Prediction of capital cost of ro based desalination plants using machine learning approach

Mohamed Ibrahim Kizhisseri[1,*], *Mohamad* Mostafa Mohamed[2] and *Mohamed* A. Hamouda[3]

[1] PhD Candidate, United Arab Emirates University, Al Ain, UAE
[2] Professor, United Arab Emirates University, Al Ain, UAE
[3] Assistant Professor, United Arab Emirates University, Al Ain, UAE

Abstract. This paper presents a neural network tool for predicting the capital cost of desalination plants based on reverse osmosis technology. A multi-layer feedforward neural network with back propagation learning method is used to model the investment cost of RO plants. The model is developed using the data sets of 1806 RO plants of capacity at least 1000 m3/day, which involved training, testing and validation. The model used six inputs that included both categorical and numerical data elements, namely: plant location, plant capacity, project award year, raw water salinity, plant types, and project financing type. The output is the capital cost of the RO plants planned. This prediction model can be used by governments, investors or other stakeholders in desalination industry to make a reasonable estimate of investment costs of upcoming RO plant projects.

Introduction

With the increase in water demands over years many of the water scarce countries are depending on desalination technologies for producing drinking water. Desalination can be defined as the process of removing dissolved salts from water with high salinity to produce water with low salinity that meets the quality (salinity) requirements of various purposes(1). There are different technologies by which desalination is carried out which include but not limited to major technologies like multi-stage flash distillation (MSF), multiple effect distillation (MED), reverse osmosis (RO) and hybrids of these. RO is one of the newer desalination technologies that's on the rise all over the world. During last few decades, the use of RO has gained wide acceptance in the Middle East countries because of lower cost and easy operation (2,3). A desalination plant is a large-scale project which is usually undertaken by governments; hence, an estimation of the project cost is important for governments, investors and stakeholders. When new RO plants are proposed, an estimation of the expected project cost is key in deciding the budget allocation. This can also help identify whether the project is cost effective or an alternative is to be looked for.

The project cost of an RO plant consists of two major components: the capital cost which is a one-time cost, and the annual operating costs (4,5). The capital cost also called as capital expenditure (CAPEX), can be direct or indirect costs. Direct CAPEX includes land cost, costs of take into consideration some parameters like year of construction and geographical location of the plant, which are already established to have impact on the cost.

major equipment ,engineering cost and so on (6). The indirect capital costs include elements such as insurance, construction, and overhead (7). The operating cost, also called as OPEX, is the cost required to run the desalination plant which includes energy cost and other costs such as labour cost, spare parts, chemicals, membrane replacement, and so on (6,8). There are different methodologies and assumptions adopted by various researchers in order to model the capital and production costs. However, this study is limited to develop a predictive model to estimate the capital cost of future RO plants using a methodology which is different from past studies.

This study began with gathering information on various methodologies available for the cost estimation of desalination plants. Literature review showed that comparative cost models, statistical models, semi-empirical cost models and parametric cost estimation are the most commonly used approaches. Comparative cost methods are old estimation methods, in which estimates are based on the published cost data available. In 1976, Glueckstern and Reed (9) published a report on the product water cost where plant sizes in the range of 1 to 200MGD were investigated. Comparative method was adopted by Shatat and Riffat (10), and Al-Karaghouli and Kazmerski (11) for developing a table that provided cost ranges for major desalination technologies. One major drawback of this method is that it didn't

Statistical models are those in which the correlation of capital cost with various parameters are determined. Multiple linear regression (MLR) modelling or other statistical forecasting techniques like autoregressive

integrated moving average (ARIMA) are most commonly used. In 2008, Wittholz et al. (5) used a database of more than 300 plants and developed a capital cost correlation as a function of plant capacity for large scale RO, MED and MSF desalination plants. The study conducted by Zhou and Tol (12) used data of 2514 RO plants implemented worldwide between the years 1970 and 2001, and developed a correlation equation for unit cost as a function of raw water quality, cumulative installed capacity, plant capacity and year of installation. A similar regression model was developed by Loutatidou et al. (13) which can predict the capital cost of RO plants GCC countries and European countries. Curve fitting was used by Lamei et al. (4) for developing a basic cost equation for RO plants, with the help of data of 21 RO plants implemented in Egypt and outside. ARIMA technique was used by Dore (4) to forecast the nature of unit cost of desalination in near future. All the statistical models discussed above tried to identify the main parameters affecting the cost of desalination and tried to provide a tool for estimating the cost of desalination plants in terms of unit cost, capital cost, and O&M cost.

