Predictive Time-Variant Photovoltaic Electrodialysis Reversal: A Novel Design Optimization using Predictive Machine Learning and Control Theory

by

Grace B. Connors

Submitted to the Department of Mechanical Engineering
in partial fulfillment of the requirements for the degree of

Master of Science in Mechanical Engineering

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2021

Author .....
Department of Mechanical Engineering
May 14, 2021

Certified by .....
Amos G. Winter V
Associate Professor
Thesis Supervisor

Accepted by .....
Nicolas Hadjiconstantinou
Professor, Graduate Officer

Predictive Time-Variant Photovoltaic Electrodialysis Reversal: A Novel Design Optimization using Predictive Machine Learning and Control Theory

by

Grace B. Connors

Submitted to the Department of Mechanical Engineering
on May 14, 2021, in partial fulfillment of the
requirements for the degree of
Master of Science in Mechanical Engineering

Abstract

This paper introduces a novel control theory and system design optimization to reduce the Levelized Cost of Water (LCOW) and maximize the reliability of Photovoltaic Electrodialysis Reversal (PV-EDR) groundwater desalination systems. This work aims to exploit the relationship between water production and Specific Energy Consumption (SEC) for time-variant PV-EDR systems and to introduce a control system that optimizes the energy management strategy with a goal of maximizing water production as well as operates efficiently with respect to energy utilization. The novel control theory introduced in this paper consists of a machine learning algorithm used to predict the future solar irradiance, a model predictive controller to use this prediction to plan the best utilization of energy between the desalination system and energy storage, and a lower level controller that determines the optimal flow rate and voltage of the EDR system based on the power available for desalination. Furthermore, this paper aims to use this control theory to build a design tool that can determine optimal PV-EDR system configurations based on geographical constraints. The system design optimization is then tested for a case study of a rural village in India. As compared to previous works, this control theory reduces the LCOW by 7%, or \$0.15/m^3, while meeting the target water production every day. Moreover, this study demonstrates the flexibility of this design tool as well as the impact of the design assumptions through a sensitivity analysis. This study determines how to design and control PV-EDR systems to minimize cost and maximize reliability, improving the commercial viability of using PV-EDR systems as a primary water desalination solution.

Thesis Supervisor: Amos G. Winter V
Title: Associate Professor

Acknowledgments

I would like to thank my Principal Investigator, Professor Amos Winter, for letting me work on a problem I care deeply about, and giving me the freedom to explore solutions to the fullest. I'd also like to thank him for continuing to be my cheerleader throughout the past 2 years. I'd like to thank Dr. Simone Gelmini for his unbelievable guidance and mentorship throughout the last 2 years, and for without whom, I know that this thesis would not be finished. Simone has shown me the value of controls, and he has pushed me to more deeply understand all the ways in which we can leverage them to improve our current systems. I'd also like to thank Sahil Shah, who continues to guide my research, and my thinking, in better directions with every conversation. He has been a mentor to me both personally and professionally, and he has contributed so greatly to my experience here in the GEAR Lab.

I'd also like to thank Dr. Susan Amrose, Elizabeth Brownell, and Shane Pratt for their wise comments, helpful conversations, and continuing to push me as a researcher. I'd also like to thank the entire GEAR desal team for supporting me, asking valuable questions, and assisting me in lab countless times.

I'd also like to thank Dario Nava for his incredible contributions to the framing of the Energy Management strategy, and for being an absolute joy to work with for my first 3 months of graduate school.

I'd like to thank my family and friends, and especially Colin, who continue to support me and fill my non-working time with so many laughs.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Salinity in groundwater is increasing around the world [15, 30, 47]. Groundwater serves as the primary drinking supply for over 2.5 billion people, or over 30% of the Earth's population [45]. Over-exploitation of groundwater resources has increased the salinity of this drinking water, requiring desalination solutions [47]. According to the World Health Organization, drinkable groundwater has Total Dissolved Solids (TDS) less than 500 ppm [44]. Typical groundwater salinities can range from 350 - 3200 ppm, and therefore can require desalination to bring the salinity to within recommended drinking water ranges [43, 36]. The most widely adopted desalination solution is Reverse Osmosis (RO) [22]. RO is a mature desalination technology which is manufactured at scale. Due to the ubiquity of RO technology, it is inexpensive to implement, especially in terms of capital cost.

