



Research article

A complementary approach of response surface methodology and an artificial neural network for the optimization and prediction of low salinity reverse osmosis performance



Ryan Brooke^a, Linhua Fan^a, Mohamed Khayet^b, Xu Wang^{a,*}

^a School of Engineering, RMIT University, Australia

^b Department of Structure of Matter, Thermal Physics, and Electronics Faculty of Physics, University Complutense of Madrid, Spain

HIGHLIGHTS

- A complementary approach of response surface methodology (RSM) modelling and an artificial neural network (ANN) prediction model
- Modelling and optimizing the reverse osmosis (RO) desalination process of the low salinity water (TDS = 500–5000 mg/L)
- Sensitivity analysis of the RO desalination process parameters to the response target
- Discussion on the merits of each modelling methodology

ARTICLE INFO

Keywords:

Reverse osmosis
Response surface methodology
Artificial neural network
Low salinity

ABSTRACT

The treatment of saline water sources by reverse osmosis (RO) is being utilized increasingly to address water shortages around the world. The application of RO is energy-intensive; therefore, plant and process optimization are crucial. The desalination of low salinity water sources with total dissolved solids (TDS) of <5000 mg/L is less energy intensive than the desalination of highly saline seawater and brackish water. A gap exists in optimization studies on lower salinity water (TDS = 500–5000 mg/L). The novelty of the study is the development of a complementary approach using response surface methodology (RSM) and an artificial neural network (ANN) for performance modelling, optimization, and prediction of RO desalination of low salinity water. Feed water salinity, pressure, and temperature were controlled variables to model the performance of the RO system. A performance index incorporating salt rejection efficiency and permeate flux was used as the response target of the system. The optimal parameter combination within their modelled range for the best performance index occurred near the highest pressure input of 150.57 psi, at the temperature of 38.8 °C, and at the lowest feed salt concentration of 577 mg/L. Both the RSM and ANN models demonstrated high validity. The RSM and ANN showed R^2 values of 0.99 each and with a root mean square error of 2.41 and 5.85 respectively. The RSM showed a small benefit in model accuracy over the ANN, but the ANN has the benefit of not requiring the central composite design before experimentation and being a continuously improving prediction method as more data becomes available. Further applications of the optimization and modelling approach can be applied to RO system optimization considering membrane types and additional feedwater characteristics.

1. Introduction

Fresh drinkable water is an essential human need. As a result of climate change, droughts are getting worse and leading to an increased shortage of fresh water supply for many regions around the world. Currently, there is a high imbalance between clean water demand and supply, and around one-quarter of the global population is facing an economic water shortage [1].

One solution to this supply issue is the use of reverse osmosis (RO) to desalinate seawater, brackish groundwater, or other saline sources. As many cities are positioned near oceans and rivers, desalination may be the future of freshwater production as an increasing number of RO-based desalination plants are being planned, commissioned, and operated around the globe.

* Corresponding author.

E-mail address: xu.wang@rmit.edu.au (X. Wang).

<https://doi.org/10.1016/j.heliyon.2022.e10692>

Received 14 May 2022; Received in revised form 15 August 2022; Accepted 14 September 2022

2405-8440/© 2022 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Potable fresh water is considered to have total dissolved solids (TDS) of less than 500 mg/L [2]. All desalination techniques are energy intensive. RO is considered to have lower energy consumption per litre of fresh water produced when compared with other desalination methods such as thermal processes and electrodialysis [3]. As a result of the energy-intensive nature of RO-based desalination processes, system optimization leads to greater cost-effectiveness.

Numerous publications have researched the optimisation of brackish water reverse osmosis (BWRO) (5000–15,000 mg/L TDS) [4, 5, 6, 7, 8] and salt water reverse osmosis (SWRO) (15,000 mg/L–35,000 mg/L) [4, 5, 9, 10, 11]. However, it is much less energy-intensive to perform reverse osmosis on low salinity water between 500 and 5000 mg/L as there is less osmotic pressure to overcome. Sources of low salinity water would include rivers, groundwater sources, and some wastewater streams. This paper outlines the optimisation of low salinity RO. Research by Khayet et al. [4] found some differences in the response to changing input parameters of SWRO and BWRO, which required unique models for each. BWRO was found to be most sensitive to the feed inlet pressure whereas SWRO was more sensitive to the feed TDS concentration. Li [5] optimized plant configurations for specific energy consumption of BWRO and SWRO, further demonstrating different salinity ranges of BWRO and SWRO cannot be optimized with the same plant settings. It is with this knowledge that lower salinity RO is modelled to indicate the most sensitive input parameters on the performance response.

The present study aims to demonstrate a performance modelling, optimization, and prediction method for a small-scale reverse osmosis system desalinating low salinity feedwater. The motivation of this study is to enable the understanding of the magnitude of input parameters' effects on performance, while also using the same data to produce a continuously improving prediction model. As the RO performance output is a function of the feed solution TDS concentration, the pressure applied across the membrane, and the temperature of the feed solution, they were selected as the input parameters in this study. A simple modelling and optimization method was designed by using response surface methodology (RSM), a statistical method to model the interaction between the input parameters and predict the performance of the RO system. The use of an artificial neural network (ANN) for modelling the RO process together with RSM was explored for comparison, but also for where these methods complement each other. As well-established techniques, RSM and ANNs have often been compared for accuracy by researchers [12, 13, 14]. Few examples exist in using these methods in parallel [15, 16], taking advantage of the merits of each of the methods. RSM can produce a mathematical model with good insights into how each parameter affects the system response. ANNs are somewhat of a 'black-box' in their modelling and do not provide a numerical model, reducing the system insight available. However, ANN does not require a pre-experimental plan and often becomes a better predictor as more data become available. Therefore, by using both RSM and ANN, process insight becomes available as well as a continuously improving prediction model.

2. Experimental design

A small-scale RO system for desalinating low salinity water was developed and modelled to optimize its design and operation. An ideal membrane would remove all salt from the feed water while allowing a large amount of water to move through the membrane. This leads to the two performance indicators of permeate flux and salt rejection efficiency being measured.

Salt rejection efficiency is measured as a percentage of salt removed across the RO membrane and is calculated by:

$$\text{Rejection \%} = \left(\frac{C - C_p}{C} \right) * 100 \quad (1)$$

where C is the TDS (mg/L) concentration of the feed and C_p is the TDS (mg/L) concentration of the permeate.