Semi-empirical methods are deployed in determining the investment and production costs of RO plants. Studies conducted by Greig and Wearmouth (15) in 1987, and Frioui and Oumeddour (8) in 2008 are examples that fall in this category. The former estimated the cost of various components of desalination plants using the published data of other existing plants, while the latter focused on estimating the capital and production costs of RO plants by considering the parameters like plant capacity, plant availability, energy price and consumption, membrane service life and various process variables.

There are few commonly available commercial cost estimating tools in the market .In literatures (16–18) , tools namely, Desalination Economic Evaluation Program (DEEP), Water Treatment Cost Estimation Program (WTCost) and Global Water Intelligence Cost Estimator (GWI CE) are found. DEEP is available as a free software which can be used for performance and cost evaluation of desalination plants. GWI CE is a tool that can be used to estimate the capital expenditure (CAPEX), operating expenditure (OPEX), and resulting water prices of seawater reverse osmosis, and other desalination techniques. These tools work on the tailor-made pre-programmed rules or algorithms by the developers. Also, it requires enough knowledge on the technical and design parameters of desalination, which makes it ineffective tool for use by economists.

This research is focused to develop a tool which could predict the capital cost of RO plants with reasonably good accuracy. An analysis of expected project cost is important is making decisions on any proposed project, and therefore, same is applicable for RO plants. This will help in deciding the budget allocation and approval of a project, even though the decisions are dependent on other factors. Therefore, an efficient and easy cost predictor is required in planning and implementation of any desalination projects. The proposed methodology is based on artificial intelligence which is reliable and easy to use. Self-learning technique in artificial intelligence is already in use for many of the

scientific and engineering application. Hence, the artificial neural network approach using machine learning technique is made use in this study. Studies conducted by Elfahham (20) and Trefor (21) are some of the examples of application of artificial intelligence using neural networks (NN) in the construction industry. In this paper, a multilayer perceptron with back propagating learning method is deployed to predict the capital cost of the RO plants. The model is developed using the data available for 1806 RO plants.

Methodology

There are many parameters that can affect the capital cost of RO plants. It can be both technical and non-technical. Technical parameters include capacity, desired quality of product water, raw water quality, pre-treatment requirements, disposal of brine discharge, options of membrane and so on. Non-technical parameters can be social, environmental or geographical. In this research, the preliminary tasks comprised data collection and processing. Cost data on RO plants was collected from the database available with Global Water Intelligence (GWI) (17), who keeps the inventory of all desalination plants. The data contained information on many design parameters as well as other non-technical details like location, procurement type, capital cost, year of the project, contractual period, supplier details, consultants and so on. However, pre-processing of data identified those parameters which are identified to be relevant for our study. The pre-processing involved choosing the potential variables that can affect the capital cost and/or operational cost. Based on the comprehensive search in the literature (4,5,13,22–25), six variables from the cost database got qualified as the modelling variables. They are plant capacity, raw water type, year of project award, geographical location of the plant, plant type and project financing type. Plant capacity is established to be an important parameters through a number of studies including(4,23).Raw water quality expressed in terms of salinity level are used by (13) .There are several papers which presented geographical location proves an important role in the capital cost the RO plants, as the prices are found to vary with continents and countries (13,22,25).The year of project is found to be very important as the cost was found to vary with years (8,14,20,21,24). The type of procurement or financing and plant type is chosen to find out the impact of other possible non-technical parameters on the capital cost. Both were included in the study as they were found to be statistically important. Thus, a list of readily available potential parameters was selected to the neural training and is shown in Table 1.

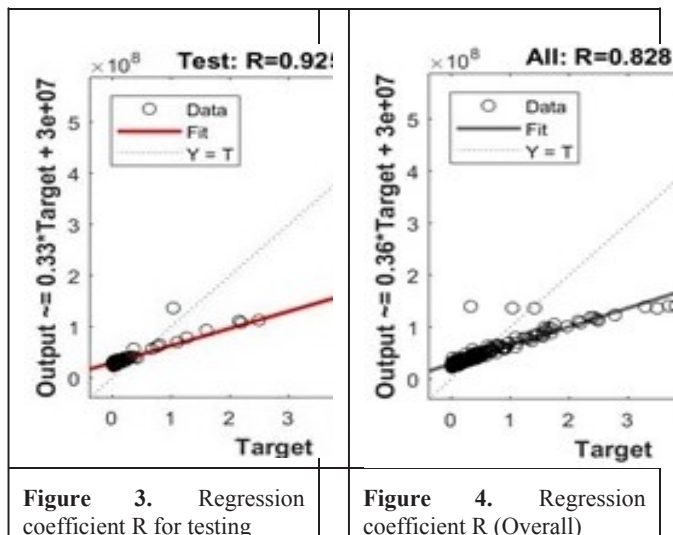**Table 1:** List of parameters used in the neural training

