

Highlights

Integrated Pumped Hydro Reverse Osmosis System Optimization Featuring Surrogate Model Development in Reverse Osmosis Modeling¹

Matthew W. Haefner, Maha N. Haji

- MDO of an Integrated Pumped Hydro Reverse Osmosis System (IPHROS) is presented
- Fractional salt rejection rate and permeate flowrate surrogate models are developed
- Tradeoffs exist between optimal IPHROS designs with respect to different objectives
- IPHROS cost savings results in 16% faster break even time from simple NPV estimates

¹The short version of the paper was presented at ICAE2021, Nov 29 - Dec 5, 2021. This paper is a substantial extension of the short version of the conference paper.

Integrated Pumped Hydro Reverse Osmosis System Optimization Featuring Surrogate Model Development in Reverse Osmosis Modeling¹

Matthew W. Haefner^{a,c}, Maha N. Haji^b

^a*Department of Systems Engineering, Cornell University, Ithaca, New York, 14853, United States*

^b*Sibley School of Mechanical and Aerospace Engineering, Cornell University, Ithaca, New York, 14853, United States*

^c*Corresponding author: mwh85@cornell.edu,*

Abstract

The pressure head generated by the upper reservoir of a pumped hydro energy storage system can be sufficient for creating the pressure gradient required for a reverse osmosis desalination plant. Combined with the fact that many drought-stricken coastal areas have nearby mountains at the necessary elevation for these upper (mountaintop) reservoirs, a symbiotic relationship can be ascertained through the co-location of a pumped storage hydropower (PSH) system with a reverse osmosis (RO) desalination system. Merging PSH and RO into one Integrated Pumped Hydro Reverse Osmosis System (IPHROS) instead of implementing each individually could result in a number of benefits, including reduced capital investment, lower maintenance costs, and a natural mechanism for diluting the highly saline brine discharge generated from the RO process. This paper extends the work of Slocum et al. in 2016, who first introduced the concept of IPHROS, by optimizing the amount of seawater sent to and diverted from the upper reservoir for maximal energy recapture and freshwater production, respectively, while also seeking to maximize the RO system recovery ratio. For this multiobjective optimization, a

¹The short version of the paper was presented at ICAE2021, Nov 29 - Dec 5, 2021. This paper is a substantial extension of the short version of the conference paper.

new reverse osmosis model is created that utilizes a blend of empirical and fundamental equations based on the solution-diffusion model of membrane transport and boundary layer effects that naturally occur along reverse osmosis membranes. Additionally, surrogate models are developed to predict the permeate flowrate and fractional salt rejection rate for a Seamaxx™-440 RO element. Optimizing the presented IPHROS model reveals a 16% decrease in the break even time for IPHROS compared to PSH and RO being implemented individually, and that at the best design with regards to the energy, freshwater, and RO system recovery objectives, 79.5 million kWh of energy and 5.79 million cubic meters of fresh water can be delivered to a population, significant amounts for a population seeking to transition to a renewable energy-based grid and alleviate dire freshwater conditions. Enhanced modeling and optimization, as was initiated in this paper, will eventually aid in IPHROS' large-scale adoption into energy and freshwater infrastructures.

Keywords: Integrated Pumped Hydro Reverse Osmosis System, reverse osmosis, multidisciplinary design optimization, surrogate modeling, pumped storage hydropower

PACS: 0000, 1111

2000 MSC: 0000, 1111

1. Introduction

By the year 2050, demand for fresh water is expected to grow by over 40% [1]. This will increase the strain on a freshwater supply already being negatively impacted by factors such as droughts, increased urbanization, and an uneven distribution of freshwater resources [2, 3]. Several recent examples highlight this dire situation, ranging from the US Government in 2021 declaring a water shortage at Lake Mead, a vital water source for the Southwestern United States, for the first time in the nation's history due to perpetuating drought conditions [4], to the "Mega Drought" in Central Chile leading to significant water rationing in the capital city of Santiago in 2022 [5, 6]. Reverse osmosis (RO) desalination plants are one means of increasing the freshwater supply by converting saline water, such as seawater, to drinking water. However, the energy requirement for pressurizing the RO process, if from fossil fuel sources, would only exacerbate the lack of fresh water due to the effects of global warming [7]. In comparison to the treatment of surface water for

drinking water, which only has energy requirements of 0.2-0.4 kWh/m³, the treatment of seawater for drinking water requires significantly more energy at 2.5-4 kWh/m³ [8]. Therefore, having the ability to incorporate renewable energy sources into the RO process would be ideal for eliminating the tradeoff between freshwater production from RO and carbon emissions from burning fossil fuels.

One means of creating the pressure gradient required for the RO process is the potential energy of water stored in pumped storage hydropower (PSH) systems. In PSH systems, excess energy from the grid is used to power a pump that drives water from a lower reservoir to an upper reservoir, consequently converting electrical energy to potential energy that can be stored for practically infinite amounts of time. This is beneficial for the integration of both intermittent and continuous sources of renewable energy into the power grid, as these energies could then more easily become a primary energy source over fossil fuels and provide enhanced grid stability [9]. PSH is especially pertinent as a cost-effective means of industrial-scale storage for regions striving for an electrical grid powered by 100% renewable energy, which requires renewable energy integration across multiple timescales [10]. PSH has the lowest levelized cost of storage (LCOS) of all mature electrical storage technologies, including diabatic compressed air energy storage and Pb batteries [11], and has a very low energy-specific capital cost of 106-200 \$/kWh [12]. This same idea also holds true with regards to an effective CAPEX metric discussed in [13], which is based on a PSH lifespan of 80 years and a 6% discount rate.

As of the end of 2019, the world has an installed PSH capacity of 153 GW, with 226 GW of expected future installed capacity in the permitting and development or construction stages [12]. In comparison, total installed battery storage worldwide, as of the end of 2021, only contributes 16 GW of capacity [14]. While battery-based storage options have the potential to eventually usurp PSH in terms of LCOS and effective CAPEX in combination with its already faster response times [11, 12, 13], PSH is a much more prevalent and mature form of energy storage [15]. However, when compared to current energy consumption, it becomes clear that much more energy storage is required for a 100% renewable energy grid. As of 2019, the latest year at time of publication not impacted by COVID-19, world energy consumption was at 1.6153×10^5 TWh, with 1.8475×10^4 TWh of that energy, approximately 10%, coming from renewable sources, including hydropower [16]. Even with

the prevalence and maturity of PSH, overcoming this gap between global energy demand and its fulfillment from renewable energy may require some form of incentive, such as the pairing of PSH with another proven technology, to convince regions to implement more storage capacity.

Individually, seawater PSH plants and RO desalination plants have been implemented to varying degrees of success. While there are hundreds of examples of pumped hydro energy storage systems worldwide that, when combined, make up 97% of the world’s large-scale energy storage, there are far fewer examples of significant pumped seawater hydro plants [17]. The 30 MW Yanbaru Power Station in Okinawa, Japan was the first and only significant seawater PSH in the world before being decommissioned in 2016. More recently, Australia in the 2010’s was developing a 225 MW facility [17]. While the inability to secure a land lease ultimately halted the project [18], there are other initiatives being undertaken regarding the feasibility of seawater PSH in other regions of the world, such as the Repower Eenergy Development Corporation in the Philippines signing a deal with the Gugler Water Turbines GmbH in 2023 to set itself up for implementing seawater PSH across the country [19]. Conversely, seawater RO desalination has seen wide-scale implementation, such as the Sorek Desalination Plant in Israel, which as of 2016 is the largest seawater RO plant in the world with a 624,000 m³/day capacity capable of supplying 20% of Israel’s freshwater demand [20].

The fact that RO desalination plants and pumped hydropower plants can utilize and require elevated reservoirs, respectively, presents an opportunity for the co-location of these two systems as an Integrated Pumped Hydro Reverse Osmosis System (IPHROS). This co-location factor provides many of the advantages an IPHROS possesses. The sharing of infrastructures such as pumps and piping provides an opportunity for a reduction in capital investments and maintenance costs compared to if the two subsystems were implemented independently, and the utilization of the ocean provides a practically infinite source of water for RO and energy storage. Also, the presence of the RO system introduces flexibility to the PSH system that would not be present otherwise. For instance, additional reservoir capacity could be created by generating freshwater if there is no need for electricity from PSH by the grid. Furthermore, co-location allows for the highly saline brine discharge from the RO subsystem, after being mixed with the seawater stream from the PSH subsystem to achieve a net reduced salinity, to be directly discharged into the ocean. This innate property of IPHROS, beyond con-

venience, presents an additional opportunity for cost savings compared to if other brine disposal methods, such as diffusing outflow pipes and evaporation ponds, were used [21]. This is also important for mitigating adverse environmental effects and for regions that regulate the salinity of discharges into the ocean. These benefits could serve as incentives for governments to adopt these technologies, and this symbiotic relationship would allow RO to occur without burning fossil fuels and provide the added benefit of freshwater production for a community, which hydropower alone would not facilitate.

Figure 1 shows a schematic of IPHROS, illustrating the flow of energy and water throughout the system. IPHROS was first proposed in the 2016 paper by Slocum et al. [21]. The emphasis of that paper was to show that IPHROS is feasible and to provide hypothetical results. Work since then published on IPHROS has looked into the potential and optimal scheduling of IPHROS for a case study in Egypt [22], and in using IPHROS as a means of reducing the specific cost for copper production powered by renewable energy systems [23]. This paper will add to these works by undertaking initial efforts to find the optimal operating conditions that maximize the amounts of fresh water and energy sent to the consumer by following multidisciplinary design optimization (MDO) principles. MDO is a methodology for using numerical optimization strategies in the design of multidisciplinary systems. Use of an MDO approach allows for the optimization of coupled subsystems of different disciplines, in this case reverse osmosis and energy storage, to be considered simultaneously in the design and optimization of IPHROS. This framework also tends to generate better designs than those generated from a sequential optimization process between individual disciplines [24], and with its emphasis on problem formulation is suited for a project that is in its relative infancy. Additionally, this paper builds upon the initial IPHROS model by [21] to better emulate the real-life complexities of its subcomponents, in this case looking specifically at the RO module.

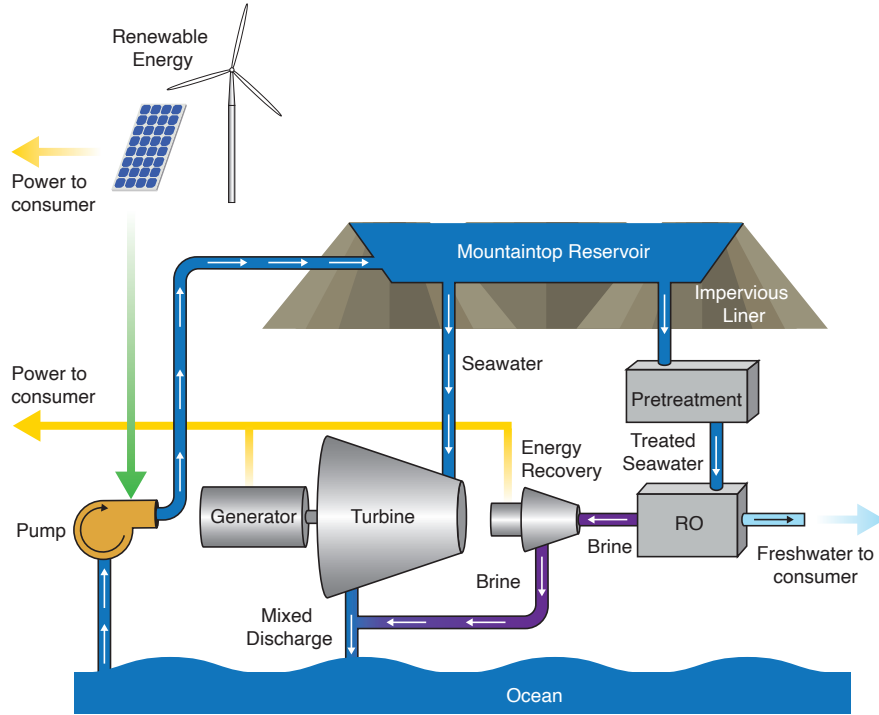


Figure 1: Schematic of IPHROS. Renewable energy from the grid is used to pump seawater to an upper (mountaintop) reservoir as a means of energy storage. While one seawater flow from the mountaintop reservoir is sent to a turbine for energy recovery, another is sent to an RO system for freshwater generation. The resulting, still highly-pressurized, brine stream is sent through an energy recovery turbine. This flow being purple as opposed to blue indicates increased salinity. The seawater stream from the mountaintop reservoir and the brine stream are then mixed to ensure dilution of the brine stream before the mixed discharge is sent into the ocean. Energy that is sent to the consumer comes from the renewable energy not sent to the grid and the energy recovered from the seawater and brine streams. Fresh water that is sent to the consumer comes from the RO system.

Previous works have investigated various components of IPHROS, but not in the context of IPHROS. A previous study conducted by Segurado et al. in [25] investigated the coupling of a PSH system with a desalination system and its optimization, but rather than pumping seawater through the PSH system as is presented for IPHROS, they suggest first using excess wind power generation to drive the desalination process and then using the generated freshwater in the PSH system. The authors of [26] and [27] also feature the

optimization of a system that incorporates renewable energy, PSH, and desalination, but the desalination plant and the PSH plant are only integrated through a control system that dictates the the flow of renewable energy to each system and the grid. While related to these architectures, IPHROS represents a unique alternative architecture to these systems, and is advantageous to these other systems by means of the benefits listed above. At a reduced level, several studies of standalone renewable energy and PSH-based microgrids, ideal for small islands and remote regions without access to a utility grid, have been conducted, including, for instance, [28], and [29, 30], which consider the optimization of the system. While site selection for IPHROS will be the subject of future work, incorporating the production of fresh water by the system through reverse osmosis will only aid in their adoption in these locations. In a similar vein, other studies such as those in [31]-[32] look into the integration of renewable energy and reverse osmosis. While [31] and [33] are case studies, which again, will be addressed in future work, [32], [34], and [35] address the optimization of these renewable energy - reverse osmosis systems. Compared to these papers focusing more on the energy modeling and cost of the system, this paper takes a much deeper dive into modeling the element-level performance of the RO system, and considers factors such as the development of surrogate models for predicting membrane performance and including concentration polarization.

The performance metrics of RO membranes are commonly based on the solution-diffusion model for membrane flow [36], which expresses water and salt fluxes across the membrane as a function of concentration gradients, pressures, and osmotic pressures. This mass transport is difficult to model. Consequently, the seemingly standard convention for authors across the literature is to make varying degrees of assumptions and make use of the commercially available RO design software Water Application Value Engine (WAVE) [37] in the calculation of membrane properties and the development of their own models. In this paper, a new RO model for incorporation into IPHROS is detailed, which solves for the conditions of the RO system flows within the overarching MDO framework, utilizes WAVE in the development of surrogate models for the fractional salt rejection rate R_k and the permeate volumetric flowrate $Q_{p,k}$ (for RO element k), and is based on the solution-diffusion model with consideration of concentration polarization effects that occur along the surface of the RO membrane.

The remainder of this paper is as follows. Section 2 presents the methods used

in developing the IPHROS simulation, including an overview of the system and the governing equations not pertaining to the RO model. Section 3 then discusses the methods used in developing the RO model, including the governing equations utilized, the process of developing the surrogate models used in the RO model, and the incorporation of these equations into an RO algorithm and the overall IPHROS model itself. Section 4 presents the optimization methodology, and section 5 presents the results stemming from the IPHROS optimization and an economic assessment of the optimization results. Finally, Section 6 provides a discussion on the obtained results.

2. IPHROS Simulation Methodology

2.1. Overview

An overview of the IPHROS simulation model is detailed in the block diagram shown below in Figure 2. Coupled with Figure 1, these figures provide a picture of the flow of energy and water throughout the system. In this block diagram, IPHROS is decomposed into disciplinary modules containing the associated governing equations. To begin with, there is some amount of renewable energy produced on a given day, \dot{E}_r . A fraction, γ , of this energy, \dot{E}_{rp} , is sent to IPHROS, while the remaining energy, \dot{E}_{rd} , is sent directly to the consumer. \dot{E}_{rp} is used to power a pump that drives a flowrate of seawater, \dot{V}_{wp} , to an upper (mountaintop) reservoir at some height h_L . From the upper reservoir, a fraction, γ_{RO} , of \dot{V}_{wp} , $\dot{V}_{w,RO}$, is sent to the RO system. The remaining seawater, \dot{V}_{swht} , is sent to the main energy recovery turbine. The gravitational potential energy associated with \dot{V}_{swht} is then converted to the electrical energy \dot{E}_{swht} and sent to the consumer. In the RO System module, illustrated in Figure 3, a certain number of RO pressure vessels in parallel and RO elements in series in each of these pressure vessels, N_{pv1} and N_{e1} respectively, are selected to receive the incoming seawater. These are all in the first stage of the RO System, and depending on the RO system (determined by the algorithm described in Section 3.5), there will be another set of pressure vessels and elements, N_{pv2} and N_{e2} respectively, with a reconfigured feed flowrate into the second stage of \dot{V}_{int} . The resulting freshwater flowrate produced, $\dot{V}_{fw,RO}$, at an RO system recovery ratio of η_{RO} , is then sent to the consumer, while the energy potential of the still significantly pressurized brine stream, $\dot{V}_{o,RO}$, at a density and salinity of $\rho_{o,RO}$ and $S_{o,RO}$, is sent to an energy recovery turbine. The energy recovered from the brine stream, $\dot{E}_{ht,RO}$, is then sent to the consumer. Lastly, the two

remaining water streams, \dot{V}_{swht} and $\dot{V}_{o,RO}$, are mixed before the net discharge flows into the ocean. This final mixture has a salinity of S_{ht} . Additionally, all the energy that is sent to the consumer on a given day is summated as one energy term, \dot{E}_{er} .