The permeate flux is the permeate flow rate across the membrane per unit area and is calculated by:

$$\text{Permeate Flux (L/m}^2\text{s)} = \frac{\text{Permeate Flowrate}}{\text{Membrane Surface Area}} \quad (2)$$

The performance index (Y) as adopted from ref [4] is the product of the salt rejection efficiency given in Eq. (1) and permeate flux given in Eq. (2) and is calculated by:

$$Y(\text{L/m}^2\text{s}) = \text{Permeate Flux} * \frac{\text{Rejection \%}}{100} \quad (3)$$

The membrane used for the study was the Dow Filmtec sw30-2521 which was 2.5-inches in diameter, 21-inches long, and with a surface area of 1.2 m² [17]. The thin-film composite reverse osmosis membrane is capable of filtering high salinity seawater. Using this membrane allows this work to be compared with the previous work by Khayet et al. [4].

The controlled inlet feed parameters were temperature, pressure, and salinity. The temperature was adjusted using a heating element within the feed solution. A high-pressure (diaphragm) pump (HPP) was fed by a low-pressure pump (LPP) to increase head pressure. Before entering the HPP, feed water moves through a screen filter to ensure the removal of particulate matter which could cause damage or clogging of the RO membrane system. The HPP then increased the pressure of the RO membrane inlet. The feed flow rate is a function of the backpressure and was not adjustable due to the use of the diaphragm pump. From the literature review, it was shown that adjusting the feed flow rate will have little effect on the performance of the RO [4]. According to the manufacturer's recommendation for the long-term performance of the Dow Filmtec sw30-2521, the minimum feed flow rate is 0.2 m³/h [18]. In this study, the feed flow rate was able to be achieved and set by tuning the backpressure valve setting. A valve on the retentate outlet produces back pressure to adjust the pressure across the membrane. The pressure is measured using a manometer on this retentate outlet. The feed solution TDS concentration was adjusted by adding NaCl to the feed tank and measured using a water quality tester (EZ-9908) TDS meter where the feed solution TDS concentration was reflected by the variation of its electrical conductivity. The feed solution TDS concentration was adjusted to the desired levels. Figure 1 depicts the designed configuration of the small-scale reverse osmosis plant. Figure 2 shows the actual experimental setup.

The housing and pump were sourced from an off-the-shelf seawater reverse osmosis (SWRO) unit, designed for use onboard seafaring boats. The whole unit includes a pressure exchanger for energy recovery; however, the energy recovery unit was not included in the setup as this energy recovery unit is designed for the much higher-pressure application of SWRO. The feed pump has the capabilities of producing feed pressure up to 155 psi which was used as the upper limit of the working range for this experiment. The maximum temperature that can be tested without the risk of damaging the membrane was set by the manufacturer as 45 °C. The feed concentration that the system is capable of desalinating is a function of the pressure and temperature. The working range of the feed concentration was found to be approximately between 0 and 4500 mg/L TDS through testing the capabilities of the system input parameters.

The three input control variables and corresponding output performance indices calculated by Eq. (3) were modelled using response surface methodology (RSM). In this case, the output selected is the performance index (Y), which is the response of the system from the series of designed experimental runs. Central composite design (CCD) was used for the design of the experiment, which is an efficient method of changing multiple input parameters and finding their response as an output. CCD can avoid a great deal of trial and error and produce effective results from minimal experimental runs, which saves experimental time and costs. By using CCD to design the experimental runs, the RSM model can be developed from the input parameters and the experimental

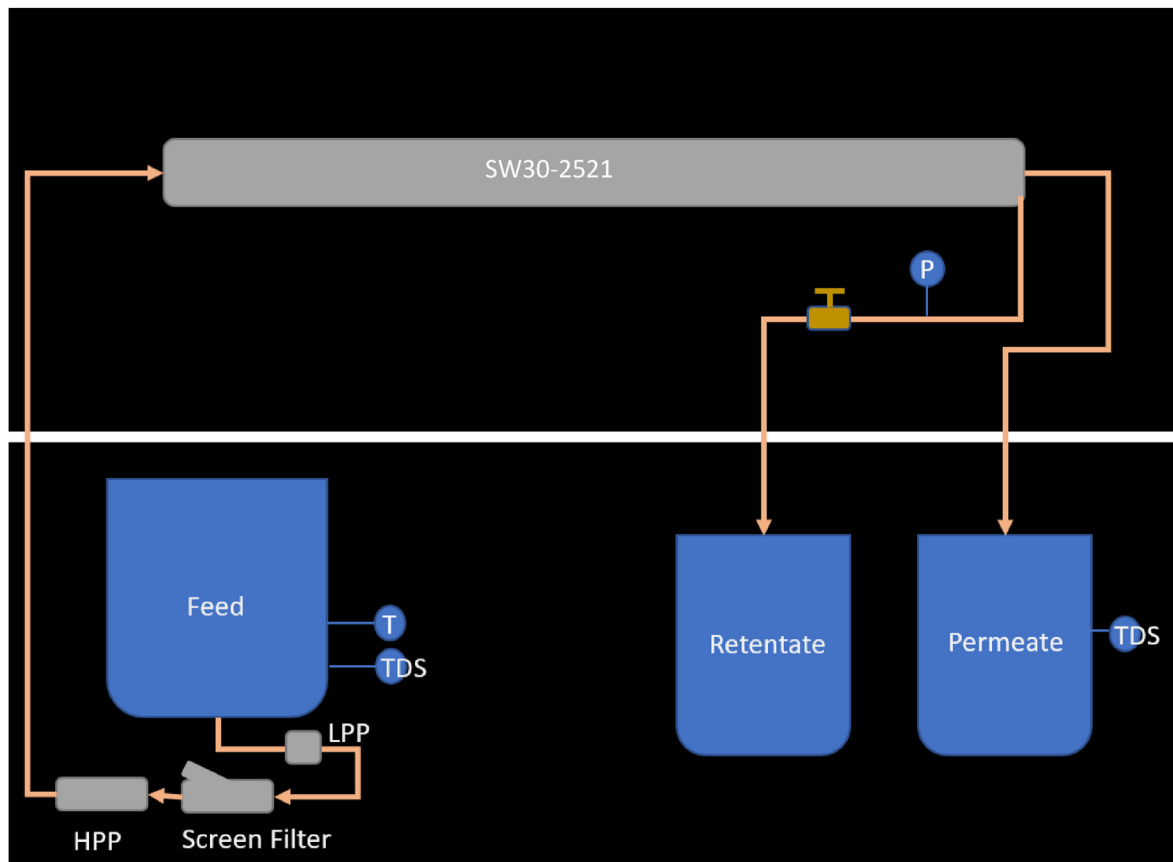


Figure 1. Design configuration for a small-scale RO testing system.