Neural networks are algorithms that learn from the input data to predict the output. This is a computing

Parameters	Range of values /categories
Plant capacity	1000- 1000000 m³/day
Raw water quality	Brackish TDS 3000ppm - <20000ppm) Seawater (TDS 20000ppm -
Plant type	Dual Purpose, Stand Alone
Year of the project	1980- 2018
Geographical location	Americas, Asia and EMEA
Project Financing /Procurement types	EPC, BOO, BOT, BOOT, DBO

process by which relationships between the input variables and outputs are established. The training data of the neural training are compiled for 1806 RO plants which were built between 1980 and 2018.As the input data elements comprised both categorical and numerical variable, all the categorical variables are first converted into an equivalent number in binary values. The output of the model is the capital cost of the RO plants. The cost values are converted to present value in US dollars corresponding to 2019, considering the inflation over a large time span. In this study, neural network toolbox in MATLAB is used to develop the model. For the training process, the data was sorted into three sets, the first set comprised 70% of the data for training, the second had 15 % for testing, and the remaining 15% was assigned for validation. All the major training algorithms available in MATLAB environment were explored, namely, Levenberg-Marquardt backpropagation, Bayesian regularization and scaled conjugate gradient. Multi-layer neural networks with different number of hidden layers tried to examine the impact on model performance. The performance of the model was evaluated using MSE, number of epochs, regression coefficient r value and training time.

Results And Discussion

A multi-layer neural network with 25 hidden layers was developed using the Levenberg-Marquardt backpropagation which gave best performance of the model. The Levenberg-Marquardt gave best training in fewer epochs and in short results. The model achieved regression coefficient (R) of 0.961 for the training (Figure 1), 0.817 for validation (Figure 2) and 0.9259 for testing (Figure 3). The overall R value for the model was 0.8285 (Figure 4), which is a good fit. The best validation result was achieved at an epoch of 70 with an MSE of 1.530 (Figure 5).

Figure 3. Regression coefficient R for testing

Figure 4. Regression coefficient R (Overall)

A sample NN predicted results of a hypothetical case is shown the Figure 6. In this prediction, the inputs used are "capacity-10000 m3 / day", "feed water quality - saline (seawater)", "plant type -standalone", "region-Middle East Asia", "project procurement type - EPC contract (Engineering, Procurement & Construction)".The prediction was made for years 2020 to 2030. The capital cost is showing a decreasing trend. This could be due to the combined effect of the parameters and is matching with the current trend where the RO plants are decreasing over years. A periodic update of the data is essential to make the predictions more accurately and over a wider span of time.
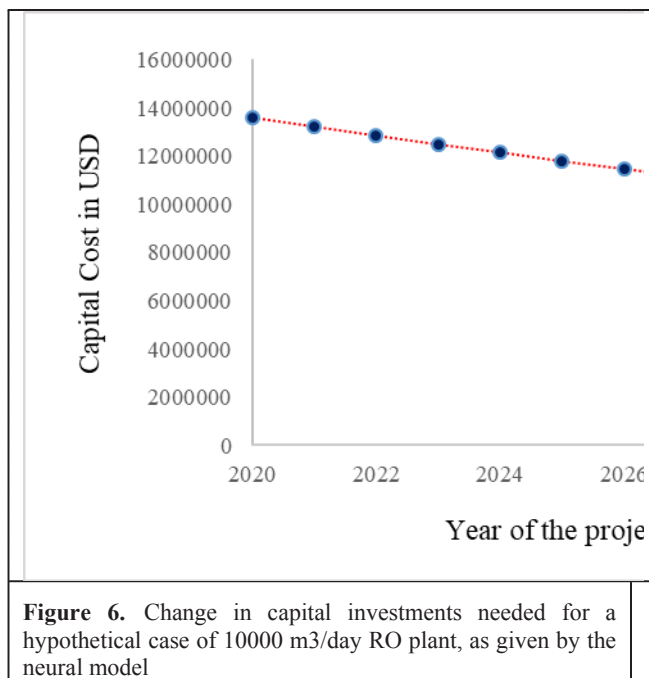

Figure 6. Change in capital investments needed for a hypothetical case of 10000 m3/day RO plant, as given by the neural model

Conclusion

In this paper, the applicability of machine learning using neural networks in predicting the capital cost trends of RO is studied. As the projects like RO plant construction

require high capital investment, there is a need to estimate the capital cost for the purpose of planning and budget allocation. A NN cost prediction model is developed to predict the capital cost of RO plants. Thus, the main contribution of this study is to provide the stakeholders of a water desalination projects with an easy and reliable tool for estimating the expected investments of coming desalination projects.