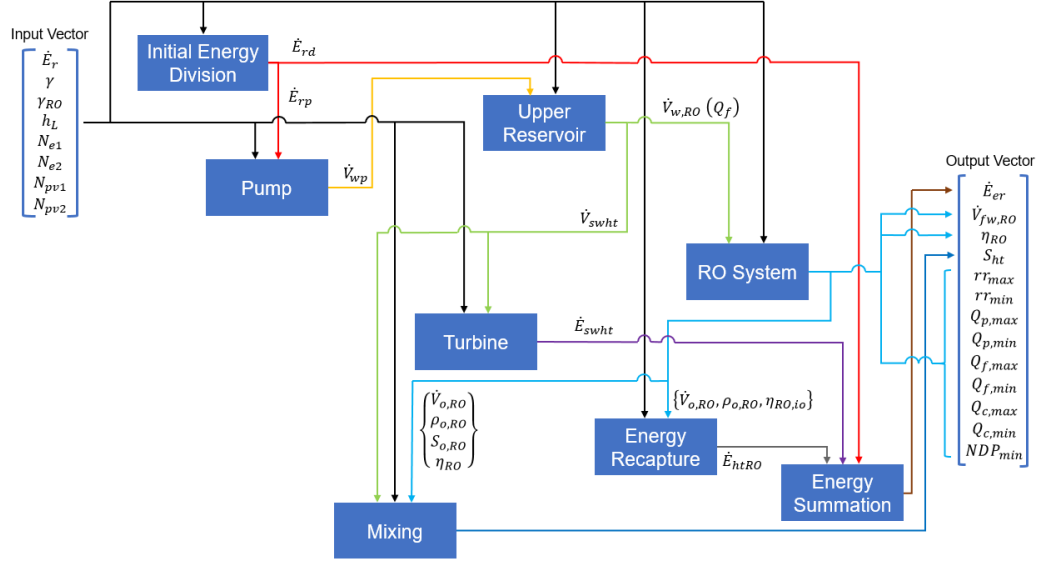


Figure 2: Block diagram of the IPHROS model. The min and max variables in the output vector represent the constraint values stemming from the RO system model.

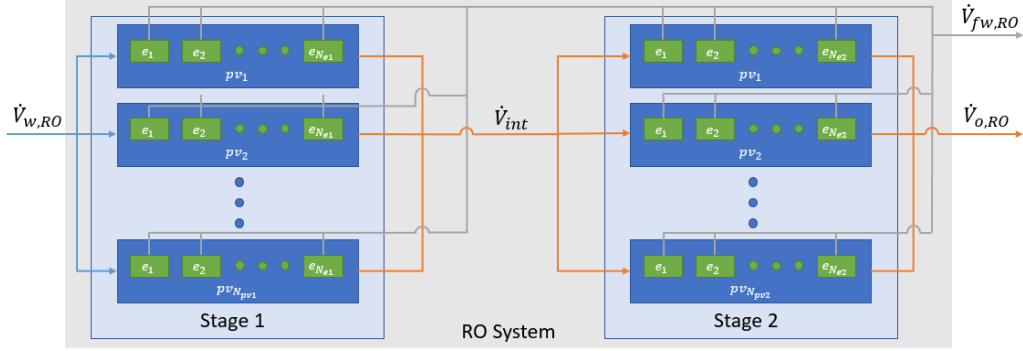


Figure 3: Illustrative view of the stages, pressure vessels, and elements within the RO system, as well as the flows between them. There can be anywhere between one and eight elements in a pressure vessel, and anywhere between 1 and 1.86×10^6 (explained in Section 4.2) pressure vessels in a stage. In the present model, there can be either one or two stages.

A full main table listing all variables relevant to IPHROS and their assignment to design, constraint, objective, parameter, and dependent variables is included in Appendix A, with an abridged version shown below in Table 1. As the system is currently formulated, there are eight design variables, which are either system operation variables (γ and γ_{RO}), or system design variables (\dot{E}_r , h_L , N_{e1} , N_{e2} , N_{pv1} , and N_{pv2}). These variables ultimately control the value of two obvious objectives, those being \dot{E}_{er} and $\dot{V}_{fw,RO}$. Additionally, the fresh water recovered from the seawater flowing through the RO system should ideally be done efficiently, both to minimize cost and to increase freshwater potential. This leads to a third objective being the RO system recovery ratio, η_{RO} , which is conventionally the focal variable in RO system optimization [38]. The remaining output variables are constraints on the system, to be described in Section 3.3. Most of the constraints in this section arise from this RO system model.

In the current model, the dynamic nature of IPHROS, such as varying consumer demands for fresh water and energy and the corresponding change in the design variables, is not considered, but will be explored in future work.

Table 1: An abridged version of the main table of all IPHROS variables, with each variable's name, units, variable type, and an associated value/range, if known. Variable bounds are discussed in Section 4.2. Included here are the design, constraint, and objective variables. See Appendix A for the dependent and parameter variables.

Variable	Name	Units	Known Value/Range	Type
\dot{E}_r	Renewable Energy Generated Per Day	kWh/day	$1 \leq \dot{E}_r \leq 100 \times 10^6$	Design
γ	Fraction of \dot{E}_r Sent to IPHROS	-	$0.01 \leq \gamma \leq 0.99$	Design
γ_{RO}	Fraction of Water in Upper Reservoir Sent to RO System	-	$0.01 \leq \gamma_{RO} \leq 0.99$	Design
h_L	Upper Reservoir Height	m	$240 \leq h_L \leq 821$	Design
N_{e1}	Number of RO Elements in Series in Stage 1	-	$1 \leq N_{e1} \leq 8$	Design
N_{e2}	Number of RO Elements in Series in Stage 2	-	$0 \leq N_{e2} \leq 8$	Design
N_{pv1}	Number of RO Pressure Vessels in Parallel in Stage 1	-	$1 \leq N_{pv1} \leq 1.86 \times 10^6$	Design
N_{pv2}	Number of RO Pressure Vessels in Parallel in Stage 2	-	$0 \leq N_{pv2} \leq 1.86 \times 10^6$	Design
S_{ht}	Discharge Salinity	g/kg	$S_{ht} \leq 40$	Constraint

NDP_k	Net Driving Pressure for RO Element k	Pa	$0 \leq NDP_k$	Constraint
$Q_{c,k}$	Concentrate Volumetric Flowrate for RO Element k	m^3/hr	$3.41 \leq Q_{c,k} \leq 15.5$	Constraint
$Q_{f,k}$	Feed Volumetric Flowrate for RO Element k	m^3/hr	$3.41 \leq Q_{c,k} \leq 15.5$	Constraint
$Q_{p,k}$	Permeate Volumetric Flowrate for RO Element k	m^3/hr	$0 \leq Q_{p,k} \leq 1.32$	Constraint
rr_k	Recovery Ratio for RO Element k	-	$0 \leq rr_k \leq 1$	Constraint
\dot{E}_{er}	Energy Sent to Consumer Per Day	kWh/day	-	Objective
$\dot{V}_{fw,RO}$	Freshwater Volumetric Flow Rate	m^3/day	-	Objective
η_{RO}	Net RO System Recovery Ratio	-	$0 \leq \eta_{RO} \leq 1$	Objective

2.2. Governing Equations

The governing equations for IPHROS that do not pertain to the RO system are detailed in this section. These are mostly drawn from [21], but with some slight modifications. First, a key equation will be defined that relates water quantities to energy quantities, which is critical for analysis because the amount of renewable energy produced is a design variable while flowrates are not. The energy sent to the pump, \dot{E}_{rp} , can be expressed either in terms of \dot{E}_r or the potential energy of \dot{V}_{wp} in the upper reservoir according to Equation 1,

$$\dot{E}_{rp} = \gamma \dot{E}_r = \frac{\rho_{sw} \dot{V}_{wp} g h_L}{3.6 \times 10^6 \eta_{hp}}, \quad (1)$$

where ρ_{sw} is the density of seawater, g is the gravitational acceleration, and h_L is the height of the upper reservoir. This equation accounts for losses on the pump-side of IPHROS, which is assumed to be constant at $\eta_{hp} = 89.4\%$ [21]. Additionally, the factor of 3.6×10^6 , here and for future equations, is a conversion factor for converting between joules and kilowatt-hours for consistent units. Solving for \dot{V}_{wp} provides the key equation relating energy quantities in the model to water quantities,

$$\dot{V}_{wp} = \frac{3.6 \times 10^6 \gamma \eta_{hp}}{\rho_{sw} g h_L} \dot{E}_r. \quad (2)$$

The governing equations for IPHROS are the equations for the three objectives noted in Table 1, those being $\dot{V}_{fw,RO}$, \dot{E}_{er} , and η_{RO} . The methods for calculating $\dot{V}_{fw,RO}$ and η_{RO} will be outlined in Section 3. The energy sent to the consumer, \dot{E}_{er} , is the sum of the energy sent directly to the consumer \dot{E}_{rd} , the energy recovered from the main seawater stream from the upper reservoir \dot{E}_{swht} , and the energy recovered from the brine stream $\dot{E}_{ht,RO}$,

$$\dot{E}_{er} = \dot{E}_{rd} + \dot{E}_{swht} + \dot{E}_{ht,RO}. \quad (3)$$

Looking at each right hand side term of Equation 3 individually, \dot{E}_{rd} is the remainder of the fraction γ of renewable energy \dot{E}_r that is sent to IPHROS,

$$\dot{E}_{rd} = (1 - \gamma)\dot{E}_r. \quad (4)$$

To express \dot{E}_{swht} in terms of design variables, it must first be expressed in terms of flowrates. This is done by relating \dot{E}_{swht} to the energy potential of the energy recovery stream from the upper reservoir, accounting for the turbine side efficiency η_{ht} ,

$$\dot{E}_{swht} = \frac{\rho_{sw}\dot{V}_{swht}gh_L\eta_{ht}}{3.6 \times 10^6}. \quad (5)$$

Equation 5 can be expanded to account for the fact that \dot{V}_{swht} is a fraction of the water in the upper reservoir,

$$\dot{V}_{swht} = (1 - \gamma_{RO})\dot{V}_{wp}. \quad (6)$$

Substituting Equation 2 into Equation 6 results in Equation 7,

$$\dot{V}_{swht} = \frac{3.6 \times 10^6(1 - \gamma_{RO})\gamma\eta_{hp}\dot{E}_r}{\rho_{sw}gh_L}, \quad (7)$$

which can then be substituted into Equation 5 to obtain the final expression for \dot{E}_{swht} ,

$$\dot{E}_{swht} = (1 - \gamma_{RO})\gamma\eta_{hp}\eta_{ht}\dot{E}_r. \quad (8)$$

A similar process can be applied to express an equation for $\dot{E}_{ht,RO}$. Beginning with the energy potential of the brine stream and scaling by the turbine side efficiency η_{ht} and the percent reduction in the pressurization of the brine stream relative to the feed conditions of the RO system $\eta_{RO,io}$,

$$\dot{E}_{ht,RO} = \frac{\rho_{o,RO}\dot{V}_{o,RO}gh_L\eta_{ht}\eta_{RO,io}}{3.6 \times 10^6}. \quad (9)$$

$\rho_{o,RO}$ and $\eta_{RO,io}$ will be taken care of by the RO model. $\dot{V}_{o,RO}$ can be expanded as follows in Equation 10,

$$\begin{aligned}\dot{V}_{o,RO} &= (1 - \eta_{RO})\dot{V}_{w,RO} \\ &= (1 - \eta_{RO})\gamma_{RO}\dot{V}_{wp},\end{aligned}\tag{10}$$

where η_{RO} is the net recovery ratio of the RO system and γ_{RO} is the fraction of the water in the upper reservoir that is sent to the RO system. Substituting Equation 2 into Equation 10, and then that into Equation 9, results in the final expression for $\dot{E}_{ht,RO}$,

$$\dot{E}_{ht,RO} = \frac{\rho_{o,RO}}{\rho_{sw}}\gamma\gamma_{RO}(1 - \eta_{RO})\eta_{RO,io}\eta_{hp}\eta_{ht}\dot{E}_r.\tag{11}$$

Substituting Equations 4, 8, and 11 into Equation 3 gives the final equation for \dot{E}_{er} as

$$\dot{E}_{er} = \left((1 - \gamma) + (1 - \gamma_{RO})\gamma\eta_{hp}\eta_{ht} + \frac{\rho_{o,RO}}{\rho_{sw}}\gamma\gamma_{RO}(1 - \eta_{RO})\eta_{RO,io}\eta_{hp}\eta_{ht} \right) \dot{E}_r.\tag{12}$$

In addition to the above governing equations, there is an additional variable that serves as one of the main constraints on IPHROS thus far, that being the salinity of the final output. Looking at the final flow, equations for the density and salinity of the discharge from IPHROS into the ocean, ρ_{ht} and S_{ht} , can be derived from conservation of mass analyses on the salt and pure water mass flowrates into and out of the Mixture module. Beginning with the salt,

$$\sum \dot{m}_{s,in} = \sum \dot{m}_{s,out}.\tag{13}$$

Expressing the mass flowrate of salts in Equation 13 as the product of the density, flowrate, and salinity of the different water streams entering and exiting the Mixing module,

$$\rho_{sw}\dot{V}_{swht}S_{sw} + \rho_{o,RO}\dot{V}_{o,RO}S_{o,RO} = \rho_{ht}\dot{V}_{ht}S_{ht},\tag{14}$$

where the first term in the left-hand side represents the salt entering from the main energy recovery stream, the second term in the left-hand side represents the salt entering from the brine stream, and the first term in the right-hand side represents the salt leaving in the mixed discharge. \dot{V}_{ht} can be expanded as the sum of \dot{V}_{swht} and $\dot{V}_{o,RO}$, and after rearranging Equation 14,

$$\dot{V}_{swht}(\rho_{ht}S_{ht} - \rho_{sw}S_{sw}) = \dot{V}_{o,RO}(\rho_{o,RO}S_{o,RO} - \rho_{ht}S_{ht}). \quad (15)$$

Substituting Equations 6 and 10 into Equation 15 and rearranging,

$$(1 - \gamma_{RO})(\rho_{ht}S_{ht} - \rho_{sw}S_{sw}) = (1 - \eta_{RO})\gamma_{RO}(\rho_{o,RO}S_{o,RO} - \rho_{ht}S_{ht}). \quad (16)$$

From the variables presented in Equation 16, ρ_{sw} and S_{sw} are parameters, γ_{RO} is a design variable, and $\rho_{o,RO}$, $S_{o,RO}$, and η_{RO} will be determined by the RO model, to be detailed in Section 3. Next, one more equation is needed such that S_{ht} (and ρ_{ht}) can be solved for, which is attainable from a conservation of mass analysis on the fresh water entering and leaving the Mixing module. Similarly to Equation 13,

$$\sum \dot{m}_{fw,in} = \sum \dot{m}_{fw,out}. \quad (17)$$

To look at the mass flowrate of fresh water rather than the mass flowrate of salt, for the salt-containing streams, the quantity $1000 - S$ can be used due to the assumption that salt is the only particulate in the seawater. This results in Equation 18,

$$\rho_{sw}\dot{V}_{swht}(1000 - S_{sw}) + \rho_{o,RO}\dot{V}_{o,RO}(1000 - S_{o,RO}) = \rho_{ht}\dot{V}_{ht}(1000 - S_{ht}), \quad (18)$$

where, like for Equation 14, the first term in the left-hand side represents the fresh water entering from the main energy recovery stream, the second term in the left-hand side represents the fresh water entering from the brine stream, and the first term in the right-hand side represents the fresh water leaving in the mixed discharge. Again using the fact that \dot{V}_{ht} is the sum of

\dot{V}_{swht} and $\dot{V}_{o,RO}$ and Equations 6 and 10, Equation 18 can be rearranged and simplified to Equation 19,

$$(1 - \gamma_{RO})(\rho_{ht}(1000 - S_{ht}) - \rho_{sw}(1000 - S_{sw})) = (1 - \eta_{RO})\gamma_{RO}(\rho_{o,RO}(1000 - S_{o,RO}) - \rho_{ht}(1000 - S_{ht})). \quad (19)$$

With two equations and two unknowns, S_{ht} can be derived from solving Equations 16 and 19 for ρ_{ht} and S_{ht} . The resulting equation for S_{ht} is shown below in Equation 20,

$$S_{ht} = \frac{S_{sw}\rho_{sw}(1 - \gamma_{RO}) + \gamma_{RO}S_{o,RO}\rho_{o,RO}(1 - \eta_{RO})}{\rho_{sw}(1 - \gamma_{RO}) + \gamma_{RO}\rho_{o,RO}(1 - \eta_{RO})}. \quad (20)$$

3. Improved RO Modeling Methodology

3.1. RO theory overview

A wide variety of membrane transport models have been developed to characterize the flow of water and solutes across osmosis membranes [39]. Of these models, the solution-diffusion model has been utilized frequently throughout the literature. A complete review of the solution-diffusion model has previously been conducted in [36].

What results is the quintessential equation for characterizing an RO element's permeate flowrate according to the solution-diffusion model, obtained by equating the chemical potentials at each water-membrane interface and utilizing Fick's first law of diffusion,

$$Q_p = A_w A_m \times NDP, \quad (21)$$

where Q_p is the permeate flowrate, A_w is the membrane water permeability coefficient, A_m is the membrane surface area of the RO element, and NDP is the net driving pressure across the membrane. Equation 21 assumes that there is no membrane fouling, which would inhibit flow across the membrane, and that the temperature is standard at 25°C. Otherwise, there would be additional fouling factors and temperature correction factors to include in Equation 21 [40].

A significant factor in the realization of flowrates across RO membranes that the solution-diffusion model does not account for is boundary layer effects along the feed side of the membrane. This phenomenon is referred to as concentration polarization. As seawater diffuses across the membrane, the salt in the seawater does not enter the membrane as readily as the water does. This difference creates a concentration gradient between the boundary layer region adjacent to the membrane and the main feed flow, resulting in salt diffusion away from the membrane, hindering flowrates across the membrane [41]. Assuming steady state allows for the salt concentration at the feed side of the membrane to be known.

Figure 4 shows approximate profiles of the water and salt concentrations, c_w and c_s respectively, inside and near the membrane, which is of thickness $x = l$. On the feed side of the membrane there is a boundary layer of thickness δ , where the effects of concentration polarization are depicted with an increase in c_s and a decrease in c_w before $x = 0$.