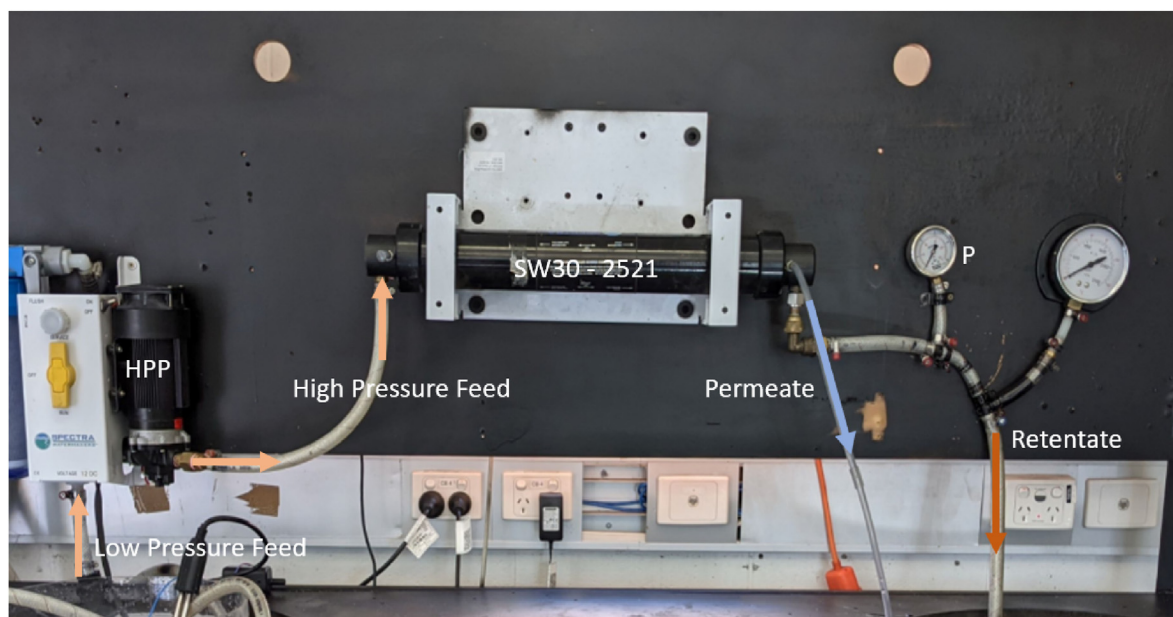


Figure 2. RO module, feed, and outlet lines in the actual built small-scale RO system.

output response. Different CCD methods can be used, however in this case an orthogonal CCD design was implemented, as the orthogonal CCD design is more centrally accurate than the other CCD methods [19]. For a CCD of three input parameters, at least 15 experimental runs are required, but more centre point experiments are used to assess the system drift. In this case, 2 centre points (0, 0, 0) were used in the experiment,

and a total of 16 experimental runs were undertaken. The CCD theory sets normalized dimensionless parameters (x_1, x_2, x_3) within the minimum and maximum range limits, then extended past this range for each of the 3 parameters to the star points. In this case, the star points are at ± 1.287 , as dictated by the CCD theory. The normalized dimensionless input parameters are shown in Table 1. The range (dz) and centre point

(z_0) are given by Eqs. (4) and (5). The dimensional minimum and maximum parameter range limits along with the calculated dz and z_0 are shown in Table 2. Finally, the experimental unnormalized dimensional run parameters (z_1, z_2, z_3) can be calculated using Eq. (6) and are listed in Table 1.

$$dz = \frac{z_{\max} - z_{\min}}{2} \quad (4)$$

$$z_0 = \frac{z_{\max} + z_{\min}}{2} \quad (5)$$

where z_{\max} is the maximum limit value of the range and z_{\min} is the minimum limit value of the range.

$$z_i = x_i * dz + z_0 \quad (6)$$

where z_i is the unnormalized dimensional parameter value and x_i is the normalized dimensionless parameter value ($i = 1, 2, 3$). All equipment and facilities were calibrated over the parameter ranges before the experimental runs were conducted.

3. Results and discussion

The CCD input parameters and the experimental response of the measured performance index (Y) are listed in Table 3. The values were input into Design Expert 12 software to develop the RSM model, with a polynomial regression relationship. The RSM model for Performance Index in L/m^2s is given by:

$$\begin{aligned} \text{Performance Index (Y)} = & 363.40 - 178.37 * x_1 + 66.43 * x_2 + 105.97 * x_3 - 69.31 * x_1 x_2 \\ & - 20.78 * x_1 x_3 + 19.93 * x_2 x_3 - 14.47 x_1^2 - 14.86 x_3^2 - 19.08 * x_1 x_2 x_3 + 14.48 x_1^2 x_2 \\ & + 10.03 x_1^2 x_3 - 14.06 x_1 x_3^2 \end{aligned} \quad (7)$$

where x_1 is the normalized feed concentration, x_2 is the normalized temperature and x_3 is the normalized feed pressure.

From the RSM model stated in Eq. (7), several significant findings can be extracted. The absolute value of the coefficient of x_1 (−178.37) is the largest among those of the three input variables. This corresponds to the feed concentration (x_1) having the greatest effect on the response, Y . Being a negative value indicates increasing the feed concentration will

Table 2. Minimum and maximum parameter values, dz and z_0 .

	min	max	dz	z_0
z_1 Feed TDS Concentration (mg/L)	1000	4000	1500	2500
z_2 Feed Temperature (°C)	20	37.5	8.75	28.75
z_3 Feed Pressure (psi)	95	145	25	120

decrease the response performance index Y . This is intuitively supported by the theory of reverse osmosis, as the feed concentration increases, the required pressure to overcome osmotic pressure for reverse osmosis also increases for maintaining a constant permeate flux. Generally, the feed concentration is relatively constant when it is collected from a stable water source. Therefore, for more cost-effectiveness and better RO performance, a lower TDS source is ideal. As a result, low salinity groundwater, river water, treated wastewater, or their mixture with seawater would all be considered beneficial.

In this experiment, the feed pressure of the system was controllable, similar to a real-world RO plant. Eq. (7) shows the absolute value of the coefficient of x_3 (105.97) has the second largest effect on the response performance index output. As the feed pressure of the system (x_3) increases, the performance index Y increases. Once the osmotic pressure is overcome, the increased pressure difference drives more fresh water through the membrane. This mechanism requires several considerations in a real-world application, including the ability of the membrane to mechanically withstand the pressure. In this case, the Dow SW30-2521 membrane has a working pressure of up to 1000 psi [17]. Also, energy cost, brine management, and membrane fouling/scaling are some of the other considerations in determining the desired operating pressure of the RO systems.