Energy is the single largest operating expense in desalination systems [22]. At typical groundwater salinities (i.e., below 2000 ppm), RO requires twice as much energy to produce the same amount of water as some alternative technologies, like Electrodialysis Reversal (EDR). Moreover, EDR can operate at higher recovery ratios than some RO systems, especially at the scales most applicable for community water solutions (i.e. $< 10\ m^3$) [49]. Small scale RO systems can have recoveries ranging from 30-60%, while small scale EDR systems consistently operate at 80-90% recoveries [49]. As a consequence, RO requires significantly more brine disposal, increasing the operation cost of RO as compared to EDR. Although more energetically and water

efficient, EDR is a more expensive technology in terms of capital cost. But, the cost gap between RO and EDR is reduced significantly when the costs are accounted for over the lifetime of the desalination system.

Alternative technologies, like Electrodialysis Reversal (EDR), are less mature and more expensive to implement, thus less popular for water desalination systems. However, the work in Wright and Winter V [49] has demonstrated that for low groundwater salinities (i.e., below 2000 ppm), EDR is 50% more energy efficient than RO. RO requires twice as much energy to produce the same amount of water as EDR, thus doubling the cost of energy. Energy is the single largest operating expense in desalination systems [22]. Moreover, EDR can operate at higher recovery ratios than some RO systems, especially at the scales most applicable for community water solutions (i.e. $< 10\ m^3$) [49]. Indeed, small scale RO systems can have recoveries ranging from 30-60%, while small scale EDR systems consistently operate at 80-90% recoveries [49]. As a consequence, RO requires significantly more brine disposal, which increases the operation cost with respect to EDR [48]. Although RO is more favorable in terms of capital cost, the cost gap between RO and EDR is reduced significantly when costs are accounted over the lifetime of the desalination system.

Photovoltaics, or PV, are a viable energy source for desalination systems because there is overlap between locations with increasing groundwater salinity and photovoltaic energy potential [17]. Moreover, in regions where grid power is non-existent or unreliable, renewable energy sources, like PV, could become the primary power source. As EDR requires less energy than RO, a smaller area of PV is needed for powering EDR desalination systems compared to RO. Thus, the capital cost for the power system in PV-EDR is significantly less than the power system for PV-RO. When determining the viability of desalination technologies, a system designer is considering both the lifetime cost of the system, accounted for as a combination of capital cost and operating costs, as well as the reliability of the system, or the ability of the system to supply people with water as they need it. In places where grid power is unreliable, PV-EDR is one promising solution to providing low-cost groundwater desalination to areas with rising groundwater salinity. However, high reliability and

competitive global costs are critical to making PV-EDR a successful technology for becoming the primary technology supplying drinking water to these areas [25].

PV-EDR systems have been examined since the second half of the 1980s through successful field pilots [16, 34]. These pioneering works designed PV-EDR systems for continuous operations without energy buffers. Improvements in the system design have been achieved when Gonzalez [28] combined the concept of water production in batches with an energy storage. That work was also the first to note the importance of decoupling power generation and consumption, which greatly increased the flexibility of the system. Based on those insights, the work in Bain et al. [19] proved through a pilot study that energy and water storage are key to meeting a high reliability target, defined as producing all the water demanded by the end users every day. The authors remarked that low-irradiance periods, especially when occurring for multiple days in a row, can only be overcome with proper energy and water storage. However, the size, and therefore cost, of energy and water storage buffers affect the commercial viability of PV-EDR systems. As these buffers can be expensive, PV-EDR has usually struggled to be cost-competitive with on-grid RO.