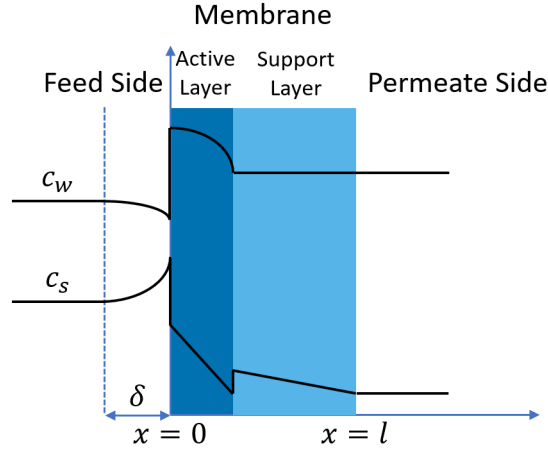


Figure 4: Water and salt concentration (c_w and c_s) profiles close to and within the RO membrane, incorporating the boundary layer δ that forms along the feed side of the membrane before $x = 0$. The membrane overall is of thickness l , with a thinner active layer followed by a thicker support layer. This figure is adapted from [41]. Profiles shown are approximate.

3.2. Model overview and assumptions

Figure 5 shows a graphical interpretation of an RO element and the significant variables that are being tracked in the model.

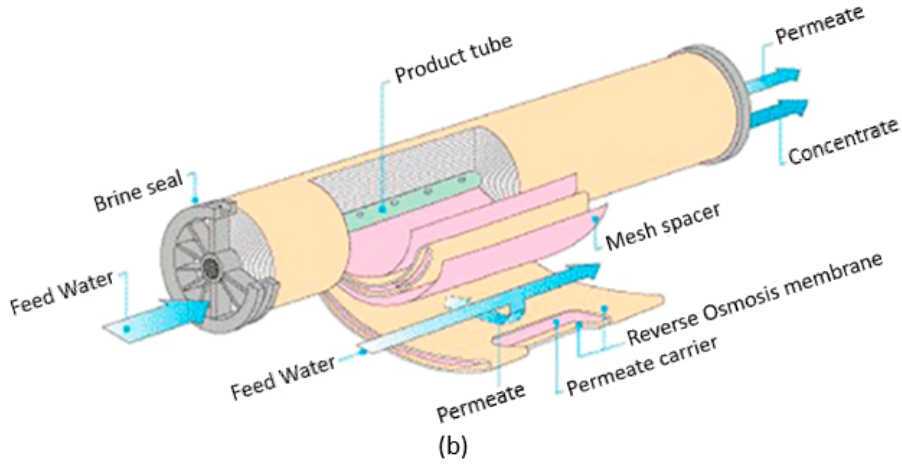
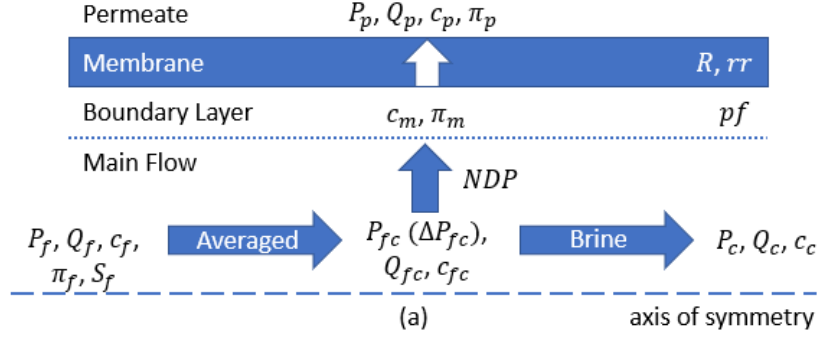


Figure 5: RO element diagrams. (a) is a cross-sectional view of the flow(s) between two membrane layers in a spiral wound RO element, showing the general groupings/regions and flow of variables within the presented RO model. The k subscripts relative to the main table in Appendix A are dropped due to one element being implied by this figure. The subscripts f , fc , c , m , and p refer to the feed, averaged, concentrate, boundary layer, and permeate regions respectively. To contextualize with a spiral-wound membrane RO element, (b) is an illustration of one from [42].

From Figure 5a, there are five “regions” of variables in this RO model. The grouping of the feed pressure P_f , feed volumetric flowrate Q_f , feed salt concentration c_f , feed osmotic pressure π_f , and feed salinity S_f are the feed conditions of the saline flow into the RO element, and the grouping of the concentrate pressure P_c , concentrate volumetric flowrate Q_c , and concentrate salt concentration c_c are the exit conditions of the more saline flow out of the RO element. Because mass transport across the membrane is occurring

along the length of the RO element, averaged values for the salt concentration and flow rate in the RO element are adopted by P_{fc} (a function of the pressure drop along the length of the membrane ΔP_{fc}), Q_{fc} , and c_{fc} . Concentration polarization pf occurs within the boundary layer that develops on the feed side of the membrane as the main flow is forced towards and across the membrane with net driving pressure NDP , resulting in the “membrane” region with osmotic pressure π_m and salt concentration c_m . As the flow is forced across the membrane, mass transport is dictated by the fractional salt rejection rate R and the membrane water permeability coefficient A_w , and, more holistically, the recovery ratio rr , resulting in the grouping of variables on the permeate side of the membrane consisting of the permeate pressure P_p , permeate volumetric flowrate Q_p , permeate salt concentration c_p , and permeate osmotic pressure π_p . These are the variables that are calculated for every element k in the RO model. In future sections, the subscript k will be used to denote the RO element number.

A list of the assumptions that are made in this model are included below in Table 2. The permeate pressure, P_p , is assumed to be equal to 1.01325×10^5 Pa (1 atm), and the salinity of seawater, S_{sw} , is assumed to be 35 g/kg. At a temperature, T , of 25°C, this corresponds to a seawater density, ρ_{sw} , of 1023.6 kg/m³. The density of fresh water, ρ_{fw} , being 996.9 kg/m³ corresponds to a temperature of 25°C as well, and the molar mass of salt, M_{salt} assumes that the only particulates in the seawater are Na and Cl.

Table 2: Parameter assumptions made in the RO model.

Variable	Value	Units
P_p	1.01325×10^5 (1)	Pa (atm)
S_{sw}	35	g/kg
ρ_{sw}	1023.6	kg/m ³
ρ_{fw}	996.9	kg/m ³
T	25	°C
M_{salt}	58.44	g/mol

One thing worthwhile noting is that the density of the permeate flow from each element is not equal to ρ_{fw} , since there are still slight quantities of salt that pass through the membrane. To ensure high-quality drinking water, salt

concentrations would have to be tracked, which is something to consider in future work.

3.3. Governing equations

Below are the key governing equations that define this RO model, and are applicable to any element in an arbitrary position in the RO system. This is specified by the subscript k within all subsequent equations. Rewriting Equation 21 in terms of k ,

$$Q_{p,k} = A_{w,k} A_m \times NDP_k, \quad (22)$$

In the RO model presented in this paper, Equation 22 will not be explicitly solved for in order to attain the permeate flowrate for an RO element. Instead, a surrogate model was developed for calculating $Q_{p,k}$ for some input flow into an RO element, and is detailed in Section 3.4. This also allows for the determination of $A_{w,k}$ to be circumvented. However, from Equation 22, NDP_k still needs to be calculated because for any flow configuration, the net driving pressure must be greater than or equal to zero for reverse osmosis to occur. The net driving pressure is the difference between the nominal pressure difference across the membrane and the osmotic pressure difference across the membrane, which expanded even further results in Equation 23,

$$NDP_k = (P_{fc,k} - P_{p,k} - \pi_{m,k} + \pi_{p,k}), \quad (23)$$

where, all for RO element k , $P_{fc,k}$ is the average pressure of the bulk flow along the length of an RO element, $P_{p,k}$ is the permeate pressure, $\pi_{m,k}$ is the osmotic pressure along the feed side of the membrane, and $\pi_{p,k}$ is the osmotic pressure along the permeate side of the membrane.

For applications to the reverse osmosis of seawater, the following equation is a widely used approximation for calculating the osmotic pressure in a given region of an RO element [41], using a factor of 6894.76 to convert from psi to Pa,

$$\pi_{region,k} = 6894.76 \times \left(1.12(273 + T) \sum \bar{m}_{i,region,k} \right), \quad (24)$$

where $\pi_{region,k}$ is the osmotic pressure for a region in RO element k . These regions include the feed region (f), permeate region (p), averaged feed region (fc), concentrate region (c), and boundary layer region (m). T is the temperature, and $\bar{m}_{i,region,k}$ is the molarity of a dissolved ionic or nonionic species in seawater. However, assuming that the salt in seawater is strictly NaCl and with concentrations in this paper being defined in mol/L, Equation 24 simplifies to

$$\pi_{region,k} = 6894.76 \times (1.12(273 + T)(2 * c_{region,k})), \quad (25)$$

where $c_{region,k}$ is the salt concentration in RO element k for a given region of the RO element.

The FilmTec™ Seamaxx™-440 Element RO membrane made by Dupont is used in this analysis due to its significantly higher flowrates than other seawater RO membranes and therefore its lower energy consumption [43]. The properties for this element are shown below in Table 3.

Table 3: Seamaxx™-440 element properties.

Variable	Lower Bound	Value	Upper Bound	Units
rr	0	-	0.13	-
Q_f, Q_c	3.41	-	15.5	m ³ /hr
Q_p	0	-	1.32	m ³ /hr
A_m	-	40.9 (440)	-	m ² (ft ²)

For 8-inch diameter FilmTec™ membranes, which the Seamaxx™-440 is, the concentration polarization factor pf_k for RO element k can be approximated, as is done by [40], as

$$pf_k = e^{0.7rr_k}, \quad (26)$$

where rr_k is the recovery ratio for that element. A similar approach, using [40], can be used for approximating the pressure drop along element k , $\Delta P_{fc,k}$,

$$\Delta P_{fc,k} = 0.01(4.403Q_{fc,k})^{1.7}, \quad (27)$$

where $Q_{fc,k}$ is the average flowrate along RO element k and 4.403 is a unit conversion from m^3/hr to gpm . The feed salt concentration for the first RO element of the first stage can be derived from known quantities by dimensional analysis,

$$c_{f,(k=1)} = \frac{S_{sw}\rho_{sw}}{1000M_{salt}}, \quad (28)$$

where S_{sw} is the salinity seawater, ρ_{sw} is the density of seawater, and M_{salt} is the molar mass of salt (NaCl). The permeate salt concentration for each individual element, $c_{p,k}$, is related to the fractional salt rejection rate of the membrane for that element, R_k , and to the feed salt concentration for an element $c_{f,k}$ by

$$c_{p,k} = c_{f,k}(1 - R_k). \quad (29)$$

The fractional salt rejection rate will be derived in Section 3.4. The following equation for the concentrate salt concentration for RO element k , $c_{c,k}$, is derived from a conservation of mass analysis on an RO element, solving for the salt concentration of the concentrate discharge from an RO element,

$$c_{c,k} = \frac{Q_{f,k}c_{f,k} - Q_{p,k}c_{p,k}}{Q_{c,k}}, \quad (30)$$

where $Q_{c,k}$ is the difference between the feed flowrate and permeate flowrate for RO element k ,

$$Q_{c,k} = Q_{f,k} - Q_{p,k}. \quad (31)$$

The average salt concentration in the bulk flow along the length of an RO element is approximated as a linear average of the inlet and outlet salt concentrations of an RO element,

$$c_{fc,k} = \frac{c_{f,k} + c_{c,k}}{2} \quad (32)$$

A similar approach can be taken for the average flowrate of the bulk flow along the length of an RO element,

$$Q_{fc,k} = \frac{Q_{f,k} + Q_{c,k}}{2}. \quad (33)$$

While Equation 26 already defines pf_k as a function of rr_k , pf_k can also be expressed as the ratio of the difference in salt concentrations between the membrane surface and the permeate to the difference in salt concentrations between the bulk flow and the permeate [42]. This equation can then be rearranged to provide an equation for the salt concentration along the membrane,

$$c_{m,k} = pf_k(c_{fc,k} - c_{p,k}) + c_{p,k}. \quad (34)$$

The recovery ratio for an element, rr_k , is defined as the permeate flowrate divided by the feed flowrate for that element,

$$rr_k = \frac{Q_{p,k}}{Q_{f,k}}. \quad (35)$$

Lastly, the pressure of the concentrate flow leaving RO element k is the difference between the feed pressure and the pressure drop that occurs along the length of that element,

$$P_{c,k} = P_{f,k} - \Delta P_{fc,k}, \quad (36)$$

and the average feed pressure of the bulk flow along the length of an RO element is approximated, similarly to Equation 32, as a linear average of the inlet and outlet pressures of an RO element,

$$P_{fc,k} = \frac{P_{f,k} + P_{c,k}}{2} \quad (37)$$

3.4. Surrogate models for Seamaxx™-440 membrane permeate flowrate and fractional salt rejection rate

The permeate flowrate, Q_p , and the fractional salt rejection rate, R , are both variables intrinsic to a membrane, as they reflect how much water passes through, and how much salt does not pass through, a membrane. Typically,

R is calculated experimentally [39], with the membrane water permeability coefficient, A_w , also calculated experimentally and used in concert with Equation 21 to obtain Q_p . However, a suitable alternative is achieved by running experimental simulations through the WAVE design software, allowing for surrogate models to be constructed for R and Q_p to predict their performances for varying feed conditions. There are three variables that together characterize the feed flow into an RO element, which include the feed pressure, P_f , the feed salinity, S_f , and the feed flowrate, Q_f . From [43], the Seamaxx™-440 has a maximum operating pressure of 8.274×10^6 Pa (1200 psi). Previous experience with running simulations on the Seamaxx™-440 element in WAVE revealed 2.068×10^6 Pa (300 psi) to be an approximate lower bound for the feed pressure, depending on the other feed conditions. Since seawater is the only feed into the RO system as the system is currently constructed, the lower bound for S_f is S_{sw} , assumed in Section 3.2 to be equal to 35 g/kg. An upper bound of 55 g/kg for these simulations was selected to account for the increasing salinity of the bulk flow as it travels through multiple RO elements. Lastly, as noted in Section 3.3, the lower and upper bounds on Q_f are 3.41 m³/hr and 15.5 m³/hr respectively.

Knowing these design variable ranges, a Latin hypercube sampling (LHS) of 50 points in the design space was first conducted to see if any trends relating to Q_p and R were easily discernible. LHS was achieved making use of the lhsmdu Python package from [44, 45], which generates LHS with multi-dimensional uniformity. Each point generated by lhsmdu together forms an LHS sampling distributions from the feed condition ranges detailed above, and is the equivalent of a simulation in WAVE with unique feed conditions.

Within WAVE, inlet conditions are specified for a pre-pressurized state, with the assumption that a pump will then pressurize the water to the desired level. However, for the IPHROS model, the seawater entering the RO system is already pressurized due the gravitational potential energy of the water in the upper reservoir. This required some slight manipulation, as WAVE does not assume that seawater is incompressible. Salinity units in WAVE being in mg/L, instead of the more standard g/kg units used in this paper, required a conversion between the two salinities,

$$S_{mgL} = S_{gkg} \times \rho_w, \quad (38)$$

where S_{mgL} is the salinity in units of mg/L, S_{gkg} is the salinity in units of g/kg, and ρ_w is the density of the saline water in units of kg/m³. Correlation equations from [46] were utilized for calculating the density of pressurized seawater from its salinity. With the desired S_f and Q_f being target values for post-pressurization in WAVE, input S_f and Q_f values were selected such that the desired values were achieved.

In WAVE, the feed water flowrate and water type (seawater) are specified, and the feed salinity is specified with the assumption that NaCl is the only salinity-contributing component in the water. The flow factor is assumed to be 1 to emulate “fresh-out-of-the-box” membrane conditions, the temperature is specified to 25°C, and the pre-stage change in pressure is assumed to be 0. Simulations were run on one Seamaxx™-440 element in one pressure vessel, with one stage and one pass. All other configurable inputs into WAVE were kept at default values.

Despite the allowable ranges for P_f , S_f , and Q_f being specified above, there are interaction effects between the feed conditions that can trigger an “RO Design Warning” message in WAVE, which, for this paper, signifies an infeasible set of input conditions. For instance, a feed pressure of 7.584×10^6 Pa (1100 psi) cannot be achieved at a feed flowrate of 14 m³/hr, as it results in an infeasible, or “illegal,” simulation. This event is triggered when operating conditions are at levels that would be detrimental to the health and durability of the membrane, such those causing delamination [40]. Possible error messages that can be raised include the average net driving pressure being less than the minimum limit, the recovery ratio being greater than the maximum limit, the concentrate flowrate being less than the minimum limit, and the permeate flowrate being greater than the maximum limit.

From running the 50 simulations in WAVE, while no immediate trends were apparent for Q_p and R relative to any single feed condition, what became clear were two distinct clusters of data points, shown below in Figure 6. Using linear Support Vector Machine (SVM) classification, a 2D plane was generated that roughly separates the two clusters, which was then shifted in a parallel direction so that all of the “legal” LHS points were contained on one side of the plane. These planes are also included in Figure 6, with the resulting plane being defined by Equation 39,

$$-0.00264288P_f + 282.748S_f + 200Q_f = -1334.934. \quad (39)$$

The left hand side of Equation 39 being greater than the right hand side signifies a feed flow condition that will (most likely) not trigger an “RO Design Warning”. Setting the granularity of P_f to 1.724×10^5 Pa (25 psi), S_f to 2 g/kg, and Q_f to 1 m³/hr results in a reduced full factorial search of the RO design space of 3363 simulations, a 36.4% reduction from the 5291 simulations that would have had to have been completed if the original variable bounds defined earlier were taken as absolute. From these 3363 simulations, removing the “illegal” WAVE simulations and any simulations that resulted in a permeate flowrate of zero due to rounding within WAVE from future analysis resulted in a total of 2507 simulations for identifying equations for R and Q_p .

