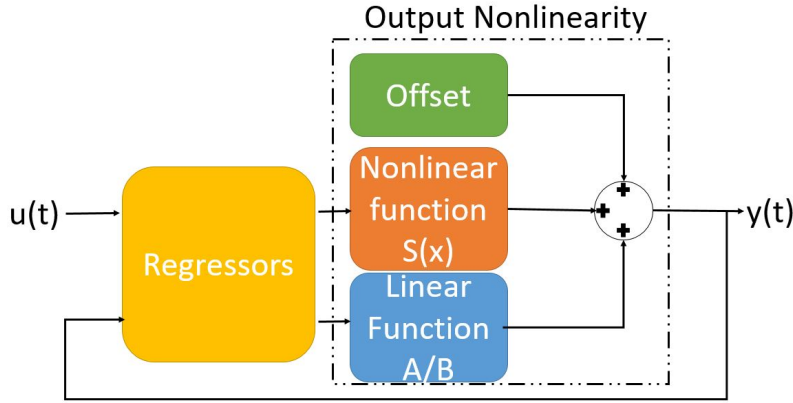


Figure 5: Structure of NARX MISO model.

3.4. NARX ANN model

The NARX model is used in this study as the foundation for the neural network structure suggested for the handling of time series data. If the approximation of the function L in the mathematical formulation of the NARX model (see Equation 7) is done using multilayer perceptrons, the resulting network is called the NARX network (Di Piazza, Di Piazza, and Vitale, 2016). In essence, a NARX network is made up of a multilayer perceptron that takes a window of previous independent inputs and outputs as input and calculates the current output (Di Piazza et al., 2016). The NARX network has limited feedback, derived only from the output neurone and not from hidden states, unlike a traditional recurrent neural network (Di Piazza et al., 2016). The typical structure of the NARX ANN model is shown in Figure 6, where IW is the input weight matrix, *offset* is the hidden layer bias vector and LW is the output layer weight matrix.

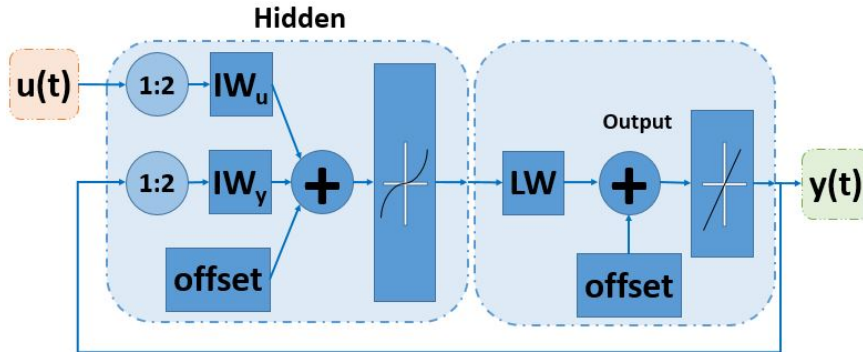


Figure 6: Schematic diagram of an NARX ANN.

3.5. Regression metrics

Regression metrics assess the effectiveness of a predictive model that estimates a numerical value. In this study, to evaluate the goodness of the model, classic error metrics such as Root Mean Square Error (RMSE) have been used along with a statistic-based measure such as the coefficient of determination (R^2). Both metrics are defined as follows:

- Root Mean Square Error (RMSE): This metric indicates the square root of the Mean Square Error (MSE) as follows:

$$RMSE = \sqrt{MSE}. \quad (8)$$

where the MSE is mathematically defined as:

$$MSE = \frac{1}{M} \cdot \sum_{i=1}^M (MV - PV)^2, \quad (9)$$

where M represents the total number of observations, MV denotes the measurement value, and PV refers to the predictive value.

- Coefficient of determination (R^2): This metric is a statistical measure that represents the proportion of the variance in the dependent variable that is predictable from the independent variables in a regression model. The calculation of this metric is given by:

$$R^2 = 1 - \frac{SSR}{SST}, \quad (10)$$

where SSR represents the sum-squared regression and SST denotes the total sum of squares.

4. Results

This section presents the results obtained. It should be noted that the input-output representation presented in Figure 3 has been used in all modelling methodologies. In this representation, the input of the model includes m_s , which represents the controllable input, $T_{s,in}$, T_a , and I , which represent the disturbance inputs, and $T_{s,out}$, which represents the solar field output. The used sub-indexes 1,...,4 stand for $T_{s,in}$, m_s , I , and T_a , respectively.

4.1. Empirical model training and validation

Before presenting the different models obtained, it should be noted that, in this study, 90% of the collected data were used for training purposes and the remaining 10% for validation.

4.1.1. ARX model

In the context of ARX models, two distinct structures were tested, distinguished by the number of regressors involved in each case. The proposed models were as follows:

- ARX121: $na = 1$, $nb_1 = 2$, $nb_2 = 2$, $nb_3 = 2$, $nb_4 = 2$ and $t_{d1} = 1$, $t_{d2} = 1$, $t_{d3} = 1$ and $t_{d4} = 1$.
- ARX221: $na = 2$, $nb_1 = 2$, $nb_2 = 2$, $nb_3 = 2$, $nb_4 = 2$ and $t_{d1} = 1$, $t_{d2} = 1$, $t_{d3} = 1$ and $t_{d4} = 1$.

During the training phase, the goodness of fit achieved by each of the preceding model structures was comparable (refer to Table 1). Among them, the ARX221 model emerged as the optimal choice, yielding the lowest RMSE metric. Detailed information on the model parameters for each structure is provided in Table 2. It is important to note that these parameters are in accordance with the ARX definition described in Equation 5.

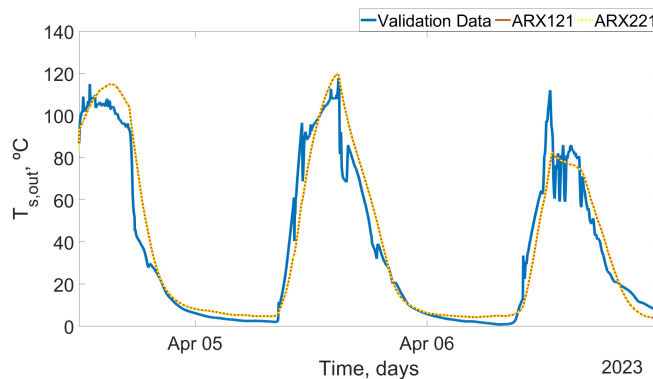


Figure 7: ARX 121 and ARX221 model validation.

ARX model	RMSE	R^2
ARX121	12.7753	0.8703
ARX221	12.5938	0.8739

Table 1

ARX model training goodness.