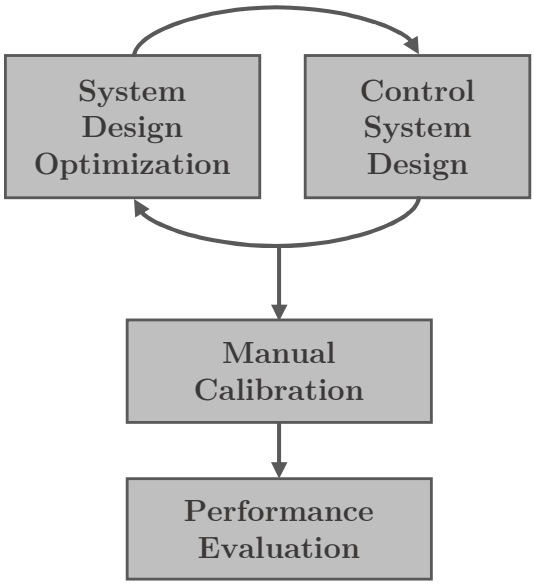
Beyond the introduction of energy storage to increase the robustness of PV-EDR systems, there have also been advancements in the control of PV-EDR systems to reduce cost. The work in Shah et al. [46] introduced a novel control strategy to vary the voltage applied to the EDR stack in order to maximize desalination rate. This innovation led to a significant increase in the amount of water desalinated for the same EDR stack, thus lowering the system cost per cubic meter of water produced. To further reduce energy storage costs, the work in LeHenaff [37] introduced a control theory specifically tailored for PV-EDR systems that adapts the system power consumption to match the solar power currently available. To accomplish this result, the control algorithm optimizes both the system flow rate and voltage to ensure the most optimal use of available power to achieve the maximum desalination rate. The work also demonstrated that this control algorithm allows to optimally size the system components, resulting in a significant reduction of the energy and water buffers size, as compared to state of the art PV-EDR systems.

Furthermore, results presented in LeHenaff [37] demonstrated that by directly coupling power generation with water production, it is possible to operate the system without the need of an energy storage unit. However, the work also determined that the SEC increased with the water production rate. Therefore, despite producing large volumes of water, operating the system at high power would result in less efficient operations. As a consequence, direct-drive systems would be subject to inefficient operations during the hours with highest solar irradiance, which are crucial to produce the daily water demand. These inefficiencies may cause the system to be oversized, and a more optimal balance between storing power and using power could maximize reliability at lowest cost.

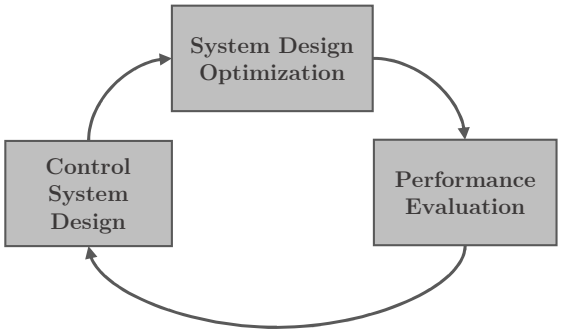
Engineering a novel control architecture to adapt water production and optimize storage use could enable the design of optimal PV-ED systems to maximize system reliability at lowest cost. Specifically, the designed system must achieve a competitive cost to be commercially viable, inclusive not only of capital costs but also operating and maintenance costs, and replacement costs over the plant lifetime. This paper investigates how to build a control strategy that could be optimal and flexible to any design conditions, including the geography of where the system is used. This objective poses an additional challenge: To implement an optimal control strategy, prediction capabilities of future system states are needed [35]. In off-grid applications, this requirement means estimating future power availability to optimally manage water production and control the system energy flows. To this end, the present study explores how to use machine learning techniques to perform this prediction with minimal uncertainty, leveraging historic weather data to reduce sensing and computing costs.

To achieve an optimal PV-EDR control theory and system design, this paper also introduces a novel workflow for PV-EDR system design and optimization, which is depicted pictorially in Figure 1-1. In a traditional workflow, the control system development is decoupled from the system design optimization, which could lead to conservative results. Instead, this workflow aims to exploit the machine performance by a continuous joint improvement of the control system and the system design based

on the performance of the system. If successful, this novel workflow would maximize control system performance and, in turn, lead the system design to price parity with respect to commercially available systems currently used in similar applications.