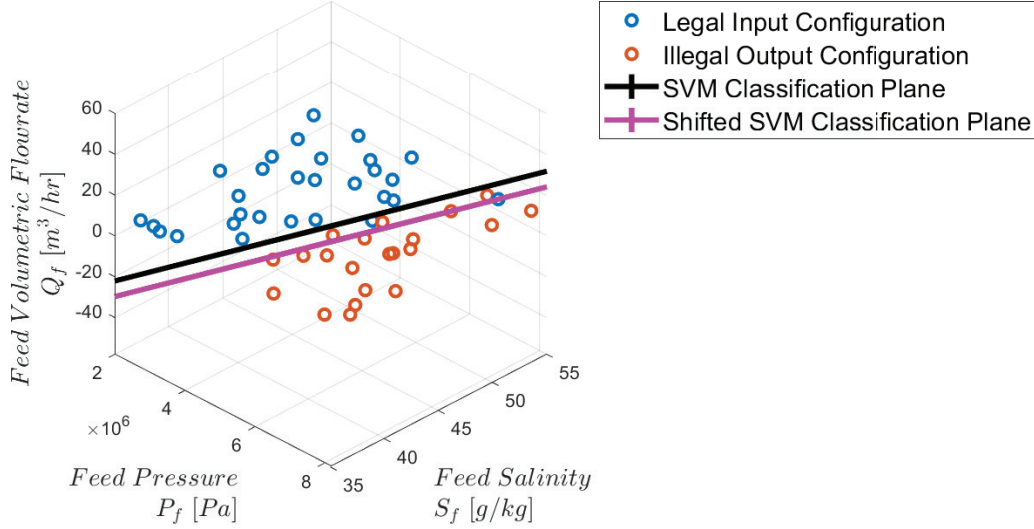


Figure 6: Clustering among WAVE simulations from LHS sampling, with a shifted 2D plane separating “legal” simulations from “illegal” simulations. An illegal simulation is one that triggered an “RO Design Warning” notification in WAVE.

Once simulations were complete, data processing was conducted using the JMP software [47], a computer program suite for statistical analysis, and MATLAB. Before beginning the analysis, from the 23 variables that were either recorded directly from the WAVE simulations or calculated dependently from those values, variables were removed from model consideration that were either the same variable as another but in different units, tracked a similar metric to another variable, or were variables that are linearly de-

pendent on other variables. This resulted in a list of 15 variables: P_f , Q_f , ρ_{sw} , rr , Q_c , P_c , c_c , c_f , c_p , pf , c_m , ΔP , Q_p , R , and A_w .

Beginning with R , bivariate scatter plots were generated in JMP, included in Appendix B, to visually investigate if any correlations, linear or nonlinear, were clearly evident between R and another variable. A distinct trend was revealed between R and Q_p , shown below in Figure 7.

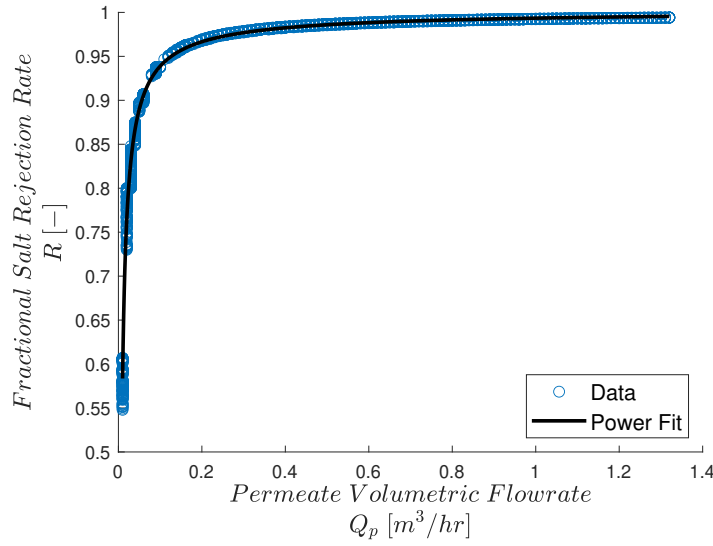


Figure 7: Scatterplot of R vs. Q_p , where the equation for the power fit is noted in Equation 40. Data points exceeding Cook's distance were not included in this figure.

This strong nonlinear correlation is supported by observing Kendall's τ coefficient, a nonparametric correlation metric between two variables based on the rank of the variable's values. A Kendall's τ of 1 indicates a positive correlation where the rank of x_i and y_i for (x_i, y_i) is the same for all i . A Kendall's τ of -1 indicates a negative correlation where the rank of x_i and y_i for (x_i, y_i) is the inverse for all i . Conversely, a Kendall's τ of 0 means there is no relationship in the rank of (x_i, y_i) . The Kendall's τ for R and Q_p was 0.9739, a strong indicator of high correlation between the two variables. This is further supported by JMP's Predictor Screening method, which uses bootstrap forest partitioning to determine the proportion of a response that

is caused by a certain predictor. Using this method with the number of trees set to 100 revealed that approximately 72% of the values of R can be attributed to Q_p .

Fitting a curve to the data points in Figure 7 was done by evaluating the R^2 value for three different models: a five parameter bi-exponential model, a power model, and a rational model with a numerator degree of 2 and a denominator degree of 1. The results are shown below in Table 4.

Table 4: R^2 values for three different models fitted to R vs. Q_p . In the General Form column, a , b , c , d , and f are coefficients.

Model	General Form	R^2 Value
Biexponential 5p (a,b,c,d,f)	$a + be^{-cx} + de^{-fx}$	0.9727
Power	$a + bx^c$	0.9734
Rational	$\frac{ax^2+bx+c}{x+d}$	0.9732

From these three models, the power model had the best fit to the data with an R^2 value of 0.9734. To minimize the effect that outliers have on the parameters, Cook's distance was utilized to identify overly influential outliers in the data. Removing these data points from consideration in parameter tuning results in a stronger fit of the model to the majority of the data. These points were also excluded from Figure 7. Eliminating data points with a Cook's distance greater than $4/n$, which in this case is approximately 0.0015955, increases the R^2 value of the power model to 0.9939, and results in Equation 40,

$$R = 1.0034 - 0.00997Q_p^{-0.8122}. \quad (40)$$

It is worthwhile noting that c_p also has a relatively strong correlation with R , with a Kendall's τ of -0.9113 and a 25% contribution to the values of R from the predictor screening mentioned above. However, with $R = 1 - c_p/c_f$ being the definition of R , it is not advantageous to use c_p in a univariate equation for R . Using Q_p , aside from having the best Kendall's τ and predictor contribution, relates R to an additional variable.

For developing an equation for Q_p , no such bivariate trends were apparent in the bivariate scatterplots mentioned above. While a multiple linear regression

approach could have been utilized to find an explicit equation for Q_p and was considered, a neural network approach for determining Q_p was instead decided upon. This allows for Q_p to be predicted directly from the input flow conditions P_f , S_f , and Q_f , in a sense mimicking WAVE in a way that is integrable with a programming language.

To develop the neural network for Q_p , the Automated Machine Learning (AutoML) package AutoKeras [48] was used. This allowed for an optimized, automated selection of neural network features, such as the architecture, model hyperparameters, and activation functions, to take place, with the end goal of arriving at a neural network that will not overfit or underfit the data. From the 2507 WAVE simulations, 80% of the simulations were randomly assigned as the training set, with the remaining 20% set aside as the test set. Using the AutoKeras structured data regression class, P_f , S_f , and Q_f were set as numerical inputs to the input layer. The mean squared error loss function was selected for error minimization in network development. The number of max trials, or the maximum number of different models tried, was set to 100, and the number of epochs, or the number of forward and back propagations done per model iteration in generating the weights for the model, was set to 1000. All other configurable variables were kept at default values.

A summary of the neural network generated by AutoKeras illustrating the forward propagation of the inputs P_f , S_f , and Q_f to the output Q_p is shown below in Figure 8. Beginning with the input layer, which is composed of three nodes (one for each RO input), the inputs get fed into a normalization layer that standardizes each feature with the mean and standard deviation of each feature. From there, for each node in the hidden layers, the inputs are scaled by weights generated by AutoKeras and added together as a weighted sum (the bias nodes, denoted by β , represent a weighted input of 1 into a layer). The resulting weighted sum gets passed into an activation function, which for this neural network are all ReLU activation functions. The three hidden layers, excluding the bias nodes, have 16, 512, and 128 nodes respectively. The output of the network, Q_p , is a weighted sum of the outputs from the final hidden layer.

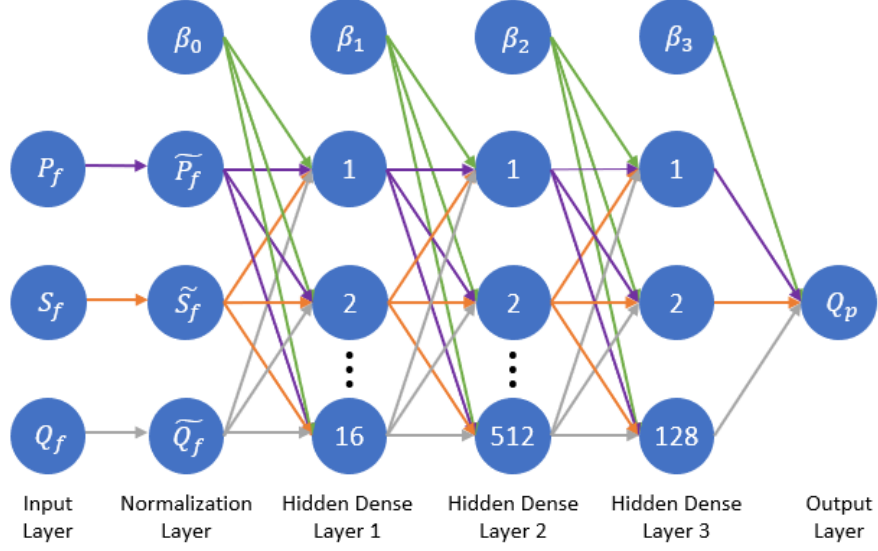


Figure 8: Diagram of neural network for Q_p generated by AutoKeras. There are six total layers, including an input layer with three nodes, a normalization layer with three nodes, three hidden dense layers with 16, 512, and 128 nodes respectively, and an output layer with one node. The normalization layer does feature standardization on the input layers, and each numbered node in the hidden layers uses a ReLU activation function. The β nodes are bias nodes that inject a weighted input of one into the downstream layer.

Using the test set to assess the accuracy of the neural network in accurately predicting Q_p reveals an R^2 value of 0.99985 and an RSME of 0.0046995. Another way of interpreting the error of the model is by looking at a histogram of the percent error distribution of the Q_p values predicted by the neural network to the corresponding Q_p values from WAVE, shown below in Figure 9. Of the WAVE simulations in the test set, approximately 70% of the simulations have a percent error less than 5%, and approximately 84% of the simulations have a percent error less than 10%. It should be noted that approximately 1% of the simulations have a percent error greater than 25%, which could possibly have a slight impact on the reliability of the predicted Q_p for an RO element. These more erroneous predictions could possibly be resolved by the use of a different machine learning technique. Ultimately, however, what these error metrics indicate is that, for the vast majority of input flow conditions, the neural network obtained from AutoKeras for Q_p is an accurate model that should be sufficient for accurately modeling RO

membrane performance in the current IPHROS.

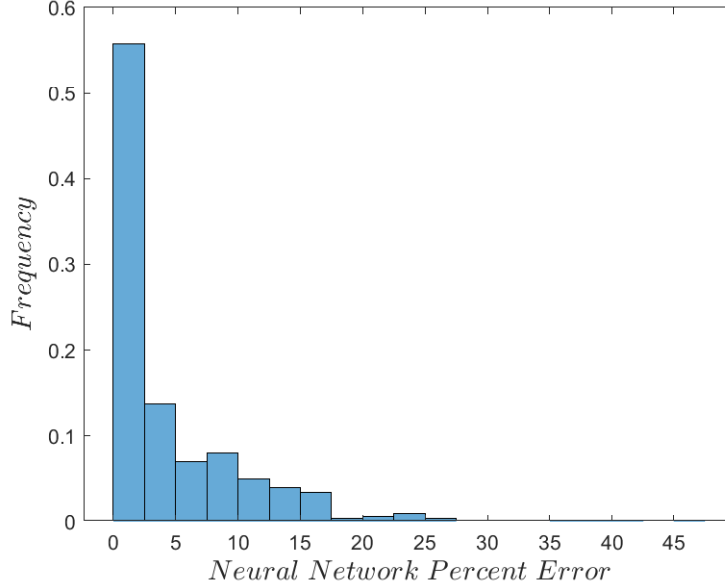


Figure 9: Histogram showing the distribution of the percent error between the Q_p values generated by the neural network and the corresponding values from WAVE. Approximately 70% of the simulations from the test set have a percent error less than 5%, and approximately 84% of the simulations from the test set have a percent error less than 10%.

The methods in this section are all in contrast to [49], for which simulations were run for three SeamaxxTM-440 elements in series in one pressure vessel, feed conditions were adjusted each simulation (increase in feed salinity matched with a decrease in the feed flowrate) such that the permeate flowrate through the first element was kept at the highest possible value of 1.32 m³/hr, and the RO system recovery ratio was set at 25%. Furthermore, the RO model in [49] predicted A_w instead of Q_p , and assumed that changing the net system recovery does not alter the previously obtained equations for A_w and R .

3.5. RO model algorithm

A visualization of the RO model algorithm is shown below in Figure 10.

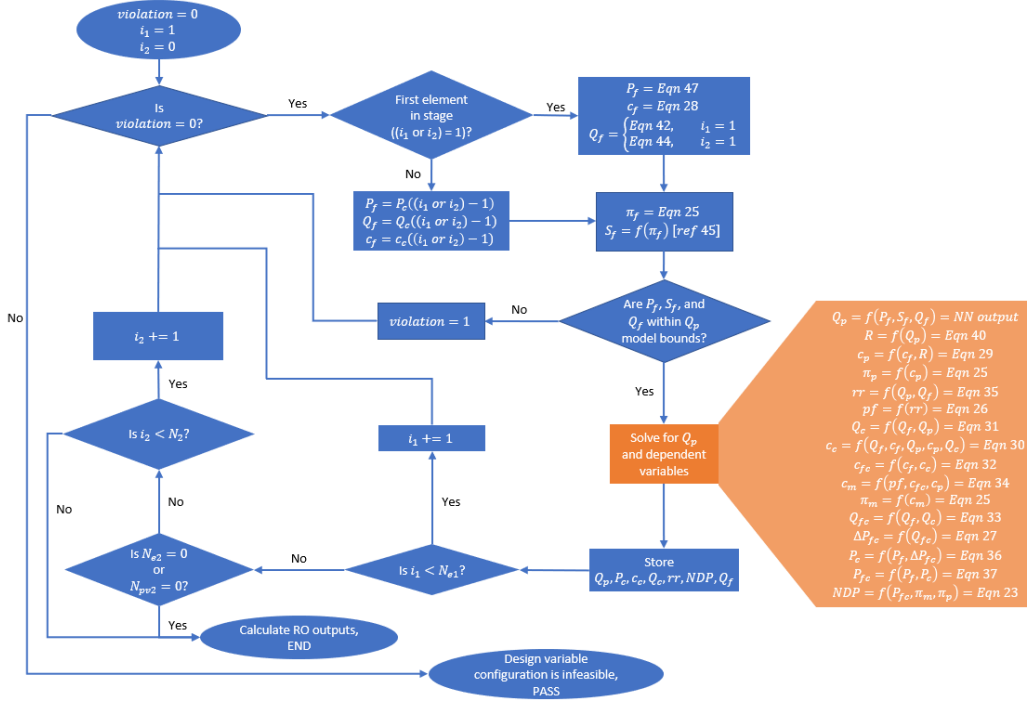


Figure 10: RO algorithm flowchart.

i_1 and i_2 are counters representing the element in stages 1 and 2, respectively, that the RO algorithm is finding the RO-related variables for. For the first element in the first stage, these are initialized to 1 and 0, and a violation variable is initialized to 0. If either i_1 or i_2 is equal to 1, then P_f , c_f , and Q_f are solved for using Equations 47, 28, and either 42 or 44 (depending on if $i_1 = 1$ or $i_2 = 1$). Otherwise, P_f , Q_f , and c_f equal the concentrate values from the upstream element. From there, Equation 25 is used to find π_f , which is then used to find S_f by using the methodology in [46]. If the solved P_f , S_f , or Q_f are not within the bounds of the model, then the violation variable is set to 1, which signifies an infeasible design variable configuration, and the RO algorithm is terminated. Otherwise, Q_p its dependent variables are solved for using the equations listed in Figure 10, and stored for later calculations. If there is another element in the first pressure vessel, then $i_1 = 1$, and the algorithm continues in a similar manner as above. If not, and there is another RO stage, an element counter i_2 is set to 1 and the algorithm continues. Otherwise, the final RO system outputs are calculated, and the

overall optimization continues.