Function	Parameters	ARX121	ARX221
A	a_1	-1.000	-1.425
	a_2	0.000	0.425
B_1	b_0	$7.159 \cdot 10^{-3}$	$6.567 \cdot 10^{-3}$
	b_1	$-6.974 \cdot 10^{-3}$	$-6.458 \cdot 10^{-3}$
B_2	b_0	$2.552 \cdot 10^{-2}$	$2.148 \cdot 10^{-2}$
	b_1	$-2.540 \cdot 10^{-2}$	$-2.142 \cdot 10^{-2}$
B_3	b_0	$6.841 \cdot 10^{-5}$	$1.164 \cdot 10^{-4}$
	b_1	$-5.506 \cdot 10^{-5}$	$-1.086 \cdot 10^{-4}$
B_4	b_0	$-1.168 \cdot 10^{-4}$	$-8.865 \cdot 10^{-4}$
	b_1	$-2.312 \cdot 10^{-4}$	$6.879 \cdot 10^{-4}$

Table 2

ARX model parameters.

ARX model	RMSE	R^2
ARX121	10.2262	0.9322
ARX221	10.0534	0.9345

Table 3

ARX model validation goodness.

HW model	RMSE	R^2
HW231	6.1308	0.9715
HW131	5.7944	0.9732

Table 4

HW model training goodness.

After the training phase was completed, model validation was performed using the validation dataset and metrics outlined in Section 3.5. The numerical metric values are presented in Table 3, while Figure 7 provides a graphical comparison between the actual and predicted temperatures. In line with the results obtained in the simulation phase, the model that yielded the best outcome was the ARX121 according to both metrics. However, the results were very similar (this is confirmed also in the visual validation, see Figure 7); therefore, for comparative purposes with other modelling techniques, the ARX121 model was used because it is a lower order linear model.

4.1.2. HW model

The second methodology examined was the HW modelling approach. Similarly to the earlier case, various model structures were employed, which are delineated as follows:

- HW231: Input non-linear term: none. Linear function: $na_1 = 2, na_2 = 2, na_3 = 2, na_4 = 2, nb_1 = 1, nb_2 = 1, nb_3 = 1, nb_4 = 1, t_{d1} = 1, t_{d2} = 1, t_{d3} = 1$ and $t_{d4} = 1$. Output non-linear term: Sigmoid with 10 units.
- HW131: Input non-linear term: none. Linear function: $na_1 = 1, na_2 = 1, na_3 = 1, na_4 = 1, nb_1 = 1, nb_2 = 1, nb_3 = 1, nb_4 = 1, t_{d1} = 1, t_{d2} = 1, t_{d3} = 1$ and $t_{d4} = 1$. Output non-linear term: Sigmoid with 10 units.

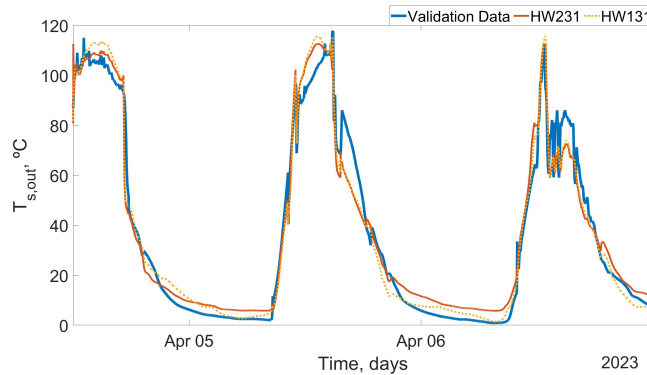
After training the models, the goodness of fit obtained for each model is detailed in Table 4. Based on the error metrics, the HW131 configuration emerged as the optimal choice in this case. With respect to the model parameters for each case, in HW models, three distinct sections are discernible. Firstly, there is the non-linearity component of the input, which predominantly influences the input of the system, as illustrated in Figure 4. It should be noted that

Function	Parameters	HW231	HW131
A_1	a_1	1.0000	1.0000
	a_2	-0.9815	0.0000
A_2	a_1	-0.0224	-0.1559
	a_2	0.0224	0.0000
A_3	a_1	0.0005	0.0058
	a_2	0.0224	0.0000
A_4	a_1	-0.0854	$5.5 \cdot 10^{-7}$
	a_2	0.0224	0.0000
B_1	b_0	-0.9671	-0.3237
	b_1	0.9192	0.3268
	b_2	-0.9465	-0.9952
B_2	b_0	-0.9949	-0.3640
	b_1	-0.9975	0.3639
	b_2	0.9925	-0.9985
B_3	b_0	-1.0040	-2.1430
	b_1	-0.9862	2.1400
	b_2	0.9907	-0.9959
B_4	b_0	-2.0900	-1.6670
	b_1	1.3510	0.3348
	b_2	-0.2609	0.3326

Table 5

Linear block parameters for each HW model.

different configurations were tested, with the most favourable results observed when excluding the input non-linearity block. Consequently, this block was omitted from both HW structures. The second section comprises the linear block, with the parameters detailed for each case in Table 5. Finally, the third section encompasses an output non-linearity block, whose resulting parameters can be found in Table 6.

**Figure 8:** HW231 and HW131 model validation.

As in the previous case, the models obtained were validated with the validation dataset and the results are summarised in Table 7 and illustrated in Figure 8. As expected, HW131 obtained better results than HW231, thus being selected for comparison with the other modelling structures. Moreover, note that the validation results are similar to those obtained in the training phase and are superior to those of the ARX structure. This indicates that HW models capture the nonlinear component of the system and demonstrate better effectiveness than linear models.

4.1.3. NARX model

The third method under investigation was the NARX modelling approach. In a similar way to the previous case, different model structures were used, with the NARX model, described as follows, being the best result obtained:

Nonlinearity term	Parameters	HW231	HW131
Offset		12.5524	3.6851
Linear function	a_{1_1}	0.3399	0.0079
Non linear function	d_i^T	(0.3371)	(0.0285)
		1.1447	0.0157
		1.0183	0.0207
		0.7659	0.0161
		0.8026	0.0723
		-0.0278	0.0046
		0.0370	0.0011
		0.1637	0.0021
		0.1634	0.0033
		0.0624	0.0024
	c_i^T	(38.2299)	(72.0485)
		111.2431	100.4171
		92.5516	86.2473
		60.8712	62.7955
		50.2676	66.6780
		-8.9822	11.0695
		-0.7554	-3.0629
		-18.0603	-10.7457
		-28.4645	2.7228
		-18.0956	-27.2002
	s_i	(-8.0606)	(-1.3656)
		-0.6457	0.5144
		-0.4411	-2.1072
		-0.1394	-0.5516
		1.2008	0.0488
		0.7048	1.4938
		-16.9892	-15.0580
		8.4333	16.7033
		-0.1028	-2.7128
		-69.5780	-52.2801
	Q	1	1

Table 6

Output nonlinearity parameters for each HW model.