(a) Traditional system workflow



(b) Novel system workflow

Figure 1-1: Traditional system workflow (a) versus proposed one (b). In a traditional workflow, system performance affects the combination of the control system and system design. Following what we propose, the performance of the system can be directly leveraged to improve the control system, which then influences the system design optimization.

Chapter 2

Background

To improve the control strategy of an EDR system, it is critical to understand the operation of EDR systems. Moreover, we need to understand the state of the art control theory for EDR technologies in order to leverage any opportunities for improvement.

2.1 General Electrodialysis Reversal System Description

ED is a membrane-based desalination technology. An ED unit, referred to as an ED stack, consists of pairs of electrodes and pairs of ion exchange membranes. Cation Exchange Membranes (CEM) are able to pass cations across the membrane, and Anion Exchange Membranes (AEM) are able to pass anions across the membrane. By arranging the membranes in pairs, alternating CEMs and AEMs, and applying a voltage to the electrodes to induce ion movement, ions can move across membranes based on their charge. This ion movement creates channels of brine, where many ions are concentrated, and diluate where there are relatively few ions. ED Reversal (EDR) operates under the same principles as ED, but the polarity of the electrodes is reversed, which reverses the brine and diluate channels. This reversal can reduce scale formation within the ED stack [29].

To fully exploit the potential of EDR technology in community-scale systems,

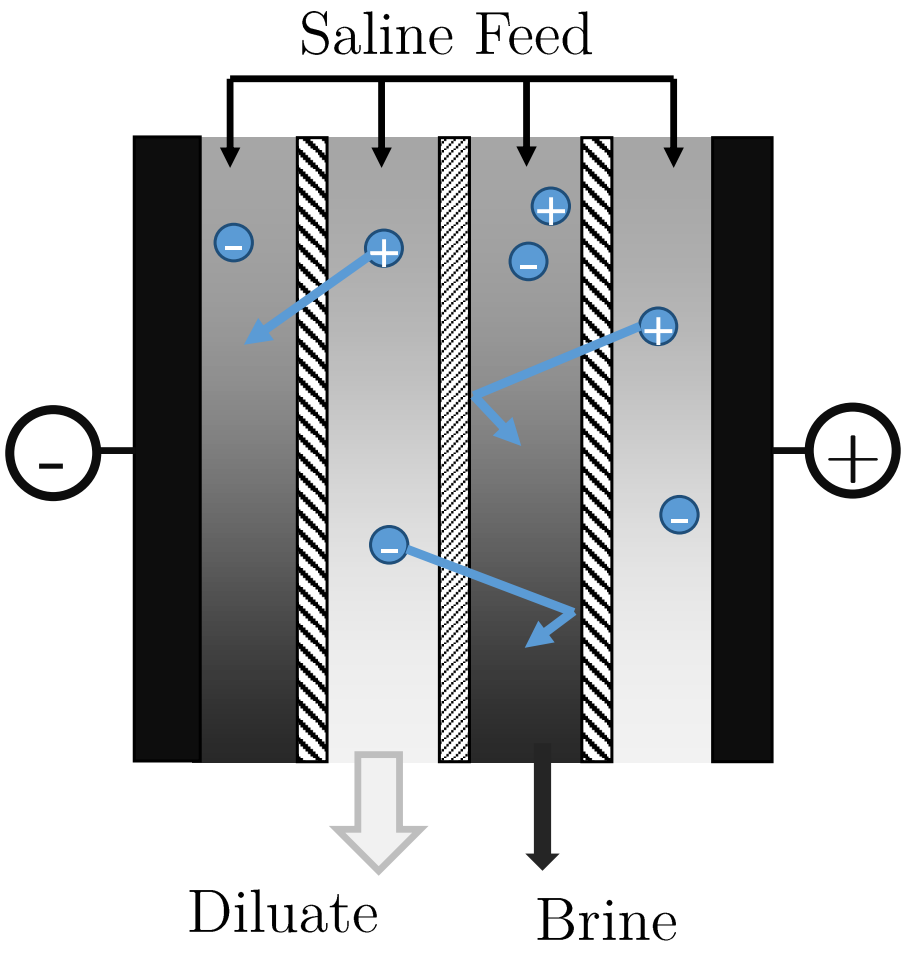


Figure 2-1: [When a voltage is applied to an ED stack with alternating Ion Exchange Membranes, ions, depending on charge, can move across either the CEM or AEM, but are trapped by the other membrane.

Wright and Winter V [49] demonstrated that there is an advantage to batch mode operation of EDR systems over conventional continuous operation. In batch mode operation, water flows are recirculated between the stack channels and the tanks until the diluate tank concentration reaches a predetermined value. Figure 2-2 shos the schematic of the batch system configuration, including the tanks needed for recirculation. Batch mode significantly decreases the electrode and membrane area needed for water production because water circulates through the same membranes multiple times, as compared to continuous EDR where the water passes through the system only once. Therefore, due to the reduction of electrode and membrane area, the LCOW for an EDR system with the batch mode is less than the LCOW of a