Table 1. Experimental run dimensionless input parameters (x_1, x_2, x_3) and corresponding dimensioned parameters (z_1, z_2, z_3).

Experimental Run	x_1	x_2	x_3	z_1 Feed Concentration(mg/L)	z_2 Temperature (°C)	z_3 Feed Pressure (psi)
1	−1	−1	−1	1000	20	95
2	1	−1	−1	4000	20	95
3	−1	1	−1	1000	37.5	95
4	1	1	−1	4000	37.5	95
5	−1	−1	1	1000	20	145
6	1	−1	1	4000	20	145
7	−1	1	1	1000	37.5	145
8	1	1	1	4000	37.5	145
9	−1.287	0	0	569	28.75	120
10	1.287	0	0	4431	28.75	120
11	0	−1.287	0	2500	17.49	120
12	0	1.287	0	2500	40.01	120
13	0	0	−1.287	2500	28.75	87.82
14	0	0	1.287	2500	28.75	152.18
15	0	0	0	2500	28.75	120
16	0	0	0	2500	28.75	120

Table 3. The experimental results and calculated Performance Index using Eq. (3).

Experimental Run	Feed TDS Concentration (mg/L) (z1)	Temperature (°C) (z2)	Feed Pressure (psi) (z3)	Permeate Flux (L/m ² s) × 10 ⁻⁵	Rejection Efficiency (%)	Performance Index (Y) (L/m ² s) × 10 ⁻⁵
1	1000	20	95	289.81	96.47	279.57
2	4000	20	95	51.85	70.85	36.74
3	1000	37.5	95	515.74	97.33	501.99
4	4000	37.5	95	71.3	81.67	58.23
5	1000	20	145	480.56	98.87	475.11
6	4000	20	145	242.59	92.94	225.47
7	1000	37.5	145	871.3	97.97	853.58
8	4000	37.5	145	276.85	90.43	250.37
9	569	28.75	120	593.52	95.25	565.36
10	4431	28.75	120	124.07	85.56	106.16
11	2500	17.49	120	288.89	96.79	279.61
12	2500	40.01	120	472.22	95.43	450.63
13	2500	28.75	88	212.04	93.93	199.17
14	2500	28.75	152	481.48	98.03	471.98
15	2500	28.75	120	371.3	96.67	358.92
16	2500	28.75	120	377.78	97.05	366.65

greater output response. Higher feed temperatures will also increase the permeate flux due to physical changes in the membrane such as changes in pore size [21]. However, the feed temperature is limited by the manufacturer's specifications of the membrane used. The upper-temperature limit is 45 °C in this case [17].

In Eq. (7), it is seen that the absolute value of the coefficient of x_1x_2 (−69.31) is the largest among those of the two coupling input variables. This means that the interaction between the feed concentration and temperature (x_1x_2) also has a large influence on the response output. The negative sign indicates that by increasing the coupling effect of the input variables x_1 and x_2 , the response output will decrease. This finding indicates although temperature itself has only a small effect on performance, due to the coupling effect it could still be an important consideration for greater performance, especially for feed sources with variable temperatures.

Interestingly, when compared to the study conducted by Khayet et al. on SWRO and BWRO [4], the RSM model for low concentration saline water RO corresponds more closely to the SWRO RSM model than to the BWRO RSM model. For their SWRO RSM model, the effect of the feed concentration had a greater effect on the performance index than the feed pressure. For their BWRO RSM model, the feed pressure had a greater effect on the performance than the feed concentration. This is another example of how the optimization process requires different models for different ranges of TDS feed sources. It would be beneficial to run all the experiments on the same RO plant, to remove discrepancies caused by the different setups and membranes.

To validate Eq. (7), the analysis of variances (ANOVA) has been conducted. The ANOVA results of the RSM modelling are listed in Table 4. The optimal model equation coefficients (Eq. (7)) were determined with a squared correlation coefficient R^2 of 0.9999 being achieved. The R^2 value is a statistical measure, being very close to one that indicates a well-fitting model.

It is seen from Table 4 that the F -value of the model is 1805.9 and the p -value is less than 0.0001. The model F -value and p -value indicate that the model is significant and not by chance. All individual variable coefficients of the polynomial regression equation have high F values and very low p values, as desired. The ANOVA data indicates that the model is statistically significant and valid. The sensitivity analysis of the ANOVA has validated the coefficients of the RSM model in Eq. (7).

The measured performance index listed in the last column of Table 3 and the predicted performance index through Eq. (7) (the RSM model) are compared in Figure 3. The RSM predicted performance indices are very close to the experimentally measured values suggesting

Table 4. The ANOVA analysis results for the RSM Modelling.

	Sum of Squares	df	Mean Square	F value	P -value
Model	6.64×10^5	12	55325.55	1805.86	<0.0001
x_1	1.05×10^5	1	1.05×10^5	3441.27	<0.0001
x_2	14623.92	1	14623.92	477.33	0.0002
x_3	37211.77	1	37211.77	1214.62	<0.0001
x_1x_2	38433.59	1	38433.59	1254.50	<0.0001
x_1x_3	3455.02	1	3455.02	112.77	0.0018
x_2x_3	3178.68	1	3178.68	103.75	0.002
x_1^2	1193.48	1	1193.48	38.96	0.0083
x_2^2	1211.33	1	1211.33	39.54	0.0081
$x_1x_2x_3$	2912.72	1	2912.72	95.07	0.0023
$x_1^2x_2$	491.08	1	491.08	16.03	0.028
$x_1^2x_3$	235.72	1	235.72	7.69	0.0693
$x_1x_2^2$	463.29	1	463.29	15.12	0.0301
Residual	91.91	3	30.64		
Lack of Fit	62.06	2	31.03	1.04	0.5699
Pure Error	29.85	1	29.85		
R^2	0.9999				
Adjusted R^2	0.9993				
Predicted R^2	0.9761				
Adeq Precision	163.7217				

performance indices predicted by the RSM model are accurate, reliable, and indicative of a closely modelled system.