HW model	RMSE	R^2
HW231	6.9612	0.9663
HW131	6.6065	0.9690

Table 7

HW model goodness.

NARX model	RMSE	R^2
NARX	8.2113	0.9479

Table 8

NARX model training goodness.

- Linear function: $na_1 = 2, na_2 = 2, na_3 = 2, na_4 = 2, nb_1 = 1, nb_2 = 1, nb_3 = 1, nb_4 = 1, t_{d1} = 0, t_{d2} = 0, t_{d3} = 0$ and $t_{d4} = 0$.
- Non-linear function: Sigmoid with 10 units.

The goodness of fit achieved after the training phase is detailed in Table 8. Regarding the model parameters, in the NARX model, two distinct sections are discernible. First, the Regressors block defines the regressors used by the

Nonlinearity term	Parameters	NARX
Offset	$offset$	0.0424
Linear function	a_{1_1}	32.0437
	a_{2_1}	14.6234
	b_{1_1}	-3.3353
	b_{2_1}	-0.1706
	b_{1_2}	0.3588
	b_{2_2}	0.0014
	b_{1_3}	-0.0002
	b_{2_3}	0.0015
	b_{1_4}	0.0002
	b_{2_4}	0.0172
Non linear function	d_i	(-0.2284 0.1791 0.2809 0.6329 0.6926 0.6711 0.2481 -0.4841 -0.0263 0.5670)
		(0.2670 0.1044 -0.2817 -0.4418 0.1526 -0.2743 0.2707 -0.0591 -0.2134 -0.2406)
		(0.0806 0.0003 0.1538 -0.0956 0.1577 -0.0612 0.0658 0.1887 0.1009 0.1677)
		(-0.2598 -0.1072 0.2210 0.1922 0.1470 -0.2599 0.0583 -0.0778 -0.2229 0.0117)
		(0.1044 -0.2135 -0.2003 -0.1247 -0.1262 0.0281 0.1678 -0.0441 -0.0973 0.2225)
		(0.0012 -0.0053 -0.0003 0.0023 -0.0019 0.0070 -0.0187 -0.0031 -0.0034 0.0052)
		(0.0410 0.0128 0.0326 -0.0064 0.0274 -0.0307 0.0470 0.0386 0.0392 -0.0231)
		(0.0026 -0.0038 -0.0048 0.0022 -0.0041 0.0023 0.0071 0.0055 0.0020 -0.0005)
		(-0.0035 0.0033 -0.0003 0.0031 0.0037 0.0007 -0.0039 0.0001 -0.0037 -0.0028)
		(-0.0016 0.0035 -0.0012 -0.0008 -0.0013 0.0016 0.0005 -0.0048 -0.0045 0.0001)
		(3.4083 -3.4563 -3.0320 -1.8626 -0.9557 0.3815 0.4594 -1.3544 -2.744 2.5244)
		(-0.0314 0.1254 -0.0264 -0.0095 0.0516 0.0620 -0.0017 0.061 0.0086 -0.0848)
		(0.0001 0.0202 -0.0567 -0.0089 0.0244 0.0005 0.0222 0.2445 0.0370 53.1752)
		(0.0001 0.0202 -0.0567 -0.0090 0.0244 0.0005 -0.0222 -0.2448 -0.0345 -53.1639)
Q	s_i	(0.0001 0.0164 0.0692 0.0280 -0.0201 -0.0010 0.0299 0.0397 22.9065 -0.0892)
		(0.0001 0.0164 0.0692 0.0280 -0.0202 0.0001 -0.0300 -0.0387 -22.9082 0.0795)
		(0.0000 0.0000 0.0211 -0.0725 0.1597 0.0002 0.0049 21.6346 -0.0470 -0.6045)
		(0.0000 0.0000 0.0211 -0.0725 0.1596 -0.0006 -0.0051 -21.6358 0.0368 0.5978)
		(0.0015 -0.0031 0.0008 0.0017 -0.0023 0.4272 0.0002 -0.0203 0.0292 0.0061)
		(0.0015 -0.0031 0.0005 0.0011 -0.0026 -0.4272 -0.0002 0.0202 -0.0292 -0.0068)
		(0.0000 0.0020 0.0066 -0.1390 -0.0890 0.0007 10.6814 -0.0106 -0.0618 -0.1007)
		(0.0000 0.0020 0.0066 -0.1390 -0.0890 0.0007 -10.6811 0.0110 0.0667 0.1193)

Table 9
NARX model parameters.

non-linear function and the linear function, both of which have two regressors for each variable. Finally, the second section encompasses an output non-linearity block, whose resulting parameters can be found in Table 9. After the training phase, the model obtained was validated with the validation dataset and the results are summarised in Table 10 and illustrated in Figure 9.

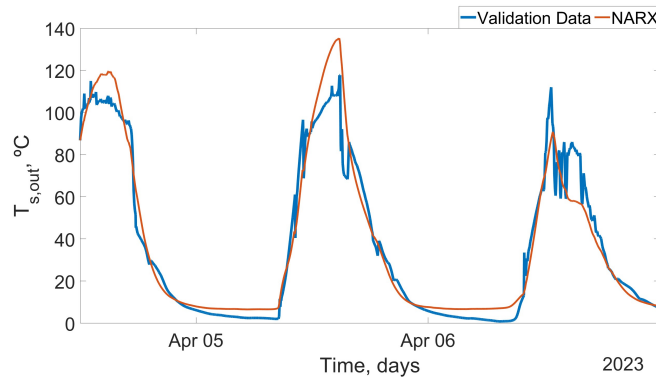


Figure 9: NARX model validation.

NARX model	RMSE	R ²
NARX	9.0790	0.9439

Table 10

NARX model goodness.

NARX ANN model	RMSE	R ²
NARX ANN	0.0182	1.000

Table 11

NARX ANN model training goodness.

Parameters	NARX ANN				
IW	1.6488	1.6483	-0.7373	-0.6687	0.9498
	0.2974	-0.0885	-0.4170	0.6028	-0.5037
	-2.7572	0.6397	3.5943	2.2517	0.9782
	-0.0184	0.0104	0.0122	-0.0121	0.0694
	0.0199	-0.0032	-0.0366	0.0324	0.3916
	0.0217	-0.0119	-0.0157	0.0155	0.1379
	1.6548	-0.7224	-0.3008	0.0214	2.9393
$offset$			-3.0267		
			-0.4968		
			0.3647		
			-0.0715		
			-0.1523		
			-0.0409		
			4.0714		
LW			0.5757		
			-0.0001		
			-0.0036		
			0.0001		
			5.7495		
			-0.2770		
			5.1400		
			0.0012		

Table 12

Optimal network weights and bias.