continuous-operated EDR system. Moreover, by operating the system in batch mode, the voltage applied to the system can be tuned to desalinate the water precisely to the target salinity throughout the batch process [46]. At the beginning of the batch, the voltage applied can be higher to desalinate water more rapidly. As the conductivity in the diluate stream decreases throughout the batch, the voltage applied is lowered. By contrast, in continuous mode, the voltage applied is constant. This technique of voltage tuning for systems operated in batch mode minimizes the energy required to produce water.

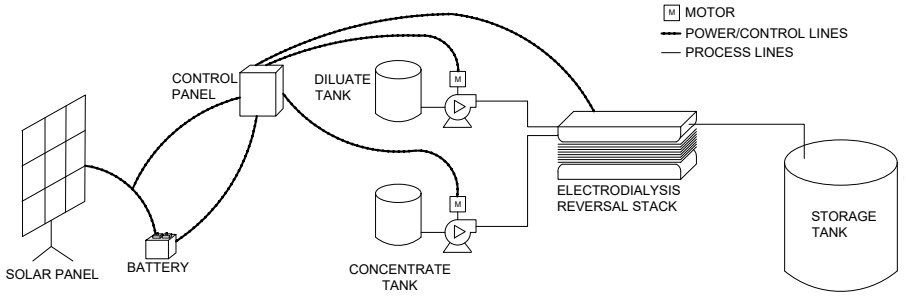


Figure 2-2: A schematic of the PV-EDR system with all of its components.

PV-EDR is the combination of EDR systems with photovoltaic panels (PV) as the power source of the system. A system utilizing PV-EDR technology consists of

- an array of solar panels;
- a battery pack, for any application requiring energy storage;
- two tanks, one for the diluate and one for brine;
- two pumps, one for each of the diluate and concentrate streams;
- storage tank for holding diluate after a batch is completed;
- and an EDR stack.

The system diagram with the power/control lines and the hydraulic process lines is shown in Figure 2-2.

2.2 Time-Variant Control Theory of EDR systems

The coupling of PV with EDR systems poses a challenge because power generated from PV varies throughout the day while EDR systems traditionally require a constant power supply. In order to manage the mismatch between the PV power supply and the EDR system power demand, energy storage is needed. To cut energy storage costs, LeHenaff [37] investigated the direct use of all the solar irradiance power available at any given time, adapting ED operations to be time variant. To maximize desalination rate, the proposed strategy introduced a method for controlling flow rate and voltage based on the instantaneous solar power. LeHenaff [37] focused on designing direct drive systems, in which the system can operate with no energy storage units. In order to achieve 99% reliability, an energy unit was added to the system retroactively. Without this energy unit, the system would not be 99% reliable. Despite the effectiveness of this approach, this strategy may be suboptimal as the energy storage system is not optimally used through the system operation. The present work builds on LeHenaff's 2019 optimal voltage and flow rate selection to improve system level operations and explicitly account for the energy storage, aiming to optimally allocate power between the energy storage and the EDR stack to maximize efficiency and minimize cost.