The RSM model of Eq. (7) was used to optimize the input variables of the small-scale reverse osmosis plant process for the maximum performance index output. The objective optimization was conducted for the three normalized input values ranging from −1 to +1, using mathematical optimization embedded in the software of Design Expert 12. The RSM model predicted the highest performance index over the normalized values ranging from −1 to 1 to be 852.25×10^{-5} L/m²s. This is compared with the highest performance index measured experimentally at 853.58 × 10⁻⁵ L/m²s for the same normalized range from −1 to +1. The combinations of the three input parameters for the highest performance index are the same for the experiments and RSM prediction. The optimal parameter combination for this range is: x_1 (feed concentration) = −1 (1000 mg/L), x_2 (feed temperature) = +1 (37.5 °C) and x_3 (feed pressure) = +1 (145 psi) which was obtained from running Design Expert 12 software.

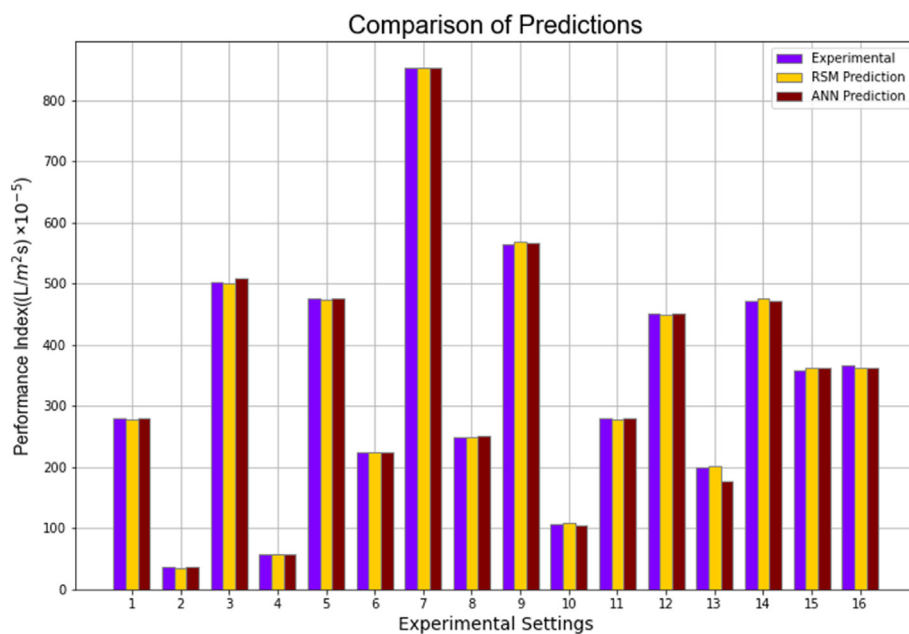


Figure 3. Comparison of the measured, RSM, and ANN predicted output performance index response.

RSM Performance Index prediction at 40 °C

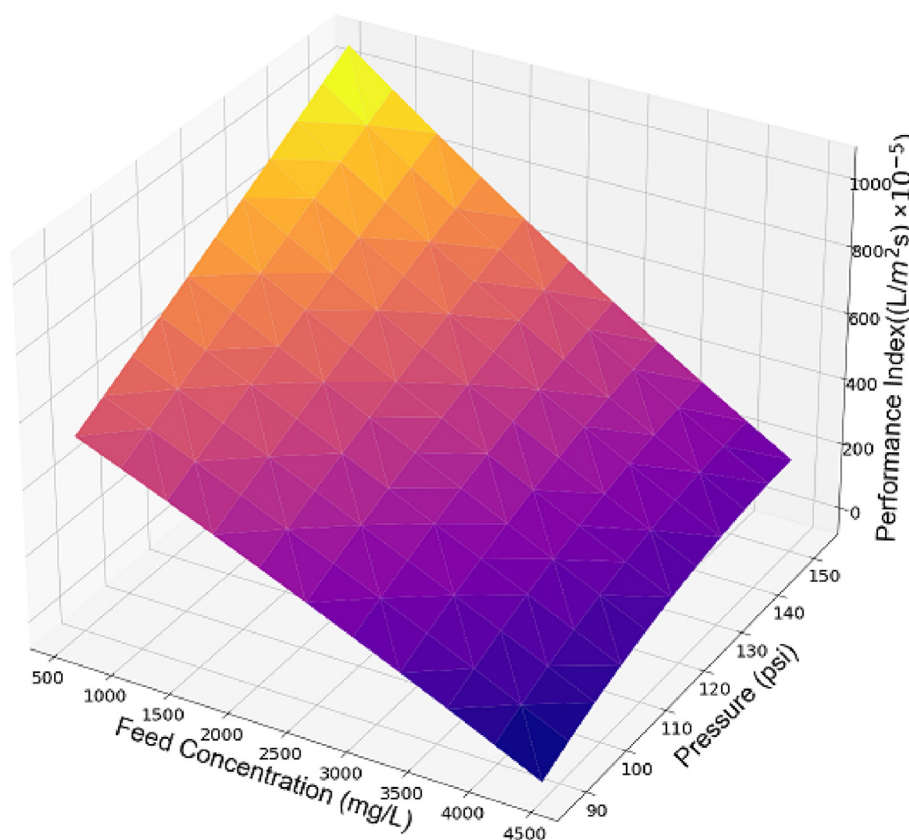


Figure 4. Three-dimensional response surface of the input parameters within the star point ranges with the temperature set to the Star point of 40 °C.

When the parameter range is extended to the largest normalized experimental value range (−1.287, 1.287) the highest performance index achieved is $1023.23 \times 10^{-5} \text{ L/m}^2\text{s}$. The highest performance index is

achieved with the input parameters: x_1 (feed concentration) = −1.282 (577 mg/L), x_2 (feed temperature) = 1.149 (38.8 °C), and x_3 (feed pressure) = 1.223 (150.57 psi).