In the light of the reasonably good results, the same methodology can be adopted to other desalination types which will help in reducing the cumbersome estimation of investment cost in the pre-approval stage of any water treatment plants.

References

1. Darre NC, Toor GS. Desalination of water: A review. Curr Pollut Rep. 2018 Jun;4(2):104–11.

2. Khawaji AD, Kutubkhanah IK, Wie J-M. Advances in seawater desalination technologies. Desalination. 221(1–3):47–69.

3. Greenlee LF, Lawler DF, Freeman BD, Marrot B, Moulin P. Reverse osmosis desalination: water sources, technology, and today's challenges. Water Res. 2009 May;43(9):2317–48.

4. Lamei A, van der Zaag P, von Münch E. Basic cost equations to estimate unit production costs for RO desalination and long-distance piping to supply water to tourism-dominated arid coastal regions of Egypt. Desalination. 2008 May 1;225(1):1–12.

5. Wittholz MK, O'Neill BK, Colby CB, Lewis D. Estimating the cost of desalination plants using a cost database. Desalination. 2008 Sep 15;229(1):10–20.

6. Ettouney HM, El-Dessouky HT, Faibish RS, Gowin PJ. Evaluating the economics of desalination. Chem Eng Prog. 2002 Dec 1; 98:32–9.

7. Reddy KV, Ghaffour N. Overview of the cost of desalinated water and costing methodologies. Desalination. 2007 Feb 5;205(1):340–53.

8. Frioui S, Oumeddour R. Investment and production costs of desalination plants by semi-empirical method. Desalination. 2008 Mar 1; 223:457–63.

9. Glueckstern P, Reed A. A comparative investigation of the economics of seawater desalting based on current and advanced distillation concepts. Oak Ridge National Laboratory,ORNL/TM-5229; 1976.

10. Shatat M, Riffat SB. Water desalination technologies utilizing conventional and renewable energy sources. Int J Low-Carbon Technol. 2014 Mar 1;9(1):1–19.

11. Al-Karaghouli A, Kazmerski L. Energy Consumption and Water Production Cost of conventional and renewable-energy-powered desalination processes. Available from: https://www.researchgate.n¬¬¬-et/publication/255814703_Energy_Consumptio n_and_Water_Production_Cost_of_Convention al_and_Renewable-Energy-Powered_Desalination_Processes

12. Zhou Y, Tol RSJ. Evaluating the costs of desalination and water transport. Water Resour Res. 2005;41(3). Available from: https://agupubs.onlinelibrary.wiley.com/doi/abs/ 10.1029/2004WR003749

13. Loutatidou S, Chalermthai B, Marpu PR, Arafat HA. Capital cost estimation of RO plants: GCC countries versus southern Europe. Desalination. 2014 Aug 15; 347:103–11.

14. Dore M HI. Forecasting the economic costs of desalination technology. Desalination. Feb 2005;172(3): 207-214

15. H.W. Greig, J.W. Wearmouth. An economic comparison of 2 × 1000 m3/day desalination plants. Desalination. 1987 Jan 1; 64:17–50.

16. Bouhelal O-K, Merrouch R, Zejli D. Costs investigation of coupling an RO desalination system with a combined cycle power plant using DEEP code. Desalination. 2004 Aug 15; 165:251–7.

17. Global Water Intelligence. Available from: www.DesalData.com.

18. Kavvadias KC, Khamis I. The IAEA DEEP desalination economic model: A critical review. Desalination. 2010 Jul 1;257(1):150–7.

19. Detailed Estimation of Desalination System Cost. Available from: https://documents.pub/document/detailed-estimation-of-desalination-system-cost.html

20. Elfahham Y. Estimation and prediction of construction cost index using neural networks, time series, and regression. Alexandria Engineering Journal. 2019; 58:499–506.

21. Trefor WP. Predicting changes in predicting changes in construction cost indexes using neural networks. J Constr Eng Manage ASCE. 1994;120(2):306–20.

22. Papapetrou M, Cipollina A, La Commare U, Micale G, Zaragoza G, Kosmadakis G. Assessment of methodologies and data used to calculate desalination costs. Desalination. 2017 Oct 1; 419:8–19.

23. Sethi V, Clark RM. Cost estimation models for drinking water treatment unit processes. Available from: https://cfpub.epa.gov/si/si_public_record_report .cfm?Lab=NRMRL&dirEntryId=96119

24. Hinomoto H. Unit and Total Cost Functions for Water Treatment Based on Koenig's Data. Water Resour Res. 1971;7(5):1064–9.

25. Atikol U, Aybar HS. Estimation of water production cost in the feasibility analysis of RO systems. Desalination. 2005 Nov 1;184(1):253–8.