4.1.4. NARX ANN model

Finally, the system's behaviour was characterised by utilising the NARX ANN model. In a similar way to the previous cases, different model structures were used, with the NARX ANN model, described as follows, being the best result obtained:

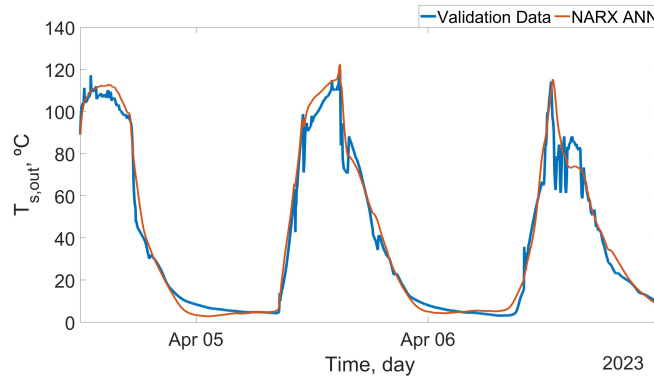
- Algorithm employed: Lavenberg-Marquadt.
- Number of hidden layers: 1.
- Number of neurones in each layer: 7.

During the training phase, the goodness of fit achieved is detailed in Table 11. As in previous cases, the model parameters were presented in Table 12, the model obtained was validated with the validation dataset, and the results are summarised in Table 13 and illustrated in Figure 10. Moreover, note that the validation results were similar to those obtained in the training phase and superior to those of the HW structure. This indicates that the NARX ANN model captures the non-linear component of the system and shows better effectiveness than HW models.

NARX ANN model	RMSE	R ²
NARX ANN	6.3151	0.9777

Table 13

NARX ANN model goodness.

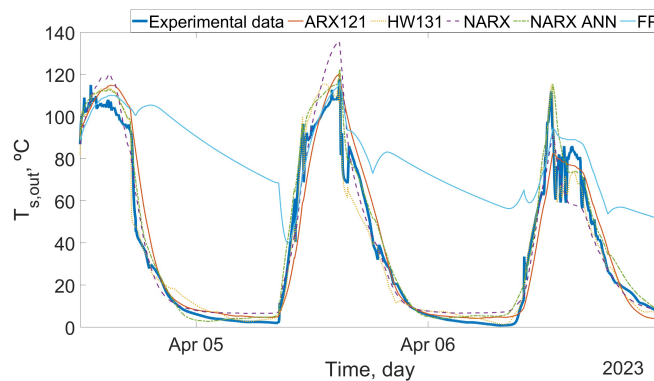
**Figure 10:** NARX ANN model validation.

Model	RMSE	RMSE Derivation	R ²	R ² Derivation
FP	50.0279	87.38%	0.3665	166.77%
ARX121	10.2262	38.25%	0.9322	4.88%
HW131	6.6065	4.47%	0.9690	0.90%
NARX	9.7090	38.25%	0.9439	3.58%
NARX ANN	6.3151	0.00%	0.9777	0.00%

Table 14Comparative model validation goodness. *RMSE* Derivation and *R*² Derivation refer to the deviation from the best model.

4.2. Model comparison

To compare the predictive capabilities of the ARX121, HW131, NARX, and NARX ANN models, their performances were analysed. In addition, the calibration of the FP model presented in Andrés-Mañas et al. (2020) was also performed using least squares to include the empirical model in the comparison. Table 14 provides a summary of their performance metrics obtained from the validation data, while Figure 11 displays the validation responses for each model.

**Figure 11:** Comparative validation model response.

Based on the results presented in Table 14, it can be concluded that the empirical models outperform the first principle model in terms of fit and capture the behaviour of the system when the pump is not working. Regarding data-driven models, the ARX model, which assumes linearity, showed the weakest performance. These results suggest that the non-linearity of the system is too high for a linear model to accurately depict its behaviour. Although the NARX model takes into account non-linear systems, it also performed poorly in comparison to the other models investigated in this paper. Nevertheless, the model that solely considers non-linear behaviour at the outlet of the system, i.e., the HW structure, accomplished a superior fit, with a value close to that of the NARX ANN model, which offers the best fit. Consequently, the NARX ANN model was selected as the best approach to model the system.

4.3. Simulation of an experimental plant for solar membrane distillation

An environmental simulation scenario, using models of the various equipment of the validated PSA facilities (see A), was used to replicate the behaviour of the system. The objective of this simulation is to examine how the model can be used to simulate the operation of the entire set-up and to test different scenarios efficiently without the necessity of conducting physical experiments. The NARX ANN model was chosen for the implementation of the solar field because of its superior fit as demonstrated in the previous section. Details of the remaining equipment model and operating modes are provided in Appendix A to Appendix D.

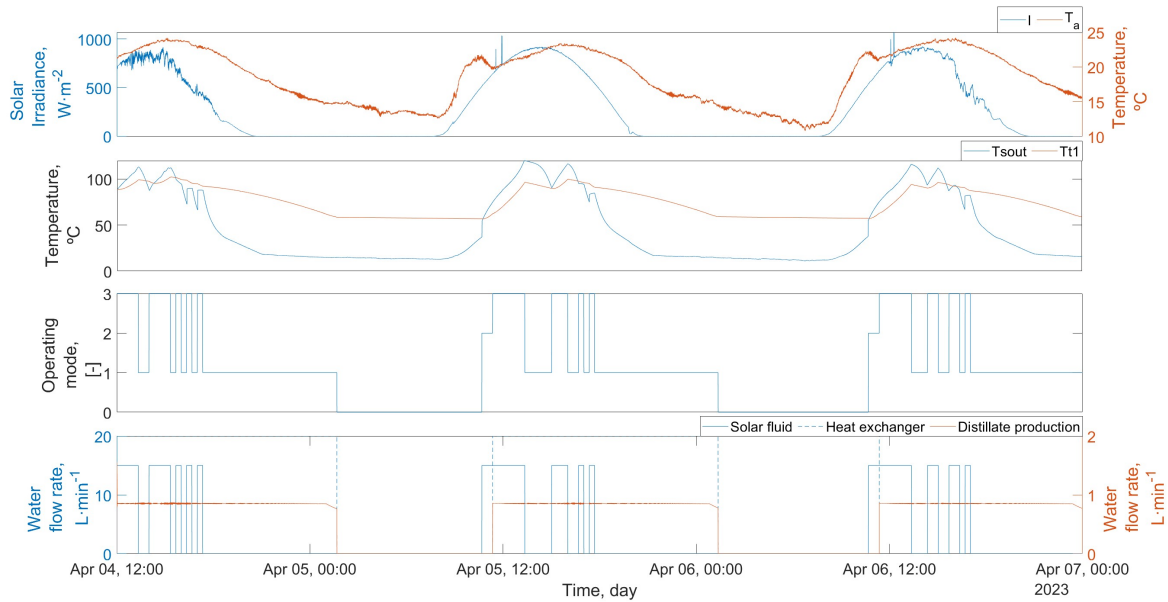


Figure 12: Operating mode simulated during the validation ambient conditions.