ANN Performance Index prediction at 40 °C

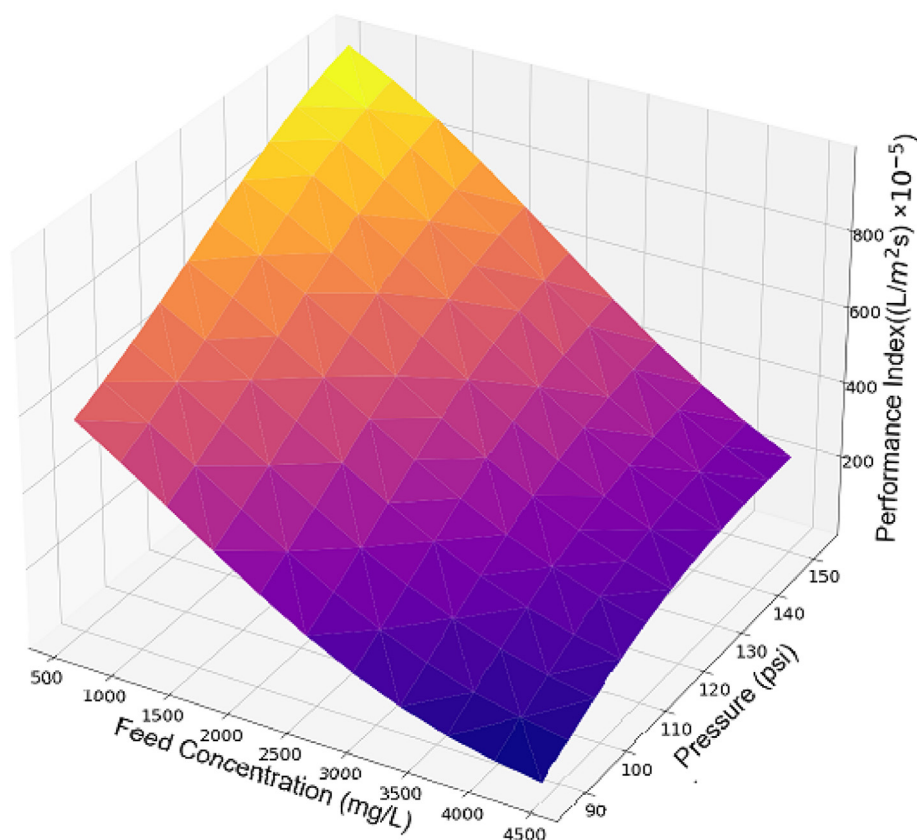


Figure 5. Three-dimensional surface plot of the ANN prediction of the input parameters within the star point ranges with the temperature set to the Star point of 40 °C.

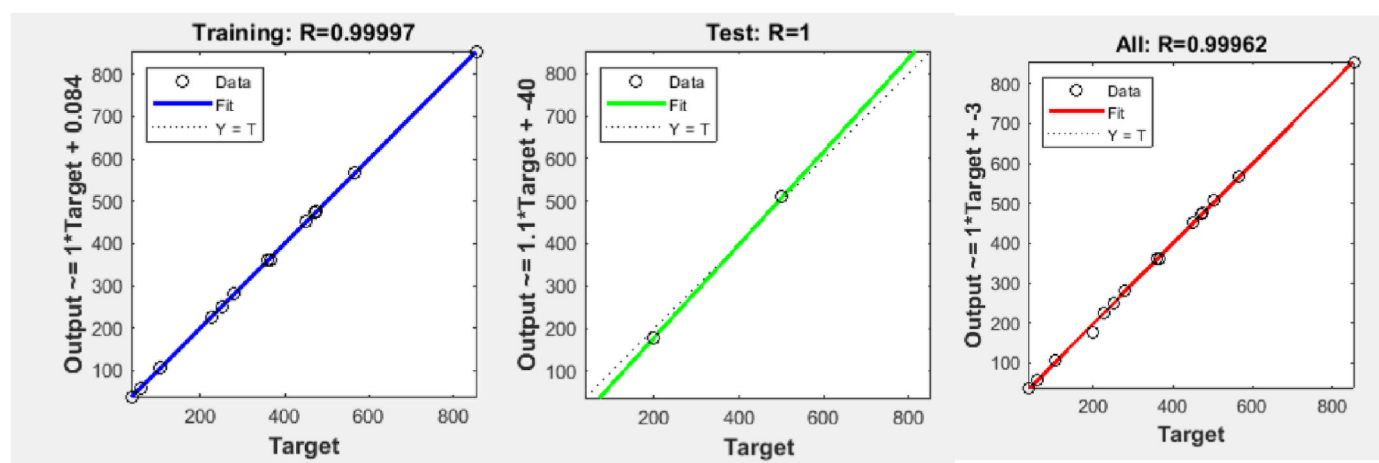


Figure 6. The regression of the ANN model.

Figure 4 depicts a three-dimensional surface response of the RSM model over the experimental range, set at the maximum experimental temperature of 40 °C.

Using an ANN designed in Matlab, composed of a single hidden layer with 10 neurons and the Bayesian regularisation training algorithm, a network model was developed. The first layer used the TRANSIG transfer function and the output layer the PURELIN function. The Bayesian regularisation was used as it is known to be robust even with a small sample set and not prone to overfit [22]. The model was found to be a close fit to

the experimental data, with an overall R^2 equal to 0.999 shown in Figure 6. 12 of the experimental runs were used for training, 2 for validation, and 2 for testing. The ANN model does not require the input parameter variables and measured responses in the carefully planned (CCD) experiments to be trained and can continually be trained using more available unplanned experimental data.

Table 5 lists the related error between the experimental and ANN prediction, most of the response output value errors are close to 0. Those used for testing and validation (runs 3, 13, 15, and 6) have a slightly

Table 5. Comparison of the response output values of the performance index predicted by the RSM and ANNs models with those measured by the experiments and their associated errors.

Experimental Run	Experimental Value $\times 10^{-5}$ L/m ² s	RSM Model Predictions $\times 10^{-5}$ L/m ² s	ANN Model Predictions $\times 10^{-5}$ L/m ² s	RSM Model Error $\times 10^{-5}$ L/m ² s	ANN Model Error $\times 10^{-5}$ L/m ² s
1	279.57	278.51	279.46	1.06	-0.12
2	36.74	35.67	37.00	1.07	0.26
3	501.99	500.93	509.41	1.06	7.42
4	58.23	57.17	58.26	1.06	0.03
5	475.11	474.05	475.21	1.06	0.10
6	225.47	224.41	225.31	1.06	-0.16
7	853.58	852.51	852.93	1.07	-0.65
8	250.37	249.31	250.48	1.06	0.11
9	565.36	569.02	565.83	-3.67	0.48
10	106.16	109.83	105.95	-3.67	-0.22
11	279.61	277.89	279.53	1.71	-0.08
12	450.63	448.91	451.01	1.72	0.38
13	199.17	202.38	177.72	-3.20	-21.45
14	471.98	475.18	472.65	-3.20	0.67
15	358.92	363.40	362.38	-4.48	3.46
16	366.65	363.40	362.38	3.25	-4.26

higher error, with the prediction of run 13 having a maximum error of -21.45. This error is likely to decrease as more training data becomes available. It would be possible to retrain the network with all data including the test and validation sets. As the model has been shown to have decent accuracy, this could improve the model further. More experimental runs could then be used as validation for the new model and then used for further training. This is the aspect of continuous improvement of the prediction model. In addition, the optimization prediction results based on the RSM model were verified for the dimensionless input parameter combination of (-1, 1, 1), producing a response output value of the performance index of 852.93×10^{-5} L/m²s. The optimization prediction results of the RSM model are very close to that of the ANN prediction and experimental results. The ANN model has been validated by the RSM model and experiments. The ANN prediction was similarly represented by a 3D surface plot at 40 °C in Figure 5. The trend observed when comparing Figures 4 and 5 is very similar, further indicating the close performance of the RSM and ANN prediction models.