3.6. Incorporation into IPHROS Model

As noted in Figure 2, the eight design variables in the IRPHO system model are E_r , γ , γ_{RO} , h_L , N_{e1} , N_{e2} , N_{pv1} , and N_{pv2} . With $\dot{V}_{w,RO}$ being a fraction, γ_{RO} , of \dot{V}_{wp} , $\dot{V}_{w,RO}$ can then be expressed in terms of design variables by use of Equation 2,

$$\dot{V}_{w,RO} = \frac{3.6 \times 10^6 \times \dot{E}_r \gamma \gamma_{RO} \eta_{hp}}{\rho_{sw} g h_L}. \quad (41)$$

Since there is a maximum limit to the feed flowrate each pressure vessel making up the RO system can allow, dividing $\dot{V}_{w,RO}$ by N_{pv1} allows for the feed flowrate into each pressure vessel in the first RO stage to be expressed in terms of overarching design variables,

$$Q_f = \frac{3.6 \times 10^6 \times \dot{E}_r \gamma \gamma_{RO} \eta_{hp}}{24 \times \rho_{sw} g h_L N_{pv1}}, \quad (42)$$

where the factor of 24 converts m^3/day to m^3/hr . When going between stages, it is necessary to redefine what the feed flowrate going into the second stage is, \dot{V}_{int} , which is a function of the number of pressure vessels in the first and second stages N_{pv1} and N_{pv2} ,

$$\dot{V}_{int} = 24 \times Q_{c, \text{last element}, 1} \times N_{pv1}, \quad (43)$$

where the 1 in $Q_{c, \text{last element}, 1}$ indicates stage 1. Rewriting Equation 43 to refer to the feed flowrate going into the first element of a single pressure vessel in the second stage results in Equation 44,

$$Q_{int} = \frac{\dot{V}_{int}}{24 \times N_{pv2}}. \quad (44)$$

The total volume per day of fresh water generated from the RO system can be calculated by the summation of the permeate flow rates from each individual element in a stage multiplied by the number of pressure vessels in that stage. For two stages,

$$\dot{V}_{fw,RO} = 24 \times \left(N_{pv1} \sum_{i=1}^{N_{e1}} Q_{p,i} + N_{pv2} \sum_{i=1}^{N_{e2}} Q_{p,i} \right). \quad (45)$$

The total volume per day of brine generated from the RO system can be calculated by the product of the brine leaving the last element of the second stage (or the first stage if there is only one stage), and the number of pressure vessels in the second (or first) stage,

$$\dot{V}_{o,RO} = 24 \times N_{pvX} \times Q_{c, \text{last element}, X}, \quad (46)$$

where the X subscript denotes the first or second stage. The interplay between the variables expressed by Equations 41-46 and the two stages is shown below in Figure 11.

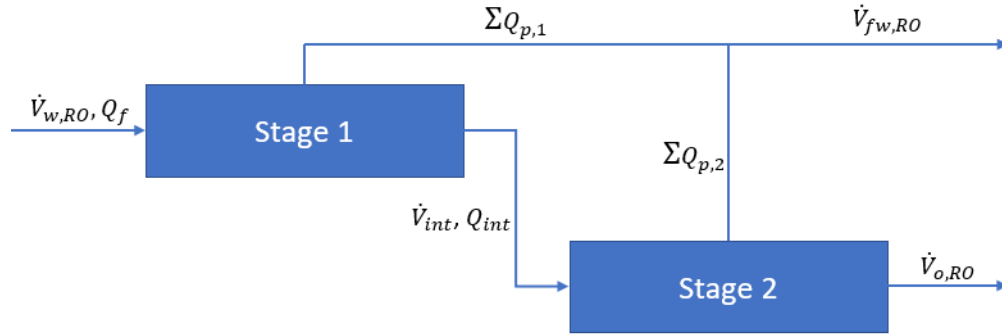


Figure 11: Illustration of the flows into and out of the two stages present in the RO model. Each stage contains a collection of RO elements and pressure vessels.

Additionally, the feed pressure for the first element of a pressure vessel in the first stage is directly related to the height of the upper reservoir by the conversion equation

$$P_f = 1000gh_L \frac{\rho_{sw}}{\rho_{fw}}. \quad (47)$$

Equation 47 assumes that there are no frictional losses to $\dot{V}_{w,RO}$ prior to the flow entering the RO system. This stems from an assumption that the pipe

this flow moves through is not long enough to significantly reduce P_f . In future work, however, more accurate modeling of this and the other flows through IPHROS will play an important role when considering site-specific case studies.

Furthermore, the RO system recovery ratio, the third objective for this IPHROS, can be calculated from the individual recoveries for each element by

$$\eta_{RO} = 1 - \prod_{i=1}^{(N_{e1}+N_{e2})} (1 - rr_i). \quad (48)$$

The fraction of pressure leaving the RO system can be calculated from the pressures computed in the RO model,

$$\eta_{RO,io} = 1 - \frac{P_{f, \text{first element}} - P_{c, \text{last element}}}{P_{f, \text{first element}}}. \quad (49)$$

Lastly, correlation equations, detailed in [46], allow for the salinity to be calculated from the osmotic pressure, and the density to then be calculated from the salinity, at elevated pressures. These correlations are used to determine S_f , $S_{o,RO}$, and $\rho_{o,RO}$.

4. IPHROS Optimization Methodology

4.1. Objectives and constraints

Looking at the model holistically, there are two pairs of overlapping competing objectives. The first set is \dot{E}_{er} and $\dot{V}_{fw,RO}$, which are both desired to be maximized, where increasing one will decrease the other. The second (less obvious) set is $\dot{V}_{fw,RO}$ and η_{RO} , where $\dot{V}_{fw,RO}$ is desired to be maximized, but may not actually occur at maximal η_{RO} . These objectives are dictated by the values of the design variables \dot{E}_r , γ , γ_{RO} , h_L , N_{e1} , N_{e2} , N_{pv1} , and N_{pv2} .

The primary constraint in this model is on the salinity of the final discharge into the ocean, S_{ht} . Different regions of the world have different requirements on the salinity of brine discharged into the ocean to mitigate adverse

environmental effects. For instance, California sets an upper limit to discharge salinity at 40 grams of salt per kilogram of seawater [50]. California’s discharge salinity limit is the constraint adopted in this paper.

Additional constraints arise from the RO model. As noted in Section 3.3, there are design limits to rr , Q_f (and Q_c), and Q_p for the SeamaxxTM-440 element, but also more generally for any RO element. Also, the net driving pressure across an RO membrane, NDP , must be greater than zero for reverse osmosis to occur. As the optimization algorithm runs, any time that the value of one of these RO constraints is violated, the responsible design variable configuration is instantly deemed infeasible by extreme values of the RO outputs being returned, which in turn makes the related objective and constraint values not even remotely optimal.

4.2. Formal problem statement

Expressing this MDO problem in standard form results in the following problem statement:

$$\min \quad J(x, p) \quad (50)$$

$$s.t. \quad S_{ht} - 40 \leq 0 \quad (51)$$

$$-rr \leq 0 \quad (52)$$

$$rr - 0.13 \leq 0 \quad (53)$$

$$3.41 - Q_f \leq 0 \quad (54)$$

$$Q_f - 15.5 \leq 0 \quad (55)$$

$$3.41 - Q_c \leq 0 \quad (56)$$

$$Q_c - 15.5 \leq 0 \quad (57)$$

$$-Q_p \leq 0 \quad (58)$$

$$Q_p - 1.32 \leq 0 \quad (59)$$

$$\epsilon - NDP \leq 0 \quad (60)$$

$$2 - (N_{pv2}^{N_{e2}} + N_{e2}^{N_{pv2}}) \leq 0 \quad (61)$$

$$N_{pv2} - N_{pv1} \leq 0 \quad (62)$$

$$\begin{bmatrix} 1 \\ 0.01 \\ 0.01 \\ 240 \\ 1 \\ 0 \\ 1 \\ 0 \end{bmatrix} \leq x = \begin{bmatrix} \dot{E}_r \\ \gamma \\ \gamma_{RO} \\ h_L \\ N_{e1} \\ N_{e2} \\ N_{pv1} \\ N_{pv2} \end{bmatrix} \leq \begin{bmatrix} 100 \times 10^6 \\ 0.99 \\ 0.99 \\ 821 \\ 8 \\ 8 \\ 1.86 \times 10^6 \\ 1.86 \times 10^6 \end{bmatrix}, \quad (63)$$

where

$$J = \begin{bmatrix} -\dot{E}_{er} \\ -\dot{V}_{fw,RO} \\ \eta_{RO} \end{bmatrix} \quad (64)$$

and

$$\epsilon = 1 \times 10^{-6}. \quad (65)$$

with \dot{E}_{er} , $\dot{V}_{fw,RO}$, and η_{RO} determined by Equations 12, 45, and 48 respectively.

The first constraint, Equation 51, corresponds to the discharge salinity limit mentioned in Section 4.1. Equation 60 enforces a positive net driving pressure for any feasible IPHROS design configuration and operation, where ϵ is a slack variable that, in this case, allows for the conversion of a “less than” constraint to a “less than or equal to” constraint. Equation 61 is included as a way of enforcing the fact that if there is not a second RO stage, or if there are no RO elements in the second RO stage, then the other should hold true as well. The constraint in Equation 62 both follows intuition and, following Equations 43 and 44, prevents the feed flowrate into the first element in a pressure vessel in the second stage from being reduced relative to the concentrate flowrate leaving the last element in a pressure vessel in the first stage, which would be counter-productive in maximizing the permeate flowrate from an RO element. The remaining constraints, Equations 52-59, are the RO element constraints included in Table 3.

Making the objectives negative in Equation 64 effectively maximizes these three variables in the optimization algorithm. The upper bound for \dot{E}_r in Equation 63 is a hypothetical value inclusive of the renewable energy requirement for the IPHROS presented in [21] to provide 2 kW/person and 500 L/day/person to one million people living a typical Californian lifestyle [21]. h_L being in the range of 240 meters and 821 meters correlates to a feed pressure range of approximately 2.413×10^6 Pa (350 psi) and 8.274×10^6 Pa (1200 psi) based on Equation 47. While the RO model was developed with a lower pressure threshold of 2.068×10^6 Pa (300 psi) in mind, having a slightly higher lower bound for the feed pressure (h_L) into the RO system allows for the expected pressure drops according to Equation 27 to occur while minimizing concerns of an RO element’s feed pressure dropping below 300 psi. The upper bounds of N_{e1} and N_{e2} are the maximum number of RO elements that can be placed in a pressure vessel, according to [40]. The lower bounds for N_{e1} and N_{pv1} are 1 because there will always be at least one RO stage (with at least one RO element in it) for generating fresh water, while the lower bounds for N_{e2} and N_{pv2} are 0 because it is possible that there is not a second RO stage. The upper bounds for N_{pv1} and N_{pv2} is such that, using Equation 42, having a maximal \dot{E}_r , γ , and γ_{RO} , and minimal h_L and Q_f , the resulting N_{pv1} is within the design variable range.

5. Results

5.1. Optimization

Optimization was completed heuristically using a multiobjective genetic algorithm, specifically a variant of NSGA-II utilized by the gamultiobj MATLAB function [51]. Convergence was reached in 102 generations to within a function tolerance of 0.0001. The population size was 200 with a Pareto fraction of 0.35, meaning that only 70 members of the population lie on the Pareto front. The constraint tolerance was 0.001, the crossover fraction was 0.8, and the elite count was 10.

Figure 12a shows the resulting Pareto front from the above optimization, with its projections included in Figure 13. To better illustrate the non-linearity of the Pareto front, an interpolated surface plot is also included in Figure 12b. Also marked on these figures are the Pareto points that coincide with the maximal \dot{E}_{er} , maximal $\dot{V}_{fw,RO}$, maximal η_{RO} , and the “best” Pareto point relative to these objectives.

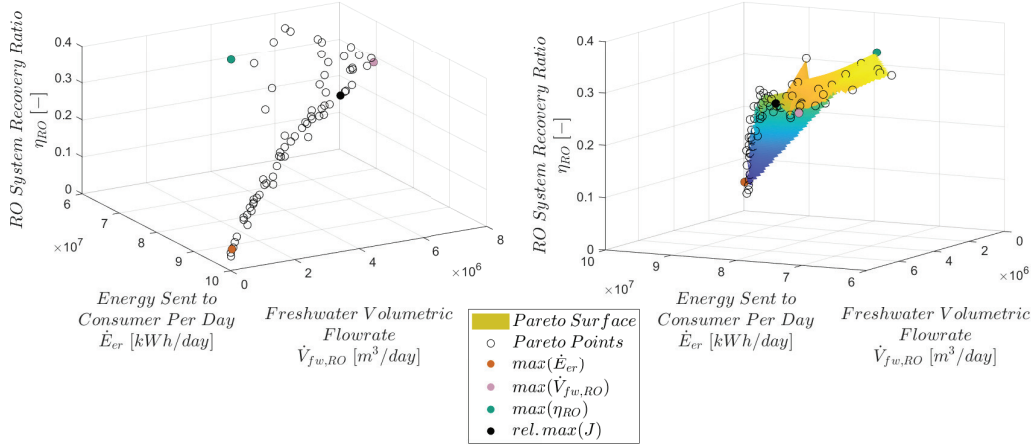


Figure 12: Pareto front generated by a genetic algorithm used to maximize $\dot{V}_{fw,RO}$, \dot{E}_{er} , and η_{RO} . To better indicate the curvature present in the Pareto front in (a), a surface plot of the Pareto front is also included in (b). Additionally, the Pareto points that lie at maximal \dot{E}_{er} , maximal $\dot{V}_{fw,RO}$, maximal η_{RO} , and the relatively “best” Pareto point are also shown.

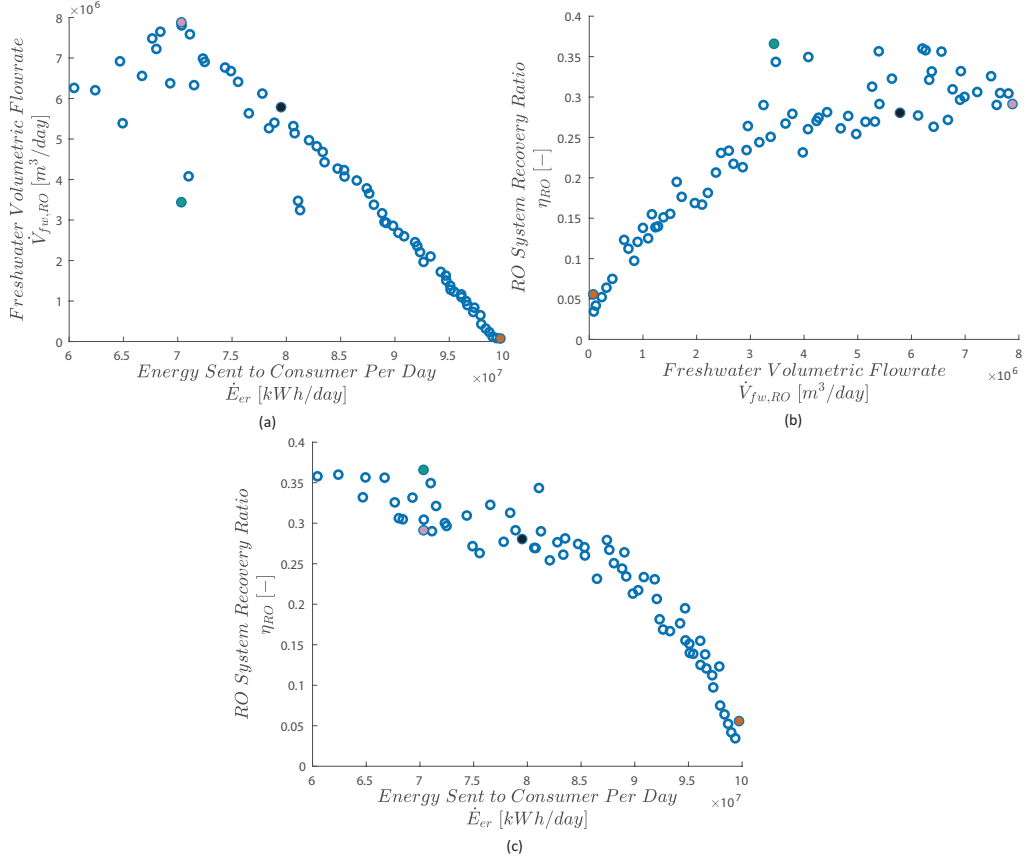


Figure 13: Projections of the Pareto front shown in Figure 12. (a) is a projection onto the \dot{E}_{er} and $\dot{V}_{fw,RO}$ axes, (b) is a projection onto the $\dot{V}_{fw,RO}$ and η_{RO} axes, and (c) is a projection onto the \dot{E}_{er} and η_{RO} axes.

The Pareto solutions are also shown as a parallel coordinates plot in Figure 14, where the axes include both the eight design variables from Equation 63 and the three objective variables from Equation 64. Also indicated on this figure are the Pareto solutions that perform in the top 10% of each of the three objectives, and the “best” Pareto solution relative to these objectives.

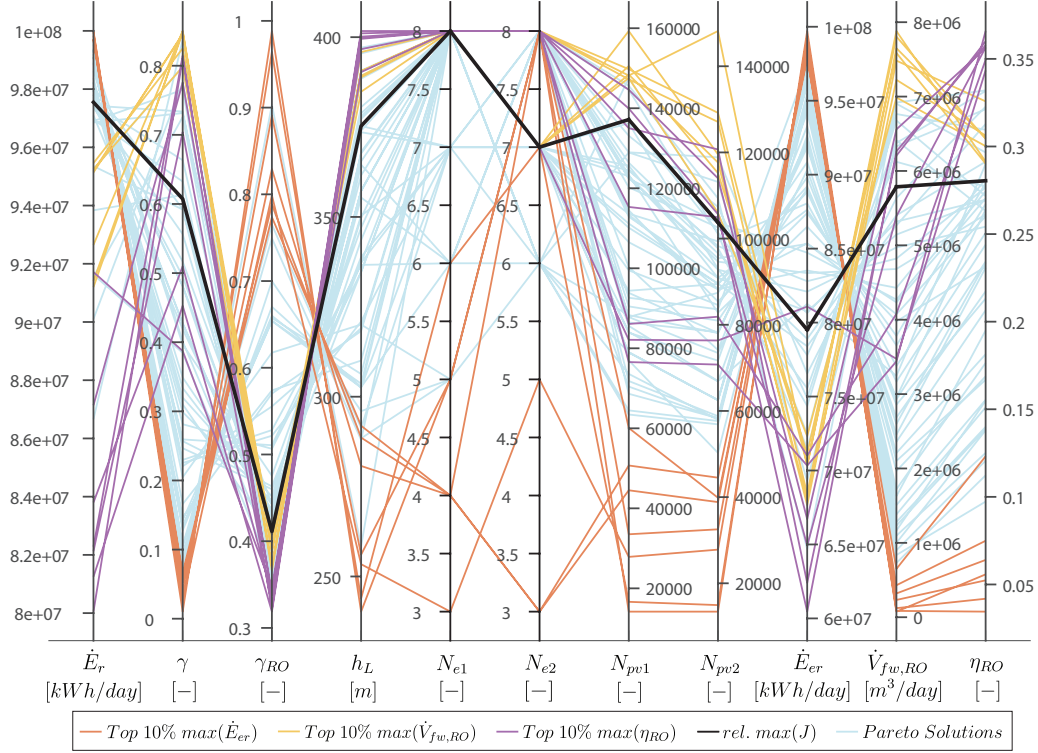


Figure 14: Parallel coordinates plot obtained from the optimization problem posed in Section 4.2. The red lines represent Pareto solutions that are in the top 10% of maximal daily energy sent to the consumer, the yellow lines represent Pareto solutions that are in the top 10% of maximal freshwater volumetric flowrate to the consumer, and the purple lines represent Pareto solutions that are in the top 10% of RO system recovery ratio. The “best” Pareto solution relative to these three objectives is represented by the black line. All other Pareto solutions are indicated by light blue lines.