Figure 12 describes the dynamics obtained with the developed NARX ANN model, taking into account the simulation conditions, the operating modes, which are explained in C, and the model used for the remaining equipment integrated in the facilities, attached in the Appendix. To assess the integrity of the simulation, a comparison was made with the results obtained by Andrés-Mañas et al. (2020).

According to Andrés-Mañas et al. (2020), the typical procedure at the beginning of the day is to activate mode 2 to switch on the solar field pump, as the temperature at the upper part of the tank outlet tends to be lower than the reference temperature of the MD operation, appreciated at the beginning of the 5th and 6th of April in the simulation illustrated in Figure 12. As soon as the solar field heats the storage tank enough, that is, when the temperature is above the reference temperature of the MD operation, the MD unit starts to work. During mode 3, the solar field is used both to increase the thermal energy in the storage tank and to supply heat to the MD unit, as described by Andrés-Mañas et al. (2020). When the solar field pump is not working and the temperature in the storage tank is higher than the reference temperature, the MD unit works due to this heat and the temperature in the tank decreases, as shown in Figure 12. If the thermal energy in the tank is not enough and the ambient conditions allow the solar field to operate, the solar field

will work and the operating mode will change from 1 to 3; otherwise if the ambient conditions do not allow the solar field to operate, the operating mode will change from 1 to 0 as shown in Figure 12.

5. Conclusions

This paper presents a comparison study between data-driven models aimed at characterising the behaviour of a solar thermal field that powers a MD desalination system. The solar field includes a novel component: mirrors that increase the solar field's effective equivalent surface area. In addition, an extensive data set was utilised that spans 25 days with 1-second intervals. This comprehensive data allows the model to predict the transient state occurring when the pump starts the fluid flow through the solar field.

Data-driven models were trained using objective inputs such as irradiance, flow rate of the solar field pump, ambient temperature, and inlet temperature in the solar field. The temperature at the outlet of the solar field was taken into account as a predictor variable. Among the different methods tested, the NARX ANN model showed the best performance, achieving a strong correlation between the predicted and experimental responses with an R^2 value greater than 0.97 (with a difference of 0.9% for the HW model and greater than 3.5% for the rest method) and an RMSE value below 6.4 (with a difference of 4.47% for the HW model and 38% for the rest method). Moreover, the empirical model accurately depicted the transient temperature behaviour when the pump was inactive, as opposed to the FP model. This shows that utilising models that account for both active and inactive pump states can facilitate the creation of more effective optimisation strategies.

The applied methodology results in a compelling model of the solar field using empirical data, which can enhance the operation of MD with solar thermal fields. The model allows to investigate various advanced control strategies within a simulation framework to enhance the efficiency of solar MD system operation. By optimizing the operation of the solar thermal field in membrane distillation systems, distillate production can be increased, which results in a reduced payback period for the initial investment in the system.

A. Appendix A: Description of the system model

The experimental solar membrane distillation plant in the PSA described in Section 2 was modelled considering the process diagram shown in Figure 13, where the main process variables are shown.

The thermal storage tank was modelled through a multimodal stratified method (dividing the tank into two sections), leading to the following model as presented by Andrés-Mañas et al. (2020).

$$\rho \cdot V_1 \cdot \frac{\partial T_{i1}(t)}{\partial t} = m_s(t) \cdot T_{s,out}(t) + m_t(t) \cdot (1 - vp(t)) \cdot T_{i2}(t) - m_s(t) \cdot T_{i1}(t) - m_t(t) \cdot (1 - vp(t)) \cdot T_{i1}(t) - \frac{UA_1 \cdot (T_{i1}(t) - T_a(t))}{C_p}, \quad (11)$$

$$\rho \cdot V_2 \cdot \frac{\partial T_{i2}(t)}{\partial t} = m_s(t) \cdot T_{i1}(t) + m_t(t) \cdot (1 - vp(t)) \cdot T_{hx,out}(t) - m_s(t) \cdot T_{i2}(t) - m_t(t) \cdot (1 - vp(t)) \cdot T_{i2}(t) - \frac{UA_2 \cdot (T_{i2}(t) - T_a(t))}{C_p}, \quad (12)$$

where each section has a volume of 0.75 m^3 (V_i); $T_{s,out}$ is the outlet temperature of the solar field, T_{i1} represents the temperature at the outlet of the upper part of the tank, m_t denotes the flow rate in the heat exchanger and vp indicates the position of the 3-way mixing valve between the storage thermal tank and the heat exchanger (a value of 1 implies that all water circulates while 0 means that the water does not recirculate). T_{i2} represents the temperature at the outlet of the bottom section of the tank, while $T_{hx,out}$ denotes the temperature at the outlet of the hot channel of the heat exchanger. The coefficient of thermal losses in each section is UA_i , and its value is $3.6 \text{ J} \cdot \text{s}^{-1} \cdot ^\circ\text{C}^{-1}$ and $3.8 \text{ J} \cdot \text{s}^{-1} \cdot ^\circ\text{C}^{-1}$, respectively.

The behaviour of the heat exchanger can be described as presented by Gil, Álvarez, Roca, Sánchez-Molina, Berenguel, and Rodríguez (2019). The ensuing equations depict the model employed:

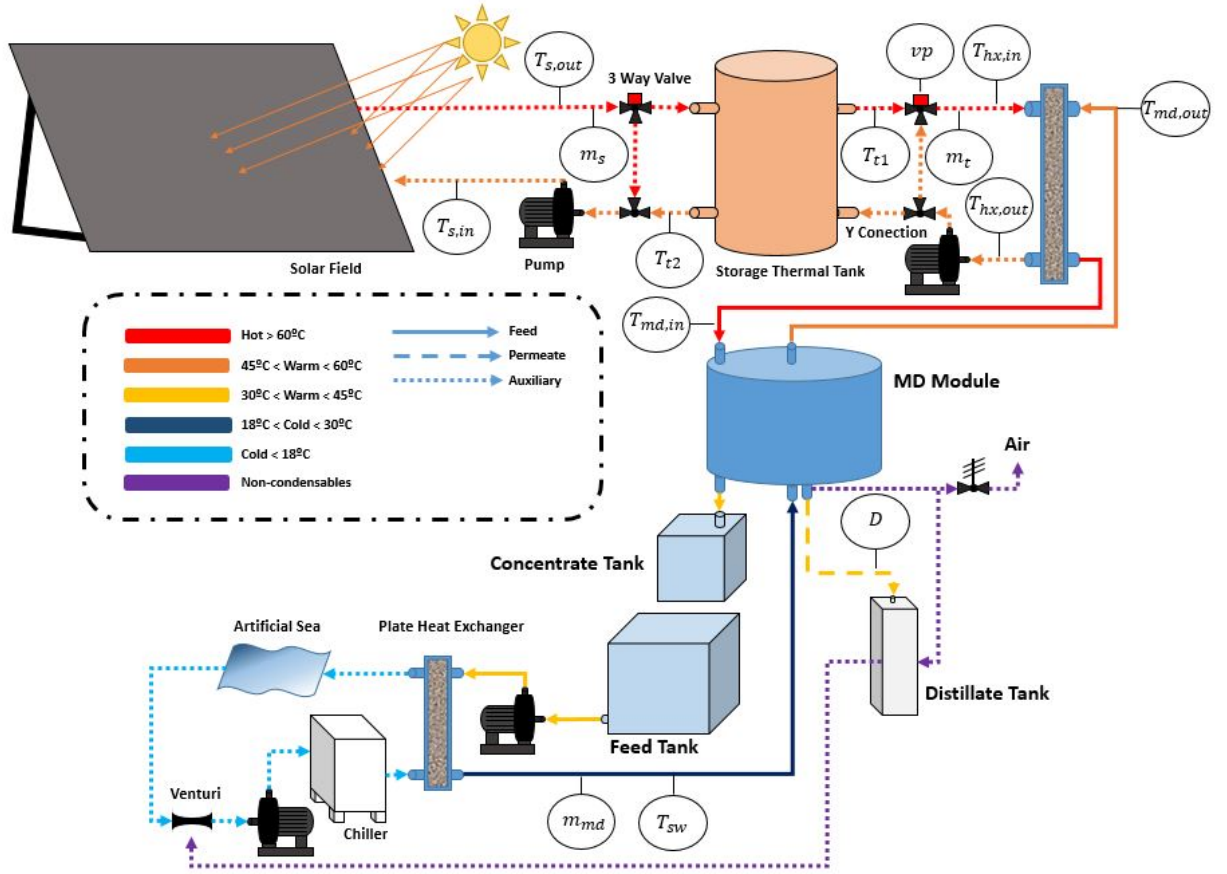


Figure 13: Schematic diagram of the PSA plant process.

$$T_{hx,out}(t) = T_{hx,in}(t) - \mu_1 \cdot (T_{hx,in}(t) - T_{md,out}(t)), \quad (13)$$

$$T_{md,in}(t) = T_{md,out}(t) + \mu_2 \cdot (T_{hx,in}(t) - T_{hx,out}(t)), \quad (14)$$

$$\mu_1 = \frac{1 - e^\theta}{1 - \frac{m_t(t) \cdot C_{p_t}}{m_{md}(t) \cdot C_{p_{md}}} \cdot e^\theta}, \quad (15)$$

$$\mu_2 = \frac{m_t(t) \cdot C_{p_t}}{m_{md}(t) \cdot C_{p_{md}}}, \quad (16)$$

$$\theta = U \cdot A_{he} \cdot \left(\frac{1}{m_t(t) \cdot C_{p_t}} - \frac{1}{m_{md}(t) \cdot C_{p_{md}}} \right), \quad (17)$$

where $T_{hx,in}$ is the temperature in the hot channel of the heat exchanger inlet, $T_{md,out}$ is the temperature in the cold channel of the heat exchanger inlet, $T_{md,in}$ is the temperature in the cold channel of the heat exchanger outlet, while θ ,

μ_1 and μ_2 are additional factors used to calculate the temperature of the heat exchanger outlet. U represents the heat exchanger transfer coefficient, with a value of $15042 \text{ W} \cdot \text{m}^{-2} \cdot \text{K}^{-1}$, while A_{he} denotes the surface area of the heat exchanger, which has a value of 1.65 m^2 .

The MD unit utilised was the AS26 module modelled by Andrés-Mañas, Requena, Ruiz-Aguirre, and Zaragoza (2022):

$$P_{flux}(t) = a_0 + a_1 \cdot T_{md,in}(t) + a_2 \cdot m_{md}(t) + a_3 \cdot T_{sw}(t) + a_4 \cdot T_{md,in}(t) \cdot T_{sw}(t) + a_5 \cdot T_{md,in}(t) \cdot m_{md}(t) + a_6 \cdot T_{sw}(t) \cdot m_{md}(t) + a_7 \cdot T_{md,in}(t)^2 + a_8 \cdot m_{md}(t)^2 + a_9 \cdot T_{sw}(t)^2, \quad (18)$$

$$T_{md,out}(t) = a_0 + a_1 \cdot T_{md,in}(t) + a_2 \cdot m_{md}(t) + a_3 \cdot T_{sw}(t) + a_4 \cdot T_{md,in}(t) \cdot T_{sw}(t) + a_5 \cdot T_{md,in}(t) \cdot m_{md}(t) + a_6 \cdot T_{sw}(t) \cdot m_{md}(t) + a_7 \cdot T_{md,in}(t)^2 + a_8 \cdot m_{md}(t)^2 + a_9 \cdot T_{sw}(t)^2, \quad (19)$$

where P_{flux} represents the permeate flux achieved in the module, $T_{md,out}$ denotes the temperature in the cold channel of the heat exchanger inlet, and a_i corresponds to the adjusted coefficients of the model provided by Andrés-Mañas et al. (2022) for each model. Furthermore, $T_{md,in}$ refers to the temperature in the cold channel of the heat exchanger outlet, and T_{sw} is representative of the temperature of the seawater. Meanwhile, m_{md} indicates the flow rate of MD.

B. Appendix B: Description of the temperature control employed in the simulation

A control method was incorporated into the model in order to regulate the temperature at the outlet of the heat exchanger's cold channel. The simulation used a set point of 70°C . To regulate the temperature, a three-way mixing valve was used to circulate the hot channel outlet of the heat exchanger and mix it with the fluid sourced from the upper section of the thermal storage tank. The valve position was gauged as per Figure 14.

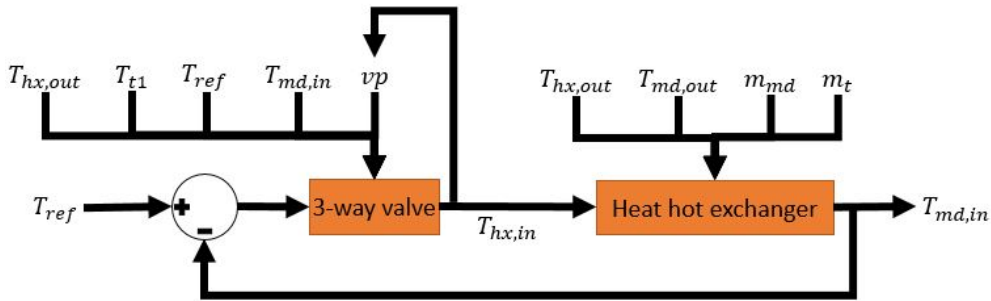


Figure 14: Scheme of the temperature control at the outlet of the cold channel of the heat exchanger.