Table 5 compares the predicted response output or performance index values of the RSM and the ANN models, and their errors with respect to the experimentally measured values. Table 6 presents the root-mean-square error (RMSE) of the calculated response output values of the RSM and ANN predictions. The RSM model is currently more accurate than the ANN model. It is seen from Table 5 that many of the ANN errors are close to 0. The larger error of the ANN response output values (runs 3, 13, 15, and 16) produces a large RMSE. If the amount of experimental data sets increases, the ANN model is expected to become more accurate than the RSM model.

The major benefits of the RSM method and CCD of input parameters are being able to fit a good response prediction model from fewer experiments and an explicit polynomial regression equation can be fit and presented, which helps to further understand the relationship between input and output for the process. The RSM model allows for parameter study and design optimization. The RSM modelling does require careful design of experiments by CCD and understanding of the capabilities of the equipment used, as the maximum and minimum input parameter values need to be defined for the experiments. Alternatively, an ANN can

take any experimental run input and output values as training data, generally, the more training data the more accurate the ANN model will be for the response output prediction. Using the RSM data provides a good platform for training the ANN as it is a good representation of the design space. ANN does not provide insight into how the process operates, but ANN provides the mechanism of the learning process with hidden layers and high levels of iteration to predict output from input.

The RSM and ANN complement each other as the RSM provides insight into the most effective input parameters on the response but also provides good, dispersed training data over the design space for training an ANN. The ANN is likely to provide better response prediction as more data is fed into the model. Therefore, for the optimization and prediction of performance for low salinity RO, the RSM can be used to better understand the system before using the ANN as an evolving, continuously improving prediction model. When researchers compare the accuracy of both RSM and ANN, it is greatly affected by the accuracy and quality of the data. Some examples show the ANN is more accurate [13, 14], especially considering it is not bound by linear or quadratic terms. But, RSM has a great deal of merit in understanding the system and providing a good representation of the design space data.

The feed flow rate was found to have no significant impact on the performance index of an RO plant [4], therefore it was not controlled for this experiment. However, in practice, the feed flow rate is an important consideration as it has an implication on membrane fouling/scaling which can degrade the long-term performance of the RO systems. If the system is run continuously with a feed flow rate that is too low, the salt build-up or scaling occurs as freshwater is extracted and the salts exceed solubility limits, leading to crystallization on the membrane surface reducing the effective area of the RO [23]. So, the feed flow rate should be great enough to avoid this premature fouling but does not drastically change performance over a certain minimum value. RO plants can mitigate scaling by maintaining a sufficient feed flow rate, lowering permeate recovery, and/or using antiscalants. In many cases, the RO retentate/brine needs to be disposed of, therefore the highest possible water recovery is ideal for reducing the volume of plant waste requiring appropriate disposal. This is just another example of the complexity of designing an RO plant, the feasibility of which is heavily reliant on the costs of waste handling, environmental impact management, and power consumption.

4. Conclusions

An RSM prediction model has been developed for low salinity feed-water to optimize input parameters for the largest predicted performance

Table 6. The root-mean-square error of the response output calculated by the RSM and ANN models.

Root Mean Square Error	
RSM	2.41
ANN	5.85

index of $1023.23 \times 10^{-5} \text{ L/m}^2\text{s}$. The predicted performance index was achieved with low feed concentration, high feed pressure, and high feed temperature. The corresponding parameters were obtained through the polynomial regression relationship between the input parameters and output response target using RSM. They were x_1 (feed concentration) = -1.282 (577 mg/L), x_2 (feed temperature) = 1.149 (38.8 °C) and x_3 (feed pressure) = 1.223 (150.57 psi). An ANN was also used to model the data, showing slightly less, but still acceptable accuracy. The close tie between the RSM and ANN results shows the validity of the methods for this system. The paper outlines where the RSM and ANN models can be complementary to each other to optimize the system parameters for the best performance while also producing a continuously improving prediction model.

Future work should look at using a BWRO membrane and actual brackish water in place of potable water-NaCl solution. The method proposed for performance modelling, optimization, and prediction can be used and further validation experiments can be undertaken. To further progress this research a single system could be set up to run a complementary RSM and ANN performance modelling, optimization, and prediction process as outlined in this paper for the low salinity RO, BWRO, and SWRO to remove discrepancies that can occur from different equipment.

5. Experimental method

The temperature sensor was calibrated before the experimental runs. The temperature was measured using a handheld thermometer and was compared with that measured by a K-type thermocouple over the temperature range of 17–40 °C. It was found that the difference in the readings was within ± 0.1 °C. When the feed solution was heated, time was taken to ensure the heat had diffused throughout the solution.

The TDS meter/water quality tester (EZ-9908) was calibrated for measuring the Total Dissolved Solids (TDS) before the experimental runs. Electrical conductivity (EC) was measured against a known calibration solution. For the handheld device, TDS is calculated from the measured EC. The handheld meter calibrated for EC has an accuracy of $\pm 2\%$ per the manufactures specification. The conversion formula from EC to TDS is approximately equal to $\text{EC} \times 0.670$ up to the seawater salinity. This formula is considered adequate when the solute was NaCl as used in this experiment [24].

The readings of multiple manometers/pressure gauges were calibrated over the range of the measured pressure, from 95 psi to 152 psi. Two gauges made by different manufacturers were used to measure the same pressure for each experimental run. The difference in the readings of the two gauges was ± 2 psi. Flowrate was measured manually. The permeate and retentate were collected in separate vessels for measurement.