The three Pareto points that result in the most maximal \dot{E}_{er} , $\dot{V}_{fw,RO}$, and η_{RO} , as well as the corresponding values of the other objectives, are shown below in Table 5. Using these three values of \dot{E}_{er} , $\dot{V}_{fw,RO}$, and η_{RO} to define a Utopian point, or the set of optimal values for the cases that each objective were to be optimized independently, the “best” Pareto point can be determined by finding the point with the smallest Euclidean distance from the Utopia point. This point, the relative maximum of J (from Equation 64), is also included in Table 5. Additionally, the values of the design variables that

correspond to these four points are also indicated in the table.

Table 5: Comparison of select Pareto points shown in Figure 12.

Pareto Point:	At $\max(\dot{E}_{er})$	At $\max(\dot{V}_{fw,RO})$	At $\max(\eta_{RO})$	At rel. $\max(J)$
\dot{E}_{er} Value	9.9710×10^7	7.0331×10^7	7.0331×10^7	7.9513×10^7
$\dot{V}_{fw,RO}$ Value	7.5199×10^4	7.8811×10^6	3.4401×10^6	5.7866×10^6
η_{RO} Value	0.0558	0.2913	0.3659	0.2804
$x^* = \begin{bmatrix} \dot{E}_r \\ \gamma \\ \gamma_{RO} \\ h_L \\ N_{e1} \\ N_{e2} \\ N_{pv1} \\ N_{pv2} \end{bmatrix}^*$	$\begin{bmatrix} 9.9995 \times 10^7 \\ 0.0102 \\ 0.9883 \\ 240.2194 \\ 5 \\ 8 \\ 14182 \\ 13422 \end{bmatrix}$	$\begin{bmatrix} 9.5181 \times 10^7 \\ 0.8453 \\ 0.4037 \\ 384.8242 \\ 8 \\ 8 \\ 148171 \\ 148171 \end{bmatrix}$	$\begin{bmatrix} 8.1250 \times 10^7 \\ 0.4528 \\ 0.3188 \\ 399.8290 \\ 8 \\ 8 \\ 76557 \\ 70760 \end{bmatrix}$	$\begin{bmatrix} 9.7561 \times 10^7 \\ 0.6074 \\ 0.4077 \\ 375.2195 \\ 8 \\ 7 \\ 137130 \\ 103563 \end{bmatrix}$

Another key result that can be ascertained from the above optimization are plots of the values of the constraint variables for the Pareto solutions, especially those that are influential in dictating the flows through an RO system. These variables, from Equations 51-60, are S_{ht} , rr , Q_p , NDP , Q_f , and Q_c , and are plotted below in Figure 15. The constraint plots in this figure shed light onto the activity of the constraints in determining the non-dominated solutions obtained by the genetic algorithm. Each x coordinate corresponds to one Pareto solution. The largest constraint variable values for a Pareto solution are indicated with circular markers, and the smallest constraint values are indicated with triangular markers. The constraint limits, correspondingly labeled “max” or “min”, are included as well. Similarly to Figures 12-14, the 10% best Pareto solutions with regards to the three objectives are also indicated, as well as the “best” Pareto solution.

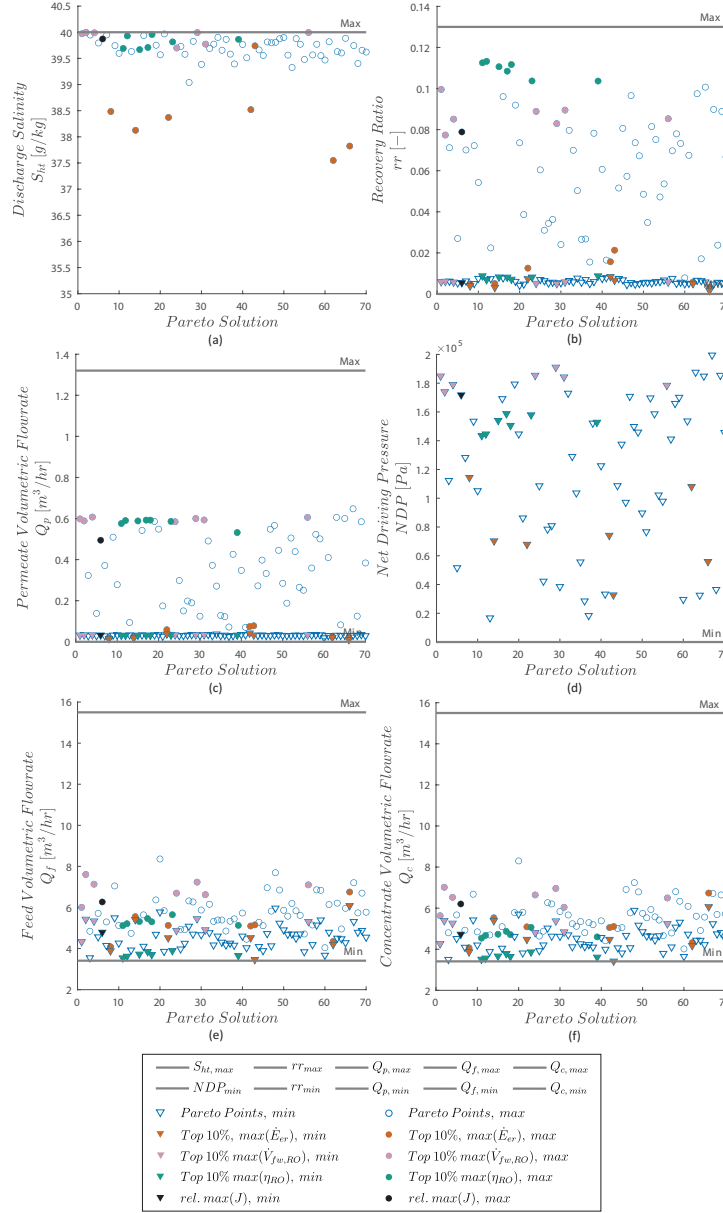


Figure 15: Plots of the constraint variable values that occur in the RO system model for each Pareto solution. (a) shows the values of S_{ht} , (b) shows the maximum and minimum values for rr , (c) shows the maximum and minimum values for Q_p , (d) shows the minimum values for NDP , (e) shows the minimum and maximum values for Q_f , and (f) shows the minimum and maximum values for Q_c .

5.2. Economic Assessment

Future work will include a full economic analysis. However, some economic insight is still beneficial for contextualizing the four Pareto solutions identified in Table 5 with each other, as well as the Pareto solutions with established PSH and RO systems. This will first be done by determining what the CAPEX and OPEX would be if the energy and freshwater production from each IPHROS configuration in Table 5 were being generated by independent PSH and RO systems. Beginning with the CAPEX calculations for RO, the capital cost and capacity for a plant (of any type) can be related to each other using the normal capacity/cost correlation rule from [52], shown below in Equation 66,

$$\frac{Capital\ Cost_{plant1}}{Capital\ Cost_{plant2}} = \left(\frac{Plant\ Capacity_{plant1}}{Plant\ Capacity_{plant2}} \right)^m, \quad (66)$$

where m is the scale factor, which for RO can be assumed to be approximately 0.8. Using the Ashdod desalination plant in Israel as an example, with a CAPEX of \$400 million (in 2015 USD, which is equal to \$504 million in 2022 USD) and a capacity of 100,000,000 m³/year [21], the CAPEX for the individual RO component of the Pareto solutions in Table 5 (in 2022 USD) can be determined using Equation 67,

$$CAPEX_{RO} = (504 \times 10^6) \left(\frac{\dot{V}_{fw,RO} \times 365}{100 \times 10^6} \right)^{0.8}. \quad (67)$$

The OPEX for RO, $OPEX_{RO}$, will simply be assumed to have a coefficient of \$0.5734/m³/day (in 2022 USD from 2017 USD) [53], resulting in Equation 68,

$$OPEX_{RO} = 0.5734 \times \dot{V}_{fw,RO} \times 365. \quad (68)$$

Transitioning to the CAPEX calculations for PSH, Equation 66 can also be used, but the scale factor m must first be determined. To do so, the capital costs and plant capacities for two plants from [54], Swan Lake North and JD Pool, can be used, resulting in a scale factor of approximately 1.1. Generating a CAPEX equation in a manner similar to Equation 67, with the Swan Lake North PSH plant used as the known example, results in

Equation 69. However, in order to isolate the PSH component from the rest of IPHROS, the component of renewable energy sent directly to the consumer, $\dot{E}_d = (1 - \gamma)\dot{E}_r$, must be subtracted to get the plant capacity.

$$CAPEX_{PSH} = (1.26 \times (2250 \times 600 \times 10^3)) \left(\frac{\dot{E}_{er} - (1 - \gamma)\dot{E}_r}{600 \times 10^3 \times 24} \right)^{1.1}, \quad (69)$$

where the factor of 1.26 is a conversion from 2014 USD to 2022 USD and the factor of 24 accounts for the fact that since there is a steady state assumption for the upper reservoir of this IPHROS model, the energy duration for the Swan Lake North PSH plant needs to be 24 hours to match. Lastly, from [54], the OPEX for a PSH plant can be calculated using the equation

$$OPEX_{PSH} = 34730 \times P^{0.32} \times AE^{0.33}, \quad (70)$$

where P is the plant capacity in MW and AE is the annual energy throughput in MWh. Modifying Equation 70 to be in terms of IPHROS variables results in Equation 71,

$$OPEX_{PSH} = 34730 \times \left(\frac{(\dot{E}_{er} - (1 - \gamma)\dot{E}_r)}{24 \times 10^3} \right)^{0.32} \times \left(\frac{(\dot{E}_{er} - (1 - \gamma)\dot{E}_r) \times 365}{10^3} \right)^{0.33}. \quad (71)$$

In addition to the CAPEX and OPEX, the annual values of electricity (AVE) and fresh water (AVW) can be calculated for the Pareto solutions in Table 5, assuming an average sale price of electricity (in California) of 0.1965 \$/kWh [55] and an average sale price of fresh water (in Los Angeles County, California) of 0.975 \$/m³ [56]. Using these unit prices of electricity and water, as well as Equations 67, 68, 69, and 71, $CAPEX_{RO}$, $OPEX_{RO}$, $CAPEX_{PSH}$, $OPEX_{PSH}$, AVE , and AVW can be calculated for, and are displayed below in Figure 6.

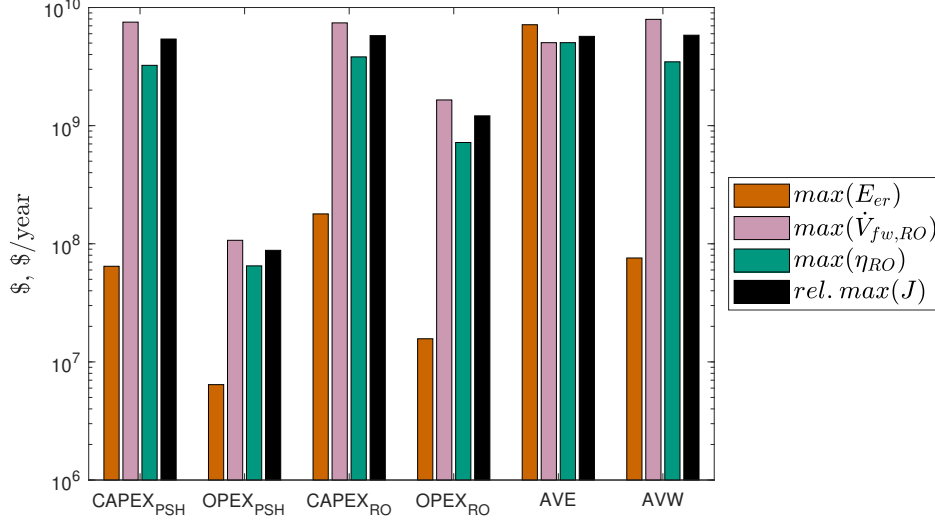


Figure 16: CAPEX and OPEX values for the four Pareto solutions from Table 5, where the values are for if the PSH and RO subsystems of that IPHROS configuration were implemented individually. The annual values of electricity and water (AVE and AVW respectively) for these points are included as well. CAPEX values are in units of \$, and OPEX, AVE , and AVW values are in units of \$/year.

The CAPEX, OPEX, AVE , and AVW values in Figure 16, to reiterate, are for the case that the PSH and RO subsystems of the four IPHROS configurations from Table 5 are implemented independently. However, with the benefits that IPHROS presents in the sharing of infrastructures, these CAPEX and OPEX values can be reduced in the form of savings coefficients $\eta_{CAPEX,PSH}$, $\eta_{OPEX,PSH}$, $\eta_{CAPEX,RO}$, and $\eta_{OPEX,RO}$. Their effects can be reflected in various economic metrics, such as the time, t , it takes for the net present value, NPV , of a project to reach \$0 for an assumed interest rate i of 5%. Assuming that the return on investment is negligible, and using the variables in Figure 16,

$$NPV = (AVE + AVW - (OPEX_{PSH} + OPEX_{RO})) \times (P/A, i, t) - (CAPEX_{PSH} + CAPEX_{RO}), \quad (72)$$

where $(P/A, i, t)$ is the uniform series present-worth factor. Incorporating the savings coefficients into Equation 72 results in Equation 73,

$$\begin{aligned}
NPV = & (AVE + AVW - (\eta_{OPEX,PSH} OPEX_{PSH} + \eta_{OPEX,RO} OPEX_{RO})) \times \\
& (P/A, i, t) - (\eta_{CAPEX,PSH} CAPEX_{PSH} + \\
& \eta_{CAPEX,RO} CAPEX_{RO}).
\end{aligned}
\tag{73}$$

With shared intake equipment and no power requirement for feed pressurization, $\eta_{CAPEX,RO}$ and $\eta_{OPEX,RO}$ are assumed to be 0.79 and 0.56, respectively [53]. While there are possible savings on the PSH side of IPHROS due to less curtailed power and the slight overlap of some operational expenditures, $\eta_{CAPEX,PSH}$ and $\eta_{OPEX,PSH}$ are assumed to both equal 1 [57]. This signifies that most of the savings from IPHROS come from the removal of components from the RO side of the system. Using these values, the times (in years) that result from setting Equations 72 and 73 equal to zero and using values from Figure 16 are shown in Table 6, as well as the percent decrease in time going from the “independent” PSH and RO configuration to an IPHROS configuration.

Table 6: Times for the NPV of individually implemented PSH and RO facilities and IPHROS facilities to reach \$0 for the IPHROS configurations in Table 5, with the percent decrease resulting from the IPHROS configuration included as well. The assumed interest rate is 5%.

Configuration	Pareto Point At $\max(\dot{E}_{er})$	Pareto Point At $\max(\dot{V}_{fw,RO})$	Pareto Point At $\max(\eta_{RO})$	Pareto Point At rel. $\max(J)$	Units
Individual	0.0347	1.41	0.958	1.15	years
IPHROS	0.0293	1.18	0.813	0.972	years
Percent Decrease	15.5	16.3	15.1	15.6	%

6. Discussion

Prior to the development of the RO model presented in this paper, strictly using equations from [21], the resulting Pareto front from optimization was linear due to a linear dependence between \dot{E}_{er} and $\dot{V}_{fw,RO}$. The fact that the Pareto front in Figure 12 is no longer linear indicates that the IPHROS model is beginning to capture some of the nonlinear complexities that exist within and between its subcomponents. This should continue to be the case as the model is built upon, with future considerations such as (renewable) power grid dynamics, consumer demands for energy and fresh water, and improved modeling of seawater flows throughout the system.

However, even with this new model, an underlying linear trend is still somewhat evident in Figure 13a, which indicates that the largest tradeoff that exists in this IPHROS model, in terms of objectives, is the amounts of energy and fresh water being sent to the consumer. This matches the intuition confirmed by Figure 14, which highlights the relatively binary choices in the design variables of IPHROS that leads to the best values of \dot{E}_{er} and $\dot{V}_{fw,RO}$. For some amount of renewable energy production, maximal energy to the consumer requires small values for γ and h_L , and large values of γ_{RO} . This follows from the idea that lower values for γ and h_L minimize any energy losses arising in IPHROS due to pump-side and turbine-side efficiencies. Higher values for γ_{RO} are paired with lower values of γ to get “relatively” maximal freshwater flowrates for the little water (energy) being sent to IPHROS. Conversely, maximal fresh water to the consumer requires large values for γ and h_L , and small values for γ_{RO} . While at first glance one would expect larger values of γ_{RO} for larger freshwater output from the RO system, what really dictates the flowrate into the RO system, from Equation 41, is the product of γ and γ_{RO} . This product being maximal allows for the coupled tuning of γ and γ_{RO} to get maximal $\dot{V}_{fw,RO}$, while still getting comparatively larger values for \dot{E}_{er} . Relatively higher upper reservoir heights of approximately 400 m, corresponding to feed pressures of 4.023×10^6 Pa (583 psi), are required for obtaining larger permeate flowrates. The effect that balancing these three design variables (γ , γ_{RO} , and h_L) has on the objectives \dot{E}_{er} and $\dot{V}_{fw,RO}$ is evident from the “best” Pareto solution in Figure 14, where the relatively best objective values is obtained from, in part, middling values for γ , γ_{RO} , and h_L .