The equations that determine the value of vp and $T_{hx,in}$ are as follows:

- vp :

- If $T_{md,in}(t) > T_{ref} + 0.2^\circ\text{C}$:

$$vp(t) = vp(t-1) + 0.01 \quad (20)$$

- If $T_{md,in}(t) < T_{ref} - 0.2^\circ\text{C}$:

$$vp(t) = vp(t-1) - 0.01 \quad (21)$$

- If $|T_{md,in}(t) - T_{ref}| \leq 0.2^\circ\text{C}$:

$$vp(t) = vp(t-1) \quad (22)$$

- $T_{hx,in}$:

– If $mode(t) = 0$ or $mode(t) = 2$:

$$T_{hx,in}(t+1) = T_{t1}(t) \quad (23)$$

– If $mode(t) = 1$ or $mode(t) = 3$:

$$T_{hx,in}(t+1) = T_{t1}(t) \cdot (1 - vp(t)) + T_{hx,out} \cdot vp(t) \quad (24)$$

C. Appendix C: Description of the operating modes used in the simulation

A genuine simulation was conducted using four distinct modes of operation, identical to those described by Andrés-Mañas et al. (2020). However, the circumstances to alter the selected mode differ. These circumstances are as follows:

- **Condition 1:** $I > 500 \text{ W} \cdot \text{m}^{-2}$ AND $T_{s,out} < 100 \text{ }^\circ\text{C}$.
- **Condition 2:** $I < 500 \text{ W} \cdot \text{m}^{-2}$ OR $T_{s,out} > 110 \text{ }^\circ\text{C}$ OR $(T_{s,out} < T_{t1}$ AND operation mode = 3).
- **Condition 3:** $T_{t1} > T_{ref} + 1$
- **Condition 4:** $T_{t1} < T_{ref}$

Here T_{ref} represents the reference temperature in the cold channel of the heat exchanger outlet. Figure 15 displays the correlation between the various operating modes, where:

- **Mode 0:** Stop mode is characterized by the inactivity of all pumps.
- **Mode 1:** In this mode, only the MD system is operational using the thermal energy stored in the tank. Both the MD and heat exchanger pumps are active.
- **Mode 2:** In solar field mode, the solar thermal field alone is active, and the generated heat is stored in the tank. The pump of the solar thermal field is the only one that is turned on.
- **Mode 3:** Run mode. All pumps are running.

D. Appendix D: Description of the initial values

With the aim of performing a simulation of three days, the following assumptions were made:

- The recorded data provided information on the solar irradiance and ambient temperature.
- The temperature of the seawater was acquired for the location of Almería in April, where it measured 16.7°C , according to Andrés-Mañas et al. (2020).
- The solar fluid flow rate was determined to be 15 litres per minute when the solar field was in operation, as stated in Andrés-Mañas et al. (2020).
- T_{ref} was 70°C .
- m_t equated to 10 litres per minute while m_{md} was 12.5 litres per minute.
- The simulation was initiated under the following conditions: $T_{t2}(0) = T_{s,in}(0)$ and $T_{t1}(0) = T_{s,out}(0)$.
- The data has a time sample interval of 1 second.
- The temperature at the solar field's inlet matches that of the lower part of the storage tank's outlet during modes 2 or 3: $T_{s,in}(t) = T_{t2}(t)$. In the other operating modes, the temperature decrease until $T_{s,in}(t) = T_a(t)$.

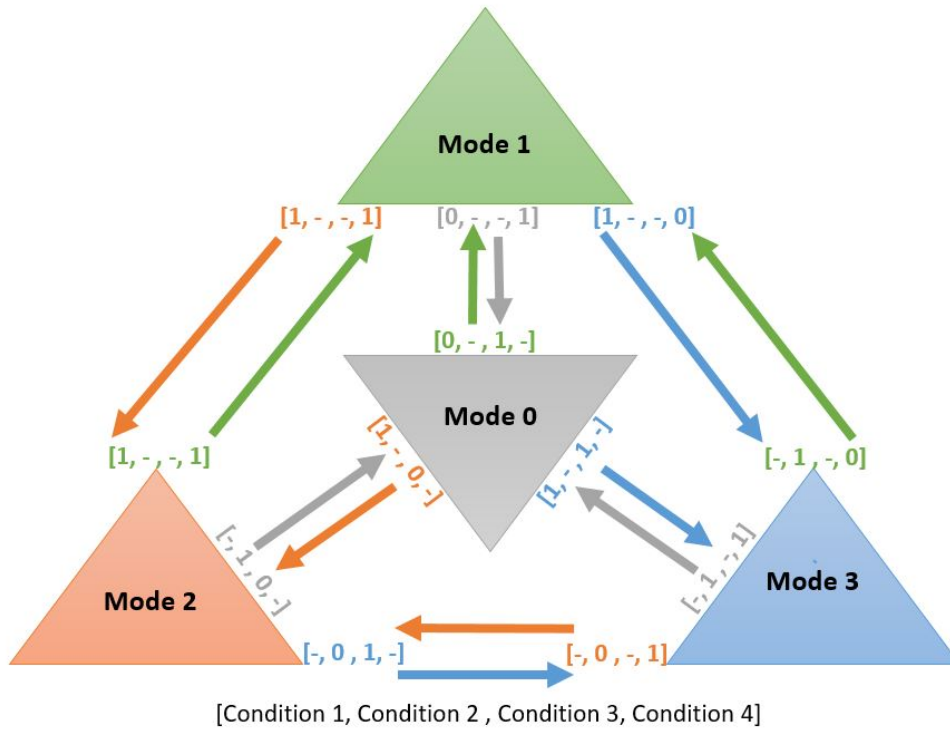


Figure 15: Relationship between different modes of operation.