He et al. [32] discussed a mathematical framework for determining the optimal flow rate and voltage that can maximize the desalination rate for the power available. Starting from the complete electrochemical model discussed in Wright et al. [50], the proposed ED model solves, in closed form, the optimal flow rate and voltage accounting for the hydraulic dynamics (i.e., mass transfer and pressure drop) and power consumption of each components of the ED system. Specifically, the desalination rate can be expressed in terms of the rate of change of the concentration of the diluate and concentrate tanks (shown in Figure 2-2) and channels through the computation of the mass balance. For each tank, the mass balance leads to:

$$\frac{dC_{\chi,t}}{dt} = \frac{q_\chi}{\upsilon_{\chi,t}} \left(C^b_{\chi,o} - C^b_{\chi,i} \right), \quad \chi = \{c, d\}, \tag{2.1}$$

where $C_{\chi,t}$ is the concentration for the diluate (i.e., $\chi = d$) and concentrate (i.e., $\chi = c$) tank respectively; $\upsilon_{\chi,t}$ is the volume of each tank; q_χ represents the flow rate for each channel; and C^b_χ is the bulk concentration of each channel, evaluated at the inlet ($C^b_{\chi,i}$) and the outlet ($C^b_{\chi,o}$) of the stack, respectively.

For the channels within the ED stack, the mass balance equations are discretized into 5 cells (denoted with $\bar{\iota}$), as discussed in [50]. Discretization allows the conductivity to change across the stack length, which more closely matches the physical behavior of the ED stack, which is not physically segmented. By denoting with ι the i-th discretization cell, the bulk concentrations for the diluate and concentrate channels are given by:

$$\begin{aligned}\frac{dC^b_{d,\iota}}{dt} = \frac{1}{NV^{cell}_\iota}\Bigg[q_d\left(C^b_{d,\iota-1} - C^b_{d,\iota}\right) - \frac{N\phi I_\iota}{zF} \\ + \frac{NA_\iota D^{AEM}\left(C^{AEM}_{c,\iota} - C^{AEM}_{d,\iota}\right)}{l^{AEM}} \\ + \frac{NA_\iota D^{CEM}\left(C^{CEM}_{c,\iota} - C^{CEM}_{d,\iota}\right)}{l^{AEM}}\Bigg]\end{aligned} \tag{2.2a}$$

$$\begin{aligned}\frac{dC^b_{c,\iota}}{dt} = \frac{1}{NV^{cell}_\iota}\Bigg[q_c\left(C^b_{c,\iota-1} - C^b_{d,\iota}\right) - \frac{N\phi I_\iota}{zF} \\ - \frac{NA_\iota D^{AEM}\left(C^{AEM}_{c,\iota} - C^{AEM}_{d,\iota}\right)}{l^{AEM}} \\ - \frac{NA_\iota D^{CEM}\left(C^{CEM}_{c,\iota} - C^{CEM}_{d,\iota}\right)}{l^{AEM}}\Bigg],\end{aligned} \tag{2.2b}$$

where N is the number of cell pairs; I_ι is the applied current to the discretization segment; z is the ion charge number; ϕ is the current leakage factor; F is Faraday's Constant; A_ι is the area of the discretization segment; D^{AEM} and D^{CEM} are the diffusion coefficients of the solute in the AEMs and CEMs; l^{AEM} and l^{CEM} are the thickness of the AEM and CEMs; and $C^{AEM}_{c,\iota}$, $C^{AEM}_{d,\iota}$, $C^{CEM}_{c,\iota}$ and $C^{CEM}_{d,\iota}$ are the concentrations of the diluate and concentrate streams at the interface with adjacent CEMs and AEMs. This differential equation is used in the ED system model to determine the change in concentration expected for one pass through the ED stack.