Intuitively, trends for $\dot{V}_{fw,RO}$ and η_{RO} should follow each other fairly closely

because, ideally, larger freshwater flowrates occur at higher RO system recovery ratios. This is generally true, and is reflected in Figure 13b by the positive correlation between $\dot{V}_{fw,RO}$ and η_{RO} . However, this is not a perfectly linear trend, and comes from the governing equations for these two objectives (Equations 45 and 48, respectively). While both are a function of the element permeate flowrate $Q_{p,k}$, the element recovery ratio rr_k is also a function of the element feed flowrate $Q_{f,k}$. Made evident by the WAVE simulations from Section 3.4 is that a diverse combination of feed flow conditions into an RO element can lead to large element recovery rates. Therefore, maximal values for rr_k , and therefore η_{RO} , can be attained by controlling the inlet conditions as much as by striving for large permeate flowrates. This explains why, for the maximal $\dot{V}_{fw,RO}$ and η_{RO} in Figure 14, there are slightly different values associated with γ , γ_{RO} , h_L , N_{pv1} , and N_{pv2} , because these are the design variables that control the flow conditions going into an RO element. This also explains the sum of N_{e1} and N_{e2} for these Pareto solutions being 15 or 16, because the more elements there are, the more opportunity there is to generate permeate. As a side note, utilizing a combination of different RO elements with different properties in each stage could lead to larger freshwater volumetric flowrates and/or recovery ratios. This will be explored in future work.

Something interesting to note from Figure 14 is the range in values for \dot{E}_r that lead to Pareto optimal solutions. Keeping this variable a design variable rather than setting it as a parameter allowed for the effects of this variable to be explored from the optimization. While more renewable energy production is intuitively favorable for all aspects of IPHROS (especially for \dot{E}_{er}), non-maximal amounts of \dot{E}_r lead to maximal $\dot{V}_{fw,RO}$ and η_{RO} objectives. This signifies a lower renewable energy requirement for an entity implementing IPHROS with the goal of, for instance, maximal freshwater production, and makes that version of IPHROS that much more attainable.

Many of the trends depicted in Figure 15 are to be expected. Starting with S_{ht} in Figure 15a, higher permeate flowrates naturally result in more highly concentrated brine, which would increase the salinity of the output from the Mixing module, so maximal objective values for $\dot{V}_{fw,RO}$ and η_{RO} naturally occur at the maximal limit for S_{ht} . Intuitively, for NDP in Figure 15d, while the lower limit of ϵ only ensure that RO can physically occur and there is no upper limit, it follows from RO theory that a higher net driving pressure results in more mass being forced across the membrane, which is more im-

portant for maximum permeate generation than it is more maximum element recovery ratios. For rr in Figure 15b, maximum individual element recovery ratios for Pareto solutions with maximal η_{RO} values not being right along the maximal limit of 0.13 is likely do to the use of a heuristic algorithm, which only approximates the Pareto front, versus a gradient-based method. Interestingly, however, maximal η_{RO} Pareto solutions have a maximum individual element recovery ratio distinctly higher than those for maximal $\dot{V}_{fw,RO}$ Pareto solutions. This adds to the somewhat surprising results from Figures 15c, 15e, and 15f, where the maximum limits for Q_p and Q_f (and Q_c) are not even close to being active constraints for these variables. This is likely because the constraints on the flow characteristics for each RO element must hold true for all the RO elements in all of the RO pressure vessels in both RO stages, so flow conditions that would lead to optimal flow conditions for an individual RO element may not be possible in the grand scheme of the entire RO system.

Comparing pumped storage hydropower to other types of energy storage technologies, two of the other main types are lithium-ion batteries and compressed air. The average levelized cost of storage (LCOS) for these technologies are \$0.13/kWh, \$0.11/kWh, and \$0.17/kWh respectively [58]. An actual comparison of the LCOS for IPHROS to these technologies cannot realistically take place until the steady state upper reservoir assumption is removed and the time-varying operating conditions for IPHROS are factored into the analysis (the subject of future work). However, with $\eta_{CAPEX,PSH}$ and $\eta_{OPEX,PSH}$ both equal to one for this study (indicating cost savings come entirely from the reduction of required components for RO), the cost savings that IPHROS presents would only really be realized in the LCOW for IPHROS. Therefore, the LCOS for IPHROS would not be lower than that of pumped storage hydropower. However, it certainly would not be higher, so the energy storage component of IPHROS would still remain competitive when compared to other storage technologies, both in terms of cost and the benefits of long-duration energy storage.

Looking at water production, the specific water production cost of reverse osmosis is \$0.857/m³ [59]. When compared to other desalination technologies, such as electrodialysis and membrane distillation, which have specific costs of \$1.532/m³ and \$6.50/m³ respectively [59]-[60], RO's cheaper cost and reliability have led to 50 % of global desalination being done with RO [61]. With $\eta_{CAPEX,RO}$ and $\eta_{OPEX,RO}$ both being less than one, this would make desali-

nation via an IPHROS implementation of RO even cheaper. Again, direct comparisons between the specific costs of water production from IPHROS to these stand-alone technologies will be the subject of future work.

While the exact time values in Table 6 should not be taken at face value, and would benefit from a site-specific derivation (to be the subject of future work), there is still value in comparing the obtained times to one another. In general, there is about a 16% decrease in the break even time for an IPHROS project compared to individually-implemented PSH and RO facilities, meaning that net project profits would be made that much more quickly. Additionally, based on the assumed selling prices of electricity and fresh water, the most profitable IPHROS configuration occurs at maximal energy being sent to the consumer, which makes sense because in this configuration, very little energy is being exposed to the various efficiencies in IPHROS, such as the pump-side and turbine-side efficiencies and the pressure drop in the RO system. The second most profitable IPHROS configuration is at maximal RO system recovery ratio, due to the optimal use of RO elements in the RO system in freshwater production.

Lastly, interpreting the values of \dot{E}_{er} and $\dot{V}_{fw,RO}$ at the relatively maximal J in Table 5 in terms of the individual consumer demands from Section 4.2, 1.66 million people could be serviced with electricity and 11.6 million people could be serviced with fresh water at this IPHROS configuration. These exact numbers would likely have more meaning if the analysis was tailored to a specific region, making use of assets such as data for renewable energy potentials, electricity and water grid data, and elevation data for siting IPHROS, which will be the subject of future work. However, just knowing that the population sizes being of the order of magnitudes indicated signifies that an optimized IPHROS is capable of operating at such a scale, which should aid in its adoption moving forward.

7. Conclusions

Developing a new RO model for the IPHROS model detailed in this paper represents an initial venture into increasing the accuracy of the model in representing real-life phenomena. For this new model, statistically sound surrogate models for the fractional salt rejection rate and permeate flowrate of a FilmTecTM SeamaxxTM-440 were developed. MDO of IPHROS using a multiobjective genetic algorithm revealed important trends in the design and

objective spaces that should be considered when IPHROS is implemented in actuality.

The most obvious tradeoff that exists for IPHROS is the one that exists between the amounts of energy and fresh water being sent to the consumer, which is fairly linear when considering the Pareto solutions. This signifies that the ultimate IPHROS configuration will be very dependent on the needs of the population being serviced, revealing that (more) site-specific case studies should be conducted. This is especially true with the future work for this IPHROS model requiring the removal of the steady state upper reservoir assumption when considering energy and freshwater markets and dynamic upper reservoir volumes. Additional future work will include accounting for the quality of the fresh water generated from the RO process, utilizing a variety of RO elements in the RO system, and further improved modeling overall, all of which will effect the determination of the optimal operating and design parameters of IPHROS.

While conceptually the optimization for maximal fresh water to the consumer and RO system recovery ratio are related, the best Pareto solutions for these objectives are distinct from one another. This means that producing maximal permeate flowrates versus producing permeate flowrates at maximal recovery ratios is a factor that future IPHROS designers should consider. For both of these objectives, which heavily pertain to the RO system, the most active constraints in the present framework are the salinity constraint on the mixed IPHROS discharge into the ocean, which places a cap on maximal RO production and recovery ratios, and the the minimum constraints for several RO-related constraints, which ensure the feasibility of RO for all RO elements in the RO system.

The optimization results show that at the optimal operating conditions of IPHROS with regards to all three objectives, over 77 million kWh of energy and 6 million m³ of freshwater could be generated on a daily basis. Considering an individual consumer uses about 2 kW and 500 liters of freshwater per day, this system could provide enough power for 1.66 million people and enough fresh water for 11.6 million people per day. Additionally, NPV analysis of the Pareto-optimal IPHROS configurations with regards to each objective and the relatively best IPHROS reveals that there is an approximately 16% decrease in the break even time for IPHROS compared to PSH and RO being implemented individually. Cumulatively, this work suggests that

IPHROS could be beneficial in helping communities transition to a renewable energy-based grid and alleviate stress on their freshwater supply.

8. Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

9. Acknowledgement

This work was supported in part by the McMullen Graduate Fellowship for first year doctoral students via Cornell University's Systems Engineering Department. Special thanks to Kapil Khanal for his help with RO data analysis, and to Owais Khalid and Wen-Jung (Joanna) Chien for their help with data acquisition from WAVE.

10. Data Availability

Supplemental material for this paper can be found at https://github.com/symbiotic-engineering/IPHROS_Applied_Energy_2022.

Appendix A. Main Table

Table A.7: Main table of all IPHROS variables, with each variable's name, units, variable type, and an associated value/range, if known. Variable bounds are discussed in Section 4.2.

Variable	Name	Units	Known Value/Range	Type
\dot{E}_r	Renewable Energy Generated Per Day	kWh/day	$1 \leq \dot{E}_r \leq 100 \times 10^6$	Design
γ	Fraction of \dot{E}_r Sent to IPHROS	-	$0.01 \leq \gamma \leq 0.99$	Design
γ_{RO}	Fraction of Water in Upper Reservoir Sent to RO System	-	$0.01 \leq \gamma_{RO} \leq 0.99$	Design
h_L	Upper Reservoir Height	m	$240 \leq h_L \leq 821$	Design
N_{e1}	Number of RO Elements in Series in Stage 1	-	$1 \leq N_{e1} \leq 8$	Design
N_{e2}	Number of RO Elements in Series in Stage 2	-	$0 \leq N_{e2} \leq 8$	Design
N_{pv1}	Number of RO Pressure Vessels in Parallel in Stage 1	-	$1 \leq N_{pv1} \leq 1.86 \times 10^6$	Design
N_{pv2}	Number of RO Pressure Vessels in Parallel in Stage 2	-	$0 \leq N_{pv2} \leq 1.86 \times 10^6$	Design
S_{ht}	Discharge Salinity	g/kg	$S_{ht} \leq 40$	Constraint

NDP_k	Net Driving Pressure for RO Element k	Pa	$0 \leq NDP_k$	Constraint
$Q_{c,k}$	Concentrate Volumetric Flowrate for RO Element k	m^3/hr	$3.41 \leq Q_{c,k} \leq 15.5$	Constraint
$Q_{f,k}$	Feed Volumetric Flowrate for RO Element k	m^3/hr	$3.41 \leq Q_{c,k} \leq 15.5$	Constraint
$Q_{p,k}$	Permeate Volumetric Flowrate for RO Element k	m^3/hr	$0 \leq Q_{p,k} \leq 1.32$	Constraint
rr_k	Recovery Ratio for RO Element k	-	$0 \leq rr_k \leq 1$	Constraint
\dot{E}_{er}	Energy Sent to Consumer Per Day	kWh/day	-	Objective
$\dot{V}_{fw,RO}$	Freshwater Volumetric Flowrate	m^3/day	-	Objective
η_{RO}	RO System Recovery Ratio	-	-	Objective
η_{hp}	Pump-Side Efficiency	-	0.894	Parameter
η_{ht}	Turbine-Side Efficiency	-	0.894	Parameter
ρ_{fw}	Freshwater Density	kg/m^3	996.9	Parameter
ρ_{sw}	Seawater Density	kg/m^3	1023.6	Parameter

g	Gravitational Acceleration	m/s^2	9.81	Parameter
M_{salt}	Molar Mass of Salt	g/mol	58.44	Parameter
P_p	Permeate Pressure	Pa	1.01325×10^5	Parameter
S_{sw}	Seawater Salinity	g/kg	35	Parameter
T	Temperature	$^{\circ}C$	25	Parameter
\dot{E}_{rp}	Energy Sent to Pump Per Day	kWh/day	-	Dependent
\dot{E}_{swht}	Energy from Turbine Per Day	kWh/day	-	Dependent
\dot{E}_{htRO}	Energy from Brine Energy Recovery Per Day	kWh/day	-	Dependent
\dot{E}_{rd}	Energy Sent Directly to Consumer Per Day	kWh/day	-	Dependent
S_f	Feed Salinity	g/kg	-	Dependent
$S_{o,RO}$	Brine Salinity	g/kg	-	Dependent
\dot{V}_{int}	Volumetric Flowrate into Second RO Stage	m^3/day	-	Dependent

\dot{V}_{wp}	Incoming Upper Reservoir Volumetric Flowrate	m^3/day	-	Dependent
$\dot{V}_{w,RO}$	Incoming RO Volumetric Flowrate	m^3/day	-	Dependent
\dot{V}_{swht}	Turbine Volumetric Flowrate	m^3/day	-	Dependent
$\dot{V}_{o,RO}$	Brine Volumetric Flowrate	m^3/day	-	Dependent
\dot{V}_{ht}	Discharge Volumetric Flowrate	m^3/day	-	Dependent
$\Delta P_{fc,k}$	Pressure Drop Along RO Element k	Pa	-	Dependent
$\eta_{RO,io}$	Fraction of Pressure Leaving RO System	-	$0 \leq \eta_{RO,io} \leq 1$	Dependent
$\pi_{f,k}$	Feed Osmotic Pressure for RO Element k	Pa	-	Dependent
$\pi_{m,k}$	Osmotic Pressure Along Membrane of RO Element k	Pa	-	Dependent
$\pi_{p,k}$	Permeate Osmotic Pressure for RO Element k	Pa	-	Dependent
$\rho_{o,RO}$	Brine Density	kg/m^3	-	Dependent
ρ_{ht}	Discharge Density	kg/m^3	-	Dependent

$c_{c,k}$	Concentrate Salt Concentration for RO Element k	mol/L	-	Dependent
$c_{f,k}$	Feed Salt Concentration for RO Element k	mol/L	-	Dependent
$c_{m,k}$	Salt Concentration Along Membrane in RO Element k	mol/L	-	Dependent
$c_{p,k}$	Permeate Salt Concentration for RO Element k	mol/L	-	Dependent
$c_{fc,k}$	Average Salt Concentration in RO Element k	mol/L	-	Dependent
$P_{c,k}$	Concentrate Pressure for RO Element k	Pa	-	Dependent
$P_{f,k}$	Feed Pressure for RO Element k	Pa	-	Dependent
$P_{fc,k}$	Average Pressure in RO Element k	Pa	-	Dependent
pf_k	Concentration Polarization Factor for RO Element k	-	$pf_k \geq 1$	Dependent
Q_{int}	Feed Volumetric Flowrate for first RO element in Second Stage of RO System	m^3/day	-	Dependent
$Q_{fc,k}$	Average Volumetric Flowrate in RO Element	m^3/hr	-	Dependent

R_k	Fractional Salt Rejection Rate for RO Element k	-	$0 \leq R \leq 1$	Dependent
-------	--	---	-------------------	-----------

Appendix B. Bivariate Scatterplots

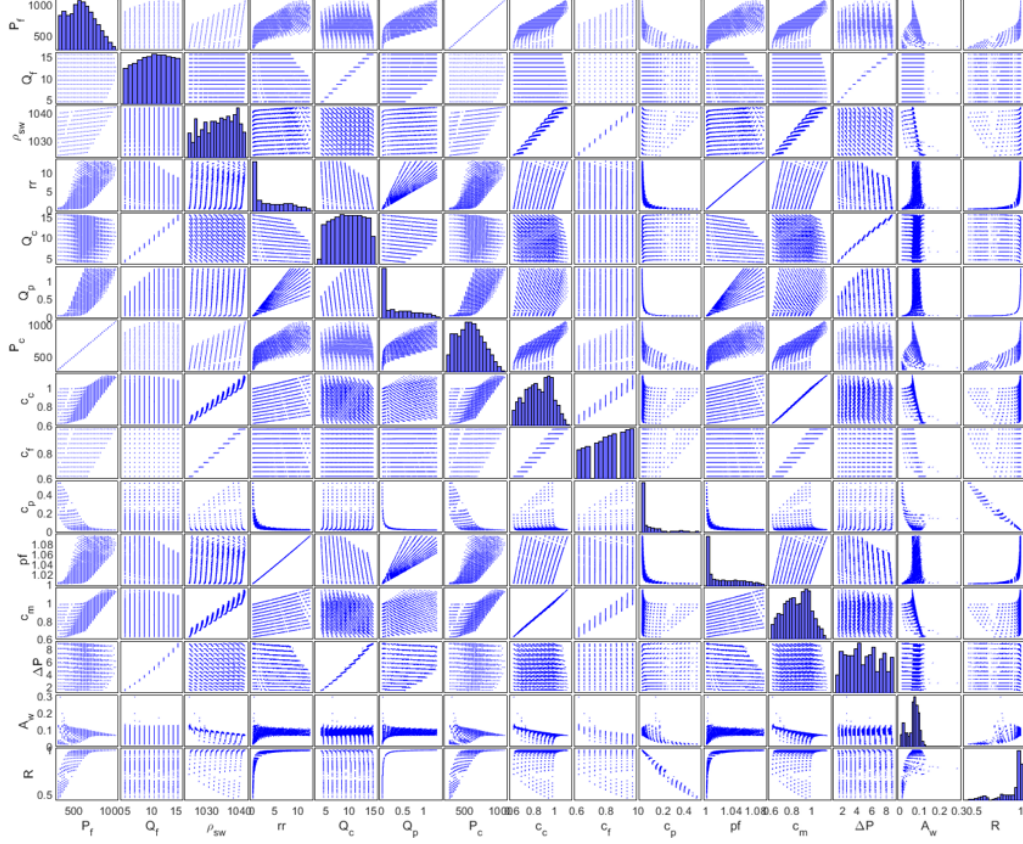


Figure B.17: Scatterplots from bivariate analysis in JMP. The variables P_f , Q_f , ρ_{sw} , rr , Q_c , Q_p , P_c , c_c , c_f , c_p , pf , c_m , ΔP , A_w , and R are plotted against each other, with histograms along the diagonal axis.