References

- C. Fritzmann, J. Löwenberg, T. Wintgens, T. Melin, State-of-the-art of reverse osmosis desalination, *Desalination* 216 (2007) 1–76.
- M. A. Shannon, P. W. Bohn, M. Elimelech, J. G. Georgiadis, B. J. Mariñas, A. M. Mayes, Science and technology for water purification in the coming decades, *Nature* 452 (2008) 301–310.
- H. Sharon, K. Reddy, A review of solar energy driven desalination technologies, *Renewable and Sustainable Energy Reviews* 41 (2015) 1080–1118.
- A. D. Khawaji, I. K. Kutubkhanah, J.-M. Wie, Advances in seawater desalination technologies, *Desalination* 221 (2008) 47–69.
- A. Al-Karaghoul, L. L. Kazmerski, Energy consumption and water production cost of conventional and renewable-energy-powered desalination processes, *Renewable and Sustainable Energy Reviews* 24 (2013) 343–356.
- L. F. Greenlee, D. F. Lawler, B. D. Freeman, B. Marrot, P. Moulin, Reverse osmosis desalination: water sources, technology, and today's challenges, *Water research* 43 (2009) 2317–2348.
- N. Ghaffour, J. Bundschuh, H. Mahmoudi, M. F. Goosen, Renewable energy-driven desalination technologies: A comprehensive review on challenges and potential applications of integrated systems, *Desalination* 356 (2015) 94–114.
- A. Pugsley, A. Zacharopoulos, J. D. Mondol, M. Smyth, Global applicability of solar desalination, *Renewable energy* 88 (2016) 200–219.
- H. M. Qiblawey, F. Banat, Solar thermal desalination technologies, *Desalination* 220 (2008) 633–644.
- A. Chorak, P. Palenzuela, D.-C. Alarcón-Padilla, A. B. Abdellah, Experimental characterization of a multi-effect distillation system coupled to a flat plate solar collector field: Empirical correlations, *Applied Thermal Engineering* 120 (2017) 298–313.
- D. S. Ayou, G. Zaragoza, A. Coronas, Small-scale renewable polygeneration system for off-grid applications: Desalination, power generation and space cooling, *Applied Thermal Engineering* 182 (2021) 116112.
- J. Zhou, X. Zhang, B. Sun, W. Su, Performance analysis of solar vacuum membrane distillation regeneration, *Applied Thermal Engineering* 144 (2018) 571–582.
- F. Diez, L. Navas-Gracia, A. Martínez-Rodríguez, A. Correa-Guimaraes, L. Chico-Santamarta, Modelling of a flat-plate solar collector using artificial neural networks for different working fluid (water) flow rates, *Solar Energy* 188 (2019) 1320–1331.
- J. Deng, Y. Xu, X. Yang, A dynamic thermal performance model for flat-plate solar collectors based on the thermal inertia correction of the steady-state test method, *Renewable Energy* 76 (2015) 679–686.
- R. Kicsiny, Improved multiple linear regression based models for solar collectors, *Renewable Energy* 91 (2016) 224–232.
- J. D. Gil, L. Roca, G. Zaragoza, M. Berenguel, A feedback control system with reference governor for a solar membrane distillation pilot facility, *Renewable Energy* 120 (2018) 536–549.
- M. Pasamontes, J. Álvarez, J. Guzman, M. Berenguel, E. F. Camacho, Hybrid modeling of a solar-thermal heating facility, *Solar Energy* 97 (2013) 577–590.

- J. Andrés-Mañas, L. Roca, A. Ruiz-Aguirre, F. Ación, J. D. Gil, G. Zaragoza, Application of solar energy to seawater desalination in a pilot system based on vacuum multi-effect membrane distillation, *Applied Energy* 258 (2020) 114068.
- A. Kabeel, M. Abdelgaied, E. M. El-Said, Study of a solar-driven membrane distillation system: Evaporative cooling effect on performance enhancement, *Renewable energy* 106 (2017) 192–200.
- A. Shafieian, M. Khiadani, A novel solar-driven direct contact membrane-based water desalination system, *Energy Conversion and Management* 199 (2019) 112055.
- K. Zarzoum, K. Zhani, H. B. Bacha, J. Koschikowski, Experimental parametric study of membrane distillation unit using solar energy, *Solar Energy* 188 (2019) 1274–1282.
- A. K. Palit, D. Popovic, Computational intelligence in time series forecasting: theory and engineering applications, Springer Science & Business Media, 2006.
- R. Porrazzo, A. Cipollina, M. Galluzzo, G. Micale, A neural network-based optimizing control system for a seawater-desalination solar-powered membrane distillation unit, *Computers & Chemical Engineering* 54 (2013) 79–96.
- P. Behnam, M. Zargar, A. Shafieian, A. Razmjou, M. Khiadani, Harnessing the power of neural networks for the investigation of solar-driven membrane distillation systems under the dynamic operation mode, *Solar Energy* 261 (2023) 63–82.
- I. Farkas, P. Geczy-Vig, Neural network modelling of flat-plate solar collectors, *Computers and electronics in Agriculture* 40 (2003) 87–102.
- G. Ampuño, L. Roca, M. Berenguel, J. D. Gil, M. Pérez, J. E. Normey-Rico, Modeling and simulation of a solar field based on flat-plate collectors, *Solar Energy* 170 (2018) 369–378.
- I. M. Pataro, J. D. Gil, J. L. Guzmán, M. Berenguel, J. M. Lemos, Hierarchical control based on a hybrid nonlinear predictive strategy for a solar-powered absorption machine facility, *Energy* 271 (2023) 126964.
- Mathworks, Estimate parameters of arx, arix, ar, or ari model, Available from: <https://es.mathworks.com/help/ident/ref/arx.html>, 2023. Accessed on: October, 26, 2023.
- H. D. Karaca, G. Özen, C. Kasnakoğlu, Nonlinear modelling and control of the flow over aerofoils using cfd simulations, *Simulation Modelling Practice and Theory* 67 (2016) 29–43.
- S. Koroglu, P. Sergeant, N. Umurkan, Comparison of analytical, finite element and neural network methods to study magnetic shielding, *Simulation Modelling Practice and Theory* 18 (2010) 206–216.
- Mathworks, Sigmoid network function for nonlinear arx and hammerstein-wiener models, Available from: <https://es.mathworks.com/help/ident/ref/idsigmoidnetwork.html>, 2023a. Accessed on: October, 26, 2023.
- Mathworks, Estimate parameters of nonlinear arx model, Available from: <https://es.mathworks.com/help/ident/ref/nlarx.html>, 2023b. Accessed on: October, 26, 2023.
- A. Fernández-Ares, A. Mora, S. M. Odeh, P. Garcia-Sanchez, M. G. Arenas, Wireless monitoring and tracking system for vehicles: A study case in an urban scenario, *Simulation modelling practice and theory* 73 (2017) 22–42.
- A. Di Piazza, M. C. Di Piazza, G. Vitale, Solar and wind forecasting by narx neural networks, *Renewable Energy and Environmental Sustainability* 1 (2016) 39.
- J. D. Gil, J. Álvarez, L. Roca, J. Sánchez-Molina, M. Berenguel, F. Rodríguez, Optimal thermal energy management of a distributed energy system comprising a solar membrane distillation plant and a greenhouse, *Energy Conversion and Management* 198 (2019) 111791.
- J. Andrés-Mañas, I. Requena, A. Ruiz-Aguirre, G. Zaragoza, Performance modelling and optimization of three vacuum-enhanced membrane distillation modules for upscaled solar seawater desalination, *Separation and Purification Technology* 287 (2022) 120396.