Each of the 16 experimental runs was completed by filling the feed tank with tap water and then heating the tank water to the required temperature for the experimental run parameters. The next step was adding salt (NaCl), letting it dissolve, and mixing thoroughly. To achieve desired salinity, the solution salinity was measured using the TDS meter adding NaCl or more water as required. The temperature was rechecked and adjusted if required. The pumps were run with the valve open to the atmosphere for ~ 10 s, allowing the system to flush water and air. The valve was closed to produce the desired pressure, letting the permeate production commence, and allowing the line to run and flush, for ~ 10 more secs to rid the line of residual fluid. Then samples were collected for the retentate and permeate for 30 s (for the calculation of flow rate), repeating the collection three times. The TDS was measured and recorded for the permeate using the TDS meter. The weights of the retentate and permeate samples were measured and recorded. Pumps were shut down and the system rest before the next experimental run.

Declarations

Author contribution statement

Ryan Brooke: Designed and performed the experiments; Analyzed and interpreted the data; Wrote the paper.

Xu Wang: Conceived and designed the experiments; Analyzed and interpreted the data.

Linhua Fan: Contributed reagents, materials, analysis tools or data; Wrote the paper.

Mohamed Khayet: Analyzed and interpreted the data; Contributed reagents, materials, analysis tools or data.

Funding statement

Professor Xu Wang was supported by Australian Research Council [ARC DP170101039].

Data availability statement

No data was used for the research described in the article.

Declaration of interest's statement

The authors declare no conflict of interest.

Additional information

No additional information is available for this paper.

Acknowledgements

The authors would like to thank Gopi Krishnan Gopalakrishnan for their assistance in the experimentation.

References

- [1] U. Nations, Water Scarcity. <https://www.un.org/waterforlifedecade/scarcity.shtml> accessed 26 July 2022.
- [2] Dupont, FilmTec™ Reverse Osmosis Membranes Technical Manual, 2021.
- [3] S. Roy, S. Ragunath, Emerging membrane technologies for water and energy sustainability: future prospects, constraints and challenges, *Energies* 11 (11) (2018) 2997 [Online]. Available: <https://www.mdpi.com/1996-1073/11/11/2997>.
- [4] M. Khayet, C. Cojocaru, M. Essalhi, Artificial neural network modeling and response surface methodology of desalination by reverse osmosis, *J. Membr. Sci.* 368 (1) (2011) 202–214.
- [5] M. Li, A unified model-based analysis and optimization of specific energy consumption in BWRO and SWRO, *Ind. Eng. Chem. Res.* 52 (48) (2013) 17241–17248.
- [6] M. Khayet, M. Essalhi, C. Armenta-Déu, C. Cojocaru, N. Hilal, Optimization of solar-powered reverse osmosis desalination pilot plant using response surface methodology, *Desalination* 261 (3) (2010) 284–292.
- [7] M. Li, Optimal plant operation of brackish water reverse osmosis (BWRO) desalination, *Desalination* 293 (2012) 61–68.
- [8] M.A. Alghoul, P. Poovanaesvaran, K. Sopian, M.Y. Sulaiman, Review of brackish water reverse osmosis (BWRO) system designs, *Renew. Sustain. Energy Rev.* 13 (9) (2009) 2661–2667.
- [9] N. Voros, Z.B. Maroulis, D. Marinos-Kouris, Optimization of reverse osmosis networks for seawater desalination, *Comput. Chem. Eng.* 20 (1996).
- [10] A. Villafafila, I.M. Mujtaba, Fresh water by reverse osmosis based desalination: simulation and optimisation, *Desalination* 155 (1) (2003) 1–13.
- [11] J. Kim, K. Park, S. Hong, Optimization of two-stage seawater reverse osmosis membrane processes with practical design aspects for improving energy efficiency, *J. Membr. Sci.* 601 (2020), 117889.
- [12] K.M. Desai, S.A. Survase, P.S. Saudagar, S.S. Lele, R.S. Singhal, Comparison of artificial neural network (ANN) and response surface methodology (RSM) in fermentation media optimization: case study of fermentative production of scleroglucan, *Biochem. Eng. J.* 41 (3) (2008) 266–273.
- [13] A. Khamparia, B. Pandey, D.K. Pandey, D. Gupta, A. Khanna, V.H.C. de Albuquerque, Comparison of RSM, ANN and fuzzy logic for extraction of oleonolic acid from *ocimum sanctum*, *Comput. Ind. Eng.* 117 (2020), 103200.
- [14] A. Hammoudi, K. Moussaceb, C. Belebchouche, F. Dahmoune, Comparison of artificial neural network (ANN) and response surface methodology (RSM) prediction in compressive strength of recycled concrete aggregates, *Construct. Build. Mater.* 209 (2019) 425–436.

- [15] S.R. Dixit, S.R. Das, D. Dhupal, Parametric optimization of Nd:YAG laser microgrooving on aluminum oxide using integrated RSM-ANN-GA approach, *J. Ind. Eng. Int.* 15 (2) (2019) 333–349.
- [16] T. Aung, S.-J. Kim, J.-B. Eun, A hybrid RSM-ANN-GA approach on optimisation of extraction conditions for bioactive component-rich laver (*Porphyra dentata*) extract, *Food Chem.* 366 (2022), 130689.
- [17] Dow, FILMTEC™ SW30-2521 membranes. <https://www.lenntech.com/Data-sheets/Dow-Filmtec-SW30-2521.pdf> accessed 12 May 2021.
- [18] Dow, FILMTEC™ Reverse Osmosis Membranes Technical Manual. <https://www.ra.inmadesal.com/wp-content/uploads/2018/09/dow-filmtec-sw30-manual.pdf> accessed 12 May 2021.
- [19] B. Ait-Amir, P. Pougnet, A. El Hami, *Embedded Mechatronic Systems 2: Analysis of Failures, Modeling*, Elsevier, 2020.
- [20] A. Paar, Viscosity of Water. <https://wiki.anton-paar.com/au-en/water/> accessed 22 May 2021.
- [21] M.F.A. Goosen, S.S. Sablani, S.S. Al-Maskari, R.H. Al-Belushi, M. Wilf, Effect of feed temperature on permeate flux and mass transfer coefficient in spiral-wound reverse osmosis systems, *Desalination* 144 (1) (2002) 367–372.
- [22] F. Burden, D. Winkler, D.J. Livingstone (Eds.), *Artificial Neural Networks (Methods in Molecular Biology)*. Humana Totowa, NJ, 2009.
- [23] A. Matin, F. Rahman, H.Z. Shafi, S.M. Zubair, Scaling of reverse osmosis membranes used in water desalination: phenomena, impact, and control; future directions, *Desalination* 455 (2019) 135–157.
- [24] P. Coleman, *Methods Manual for Salt Lake Studies/Salinity*. https://en.wikibooks.org/wiki/Methods_Manual_for_Salt_Lake_Studies accessed 15 May 2021.