Appendix C. Nomenclature

Table C.8: Nomenclature used in this paper.

Variable	Name	Units
$\Delta P_{fc,k}$	Pressure drop along RO element k	Pa

\dot{E}_r	Renewable energy generated per day	kWh/day
\dot{E}_{er}	Energy sent to consumer per day	kWh/day
$\dot{E}_{ht,RO}$	Energy from brine energy recovery per day	kWh/day
\dot{E}_{rd}	Energy sent directly to consumer per day	kWh/day
\dot{E}_{rp}	Energy sent to pump per day	kWh/day
\dot{E}_{swht}	Energy from turbine per day	kWh/day
\dot{m}_s	Mass flow rate of salt	g/day
\dot{m}_{fw}	Mass flow rate of fresh water	g/day
$\dot{V}_{fw,RO}$	Freshwater volumetric flowrate	m ³ /day
\dot{V}_{ht}	Discharge volumetric flowrate	m ³ /day
\dot{V}_{int}	Volumetric flowrate into second RO stage	m ³ /day
$\dot{V}_{o,RO}$	Brine volumetric flowrate	m ³ /day
\dot{V}_{swht}	Turbine volumetric flowrate	m ³ /day
$\dot{V}_{w,RO}$	Incoming RO volumetric flowrate	m ³ /day
\dot{V}_{wp}	Incoming upper reservoir volumetric flowrate	m ³ /day
η_{hp}	Pump-side efficiency	-
η_{ht}	Turbine-side efficiency	-
$\eta_{RO,io}$	Fraction of pressure leaving RO system	-
η_{RO}	RO system recovery ratio	-
γ	Fraction of \dot{E}_r sent to IPHROS	-
γ_{RO}	Fraction of seawater in upper reservoir sent to RO system	-

$\pi_{f,k}$	Feed pressure for RO element k	Pa
$\pi_{m,k}$	Osmotic pressure along membrane of RO element k	Pa
$\pi_{p,k}$	Permeate osmotic pressure for RO element k	Pa
ρ_{fw}	Freshwater density	kg/m ³
ρ_{ht}	Discharge density	kg/m ³
$\rho_{o,RO}$	Brine density	kg/m ³
ρ_{sw}	Seawater density	kg/m ³
ρ_w	Saline water density	kg/m ³
$(P/A, i, t)$	Uniform series present-worth factor	-
A_m	Membrane surface area	m ²
$A_{w,k}$	Membrane water permeability coefficient for RO element k	m/hr/Pa
AE	PSH annual energy throughput	MWh
AVE	Annual value of electricity	\$
AVW	Annual value of fresh water	\$
$CAPEX_{PSH}$	Capital expenditures for PSH system	\$
$CAPEX_{RO}$	Capital expenditures for RO system	\$
$c_{c,k}$	Concentrate salt concentration for RO element k	mol/L
$c_{f,k}$	Feed salt concentration for RO element k	mol/L
$c_{fc,k}$	Average salt concentration in RO element k	mol/L
$c_{m,k}$	Salt concentration along membrane in RO element k	mol/L

$c_{p,k}$	Permeate salt concentration for RO element k	mol/L
g	Gravitational acceleration	m/s ²
h_L	Upper reservoir height	m
m	Scale factor	-
M_{salt}	Molar mass of salt	g/mol
N_{e1}	Number of RO elements in series in stage 1	-
N_{e2}	Number of RO elements in series in stage 2	-
N_{pv1}	Number of RO pressure vessels in parallel in stage 1	-
N_{pv2}	Number of RO pressure vessels in parallel in stage 2	-
NDP_k	Net driving pressure for RO element k	Pa
NPV	Net present value	\$
$OPEX_{PSH}$	Operational expenditures for PSH system	\$/year
$OPEX_{RO}$	Operational expenditures for RO system	\$/year
P	PSH plant capacity	MW
P_p	Permeate pressure	Pa
$P_{c,k}$	Concentrate pressure for RO element k	Pa
$P_{f,k}$	Feed pressure for RO element k	Pa
$P_{fc,k}$	Average pressure in RO element k	Pa
pf_k	Concentration polarization factor for RO element k	Pa
$Q_{c,k}$	Concentrate volumetric flowrate for RO element k	m ³ /hr

$Q_{f,k}$	Feed volumetric flowrate for RO element k	m ³ /hr
$Q_{fc,k}$	Average volumetric flowrate in RO element k	m ³ /hr
Q_{int}	Feed volumetric flowrate for RO element in second stage of RO system	m ³ /hr
$Q_{p,k}$	Permeate volumetric flowrate for RO element k	m ³ /hr
R_k	Fractional salt rejection rate for RO element k	-
rr_k	Recovery ratio for RO element k	-
S_f	Feed salinity	g/kg
S_{gkg}	Salinity in units of g/kg	g/kg
S_{ht}	Discharge salinity	g/kg
S_{mgL}	Salinity in units of mg/L	mg/L
$S_{o,RO}$	Brine salinity	g/kg
S_{sw}	Seawater salinity	g/kg
T	Temperature	°C
IPHROS	Integrated Pumped Hydro Reverse Osmosis System	
MDO	Multidisciplinary Design Optimization	
PSH	Pump Storage Hydropower	
RO	Reverse Osmosis	

References

- [1] J. Eliasson, The rising pressure of global water shortages, Nature 517 (7532) (2014) 6.

- [2] United Nations Department of Economic and Social Affairs, Water and cities, https://www.un.org/waterforlifedecade/water_cities.shtml, Last accessed on: 30 November 2022 (2014).
- [3] P. Rogers, Facing the freshwater crisis, *Scientific American* 299 (2) (2008) 46–53.
- [4] H. Fountain, In a first, U.S. declares shortage on Colorado River, forcing water cuts, *The New York Times* (2021, August 16).
- [5] R. D. Garreaud, J. P. Boisier, R. Rondanelli, A. Montecinos, H. H. Sepúlveda, D. Veloso-Aguila, The central chile mega drought (2010–2018): A climate dynamics perspective, *International Journal of Climatology* 40 (1) (2019) 421–439.
- [6] N. Ramos, A. Villegas, Chile announces unprecedented water rationing plan as drought enters 13th year, *Reuters* (2022, April 12).
- [7] J. Gillis, California drought is made worse by global warming, scientists say, *The New York Times* (2015, August 20).
- [8] J. Kim, K. Park, D. R. Yang, S. Hong, A comprehensive review of energy consumption of seawater reverse osmosis desalination plants, *Applied Energy* 254 (2019) 113652.
- [9] S. Rehman, L. M. Al-Hadhrani, M. M. Alam, Pumped hydro energy storage system: A technological review, *Renewable and Sustainable Energy Reviews* 44 (2015) 586–598.
- [10] O. J. Guerra, J. Eichman, P. Denholm, Optimal energy storage portfolio for high and ultrahigh carbon-free and renewable power systems, *Energy & Environmental Science* 14 (10) (2021) 5132–5146.
- [11] V. Jülch, Comparison of electricity storage options using levelized cost of storage (lcos) method, *Applied Energy* 183 (2016) 1594–1606.
- [12] R. Uría-Martínez, M. M. Johnson, R. Shan, U.S. hydropower market report, Tech. rep., U.S. Department of Energy - Office of Energy Efficiency & Renewable Energy - Water Power Technologies Office (2021).
- [13] K. Krüger, Pumped storage hydropower capabilities and costs, Tech. rep., Pumped Storage Hydropower International Forum (2021).

- [14] IEA, Grid-scale storage, <https://www.iea.org/reports/grid-scale-storage>, Last accessed on: 30 November 2022 (2022).
- [15] Y. Wu, T. Zhang, R. Gao, C. Wu, Portfolio planning of renewable energy with energy storage technologies for different applications from electricity grid, *Applied Energy* 287 (2021) 116562.
- [16] BP, Full report – statistical review of world energy 2021 - bp global, <https://www.bp.com/content/dam/bp/business-sites/en/global/corporate/pdfs/energy-economics/statistical-review/bp-stats-review-2021-full-report.pdf> (2021).
- [17] A. R. E. Agency, Cultana pumped hydro project knowledge sharing report, Tech. rep., Australian Government (2017).
- [18] ARUP, World’s largest saltwater pumped hydro for south australia, <https://www.arup.com/projects/cultana-pumped-hydroelectricity-storage>, Last accessed on: 23 March 2022.
- [19] Philstar Global, Coming to the Philippines: Seawater pumped energy storage technology, <https://www.philstar.com/business/science-and-environment/2023/06/01/2270719/coming-philippines-seawater-pumped-energy-storage-technology>, Last accessed on: 14 June 2023 (2023).
- [20] M. Faigon, Success behind advanced swro desalination plant, *Filtration + Separation* 53 (3) (2016) 29–31.
- [21] A. H. Slocum, M. N. Haji, A. Z. Trimble, M. Ferrara, S. J. Ghaemsaidi, Integrated Pumped Hydro Reverse Osmosis systems, *Sustainable Energy Technologies and Assessments* 18 (2008) 80–99.
- [22] I. M. ElNaghy, A. Y. Elbanhawy, A. H. Slocum, Assessing the potential and optimal scheduling of an Integrated Pumped Hydro Reverse Osmosis System for Egypt, in: *Applied Energy Symposium: MIT A+B*, Vol. 25, 2022.
URL <https://www.energy-proceedings.org/assessing-the-potential-and-optimal-scheduling-of-an-integrated-pumped-hydro-reverse-osmosis-system-for-egypt/>

- [23] S. Moreno-Leiva, J. Haas, W. Nowak, W. Kracht, L. Eltrop, C. Breyer, Integration of seawater pumped storage and desalination in multi-energy systems planning: The case of copper as a key material for the energy transition, *Applied Energy* 299 (2021) 117298. doi:<https://doi.org/10.1016/j.apenergy.2021.117298>. URL <https://www.sciencedirect.com/science/article/pii/S030626192100711X>
- [24] J. R. Martins, A. B. Lambe, Multidisciplinary design optimization: A survey of architectures, *AIAA Journal* 51 (9) (2013) 2049–2075.
- [25] R. Segurado, J. F. A. Madeira, M. Costa, N. Duić, M. G. Carvalho, Optimization of a wind powered desalination and pumped hydro storage system, *Applied Energy* 177 (2016) 487–499.
- [26] B. Zhou, B. Liu, D. Yang, J. Cao, T. Littler, Multi-objective optimal operation of coastal hydro-electrical energy system with seawater reverse osmosis desalination based on constrained nsga-iii, *Energy Conversion and Management* 207 (2020) 112533.
- [27] B. Liu, B. Zhou, D. Yang, Z. Yang, M. Cui, Optimal capacity planning of combined renewable energy source-pumped storage and seawater desalination systems, *Global Energy Interconnection* 2 (4) (2019) 310–317.
- [28] A. Pradhan, M. Marence, M. J. Franca, The adoption of seawater pump storage hydropower systems increases the share of renewable energy production in small island developing states, *Renewable Energy* 177 (2021) 448–460.
- [29] T. Ma, H. Yang, L. Lu, J. Peng, Pumped storage-based standalone photovoltaic power generation system: Modeling and techno-economic optimization, *Applied Energy* 137 (2015) 649–659.
- [30] T. Ma, H. Yang, L. Lu, J. Peng, Optimal design of an autonomous solar–wind–pumped storage power supply system, *Applied Energy* 160 (2015) 728–736.
- [31] Z. Al Suleimani, V. R. Nair, Desalination by solar-powered reverse osmosis in a remote area of the sultanate of oman, *Applied Energy* 65 (1-4) (2000) 367–380.

- [32] D. R. Prathapaneni, K. Detroja, Optimal design of energy sources and reverse osmosis desalination plant with demand side management for cost-effective freshwater production, *Desalination* 496 (2020) 114741.
- [33] T. Novosel, B. Ćosić, T. Pukšec, G. Krajačić, N. Duić, B. V. Mathiesen, H. Lund, M. Mustafa, Integration of renewables and reverse osmosis desalination – case study for the jordanian energy system with a high share of wind and photovoltaics, *Energy* 92 (2015) 270–278.
- [34] J. A. Carta, P. Cabrera, Optimal sizing of stand-alone wind-powered seawater reverse osmosis plants without use of massive energy storage, *Applied Energy* 304 (2021) 117888.
- [35] E. M. A. Mokheimer, A. Z. Sahin, A. Al-Sharafi, A. I. Ali, Modeling and optimization of hybrid wind–solar-powered reverse osmosis water desalination system in saudi arabia, *Energy Conversion and Management* 75 (2013) 86–97.
- [36] J. G. Wijmans, R. W. Baker, The solution-diffusion model: A review, *Journal of Membrane Science* 107 (1-2) (1995) 1–21.
- [37] DuPont Water Solutions, WAVE Design Software, <https://www.dupont.com/water/resources/design-software.html> (Version 1.82).
- [38] K. L. McMordie Stoughton, X. Duan, E. M. Wendel, Reverse osmosis optimization, Tech. rep., Pacific Northwest National Laboratory (2013).
- [39] J. Wang, D. S. Dlamini, A. K. Mishra, M. T. Pendergast, M. C. Y. Wong, B. B. Mamba, V. Freger, A. R. D. Verliefde, E. M. V. Hoek, A critical review of transport through osmotic membranes, *Journal of Membrane Science* 454 (2014) 516–537.
- [40] D. W. Solutions, Filmtec reverse osmosis membranes technical manual, Dow Inc., 8th Edition (September 2021).
- [41] J. D. Seader, E. J. Henley, Separation process principles, 2nd Edition, John Wiley & Sons, 2006.
- [42] Y. Okamoto, J. H. Lienhard, How ro membrane permeability and other performance factors affect process cost and energy use: A review, *Desalination* 470 (2019) 114064.

- [43] D. P. D. Sheet, Filmtec™ seamaxx™-440 element, <https://www.dupont.com/content/dam/dupont/amer/us/en/water-solutions/public/documents/en/R0-FilmTec-Seamaxx-440-PDS-45-D00689-en.pdf> (August 2021).
- [44] S. Moza, Sahilm89/lhsmdu: Latin hypercube sampling with multi-dimensional uniformity (lhsmdu): Speed boost minor compatibility fixes, <https://dx.doi.org/10.5281/zenodo.2578780> (2020).
- [45] J. L. Deutsch, C. V. Deutsch, Latin hypercube sampling with multidimensional uniformity, *Journal of Statistical Planning and Inference* 142 (3) (2012) 763–772.
- [46] K. G. Nayar, M. H. Sharqawy, L. D. Banchik, J. H. Lienhard V, Thermophysical properties of seawater: A review and new correlations that include pressure dependence, *Desalination* 390 (2016) 1–24.
- [47] SAS Institute Inc., JMP® (Version 16).
- [48] H. Jin, Q. Song, X. Hu, Auto-keras: An efficient neural architecture search system, in: *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, ACM, 2019, pp. 1946–1956.
- [49] M. Haefner, M. N. Haji, Integrated pumped hydro reverse osmosis system optimization with enhanced reverse osmosis modeling, in: *Volume 23: Sustainable Energy Solutions for a Post-COVID Recovery towards a Better Future: Part VI*, International Conference on Applied Energy, Energy Proceedings, 2021.
URL <https://www.energy-proceedings.org/integrated-pumped-hydro-reverse-osmosis-system-optimization-with-enhanced-reverse-osmosis-modeling/>
- [50] S. Jenkins, J. Paduan, P. Roberts, D. Schlenk, J. Weis, Management of brine discharges to coastal waters: recommendations of a science advisory panel, Tech. rep., Southern California Coastal Water Research Project (2012).
- [51] MathWorks, gamultiobj algorithm, <https://www.mathworks.com/help/gads/gamultiobj-algorithm.html>, Last accessed on: 30 November 2022 (2007).

- [52] M. K. Wittholz, B. K. O'Neill, C. B. Colby, D. Lewis, Estimating the cost of desalination plants using a cost database, *Desalination* 229 (1) (2008) 10–20. doi:<https://doi.org/10.1016/j.desal.2007.07.023>.
URL <https://www.sciencedirect.com/science/article/pii/S0011916408002683>
- [53] Advisian, The cost of desalination, <https://www.advisian.com/en/global-perspectives/the-cost-of-desalination>, Last accessed on: 30 November 2022 (2017).
- [54] K. Mongird, V. Viswanathan, J. Alam, C. Vartanian, V. Sprenkle, R. Baxter, 2020 grid energy storage technology cost and performance assessment, Tech. rep., Pacific Northwest National Laboratory, Mustang Prairie Energy (December 2020).
- [55] US Energy Information Administration, US Electricity Profile 2021, <https://www.eia.gov/electricity/state/>, Last accessed on: 30 November 2022 (2022).
- [56] UNC School of Government Environmental Finance Center, California small water systems rates dashboard, <https://dashboards.efc.sog.unc.edu/ca>, Last accessed on: 30 November 2022 (2020).
- [57] G. Oladosu, C. Sasthav, Hydropower capital and O&M costs: An exploration of the FERC Form 1 data, Tech. rep., Oak Ridge National Laboratory (February 2022).
- [58] V. Viswanathan, K. Mongird, R. Franks, X. Li, V. Sprenkle, R. Baxter, 2022 grid energy storage technology cost and performance assessment, Tech. rep., Pacific Northwest National Laboratory, Mustang Prairie Energy (August 2022).
- [59] M. M. Generous, N. A. A. Qasem, U. A. Akbar, S. M. Zubair, Techno-economic assessment of electrodialysis and reverse osmosis desalination plants, *Separation and Purification Technology* 272 (2021) 118875.
- [60] A. V. Dudchenko, T. V. Bartholomew, M. S. Mauter, Cost optimization of multi-stage gap membrane distillation, *Journal of Membrane Science* 627 (2021) 119228.

- [61] M. Qasim, M. Badrelzaman, N. N. Darwish, N. A. Darwish, N. Hilal, Reverse osmosis desalination: A state-of-the-art review, *Desalination* 459 (2019) 59–104.