When extrapolated to the entire batch, solving this differential equation can determine how long each batch would be for a given system and current. More details on how to calculate those concentrations can be found in Wright et al. 2018 [50].

Starting from the diluate concentration, it is possible to calculate the flow rate and current that maximizes the desalination rate for an ED system. The applied current is constrained by the limiting current density, which is defined as the maximum current density beyond which water molecules will disassociate at the electrode surface, and result in dangerous hydrogen gas formation. The limiting current density can be estimated by setting the concentration at either the AEM or CEM to zero, given by:

$$i_{lim,\iota}^{+,-} = \frac{C_{d,\iota}^{b} z F \kappa}{\tau^{AEM,CEM} - \tau_{+,-}}, \quad \iota = 1, \ldots, \bar{\iota}, \tag{2.3}$$

where $\tau^{AEM,CEM}$ is the transport number of the counterion in the AEM or CEM membrane; $\tau^{+,-}$ is the transport number of the cations or anions in the bulk solution; and κ is the boundary-layer mass transfer coefficient, which scales with $\sqrt{Q}$ [50]. In practice, we introduce a safety factor r (set to 0.7 [42]) on the applied current density to avoid operating the system at the limiting current density. Therefore, for each discretization cell, the applied current density is computed as:

$$i_{\iota} = r i_{lim,\iota}, \quad \iota = 1, \ldots, \bar{\iota}. \tag{2.4}$$

The voltage of the EDR stack is applied based on the adjusted maximum current density and the modeled channel resistances,

$$\begin{aligned} V_{ED} = V_{el} &+ N\big(V_{\iota}^{CEM} + V_{\iota}^{AEM}\big) \\ &+ N i_{\iota}\big(R_{d,\iota}^{b} + R_{c,\iota}^{b} + R_{\iota}^{BL} \\ &\qquad + R_{\iota}^{AEM} + R_{\iota}^{CEM}\big), \end{aligned} \tag{2.5}$$

where subscripts d and c represent the diluate and concentration channels, repsectively. b denotes the bulk flow. V_{el} is the electrode potential difference, whereas V_{ι}^{CEM} and V_{ι}^{AEM} are the potentials across CEMs and AEMs repsectively. The resistances

$R^b_{d,\iota}$, $R^b_{c,\iota}$, R^{BL}_{ι}, R^{AEM}_{ι}, and R^{CEM}_{ι} are related to the bulk flows, the boundary layers of the channels, and the membranes, respectively. The resistances used are calculated using the same equations presented in Wright et al. 2018 [50].

To calculate the total current applied to the stack, we sum up the current density for each discretized segment,

$$I_{ED} = \phi_A \left(\frac{WL}{\bar{\iota}} \right) \sum_{\iota=1}^{\bar{\iota}} i_\iota, \tag{2.6}$$

where W and L are the width and length of the flow channel, $\bar{\iota}$ is the number of discretized segments, and ϕ_A is the open-area porosity of the channel spacer. Thus, the total EDR power is calculated,

$$P_{ED} = V_{ED} I_{ED}. \tag{2.7}$$

The pumping power is also dependent on the flow rate, q,

$$P_{pump} = 2q\Delta p, \tag{2.8}$$

where Δp is the pressure drop through the stack. The pressure drop is estimated using the Darcy-Weisbach equation:

$$\Delta p = \frac{\rho_{aq} f L u_v^2}{4h}, \tag{2.9}$$

where ρ_{aq} is the density of the aqueous solution, f is the Darcy friction factor, h is the channel height, u_v is the void channel flow velocity.

This system is also constrained in terms of power, where the total power needed by the system must be equal to the power available,

$$P_{ED} + P_{pump} = P_{Total} \tag{2.10}$$

Combining these principles, it is possible to run an optimization routine to select

the optimal q and V to maximize desalination rate subject to constraints on i_{lim} and p. Figure 2-3 shows the optimization routine for how the optimal q and V are determined through an iterative model to match the power available with the power draw from the ED system [32]. Figure **??** shows the experimental results from this control strategy from a pilot scale field test performed in 2019 [37].

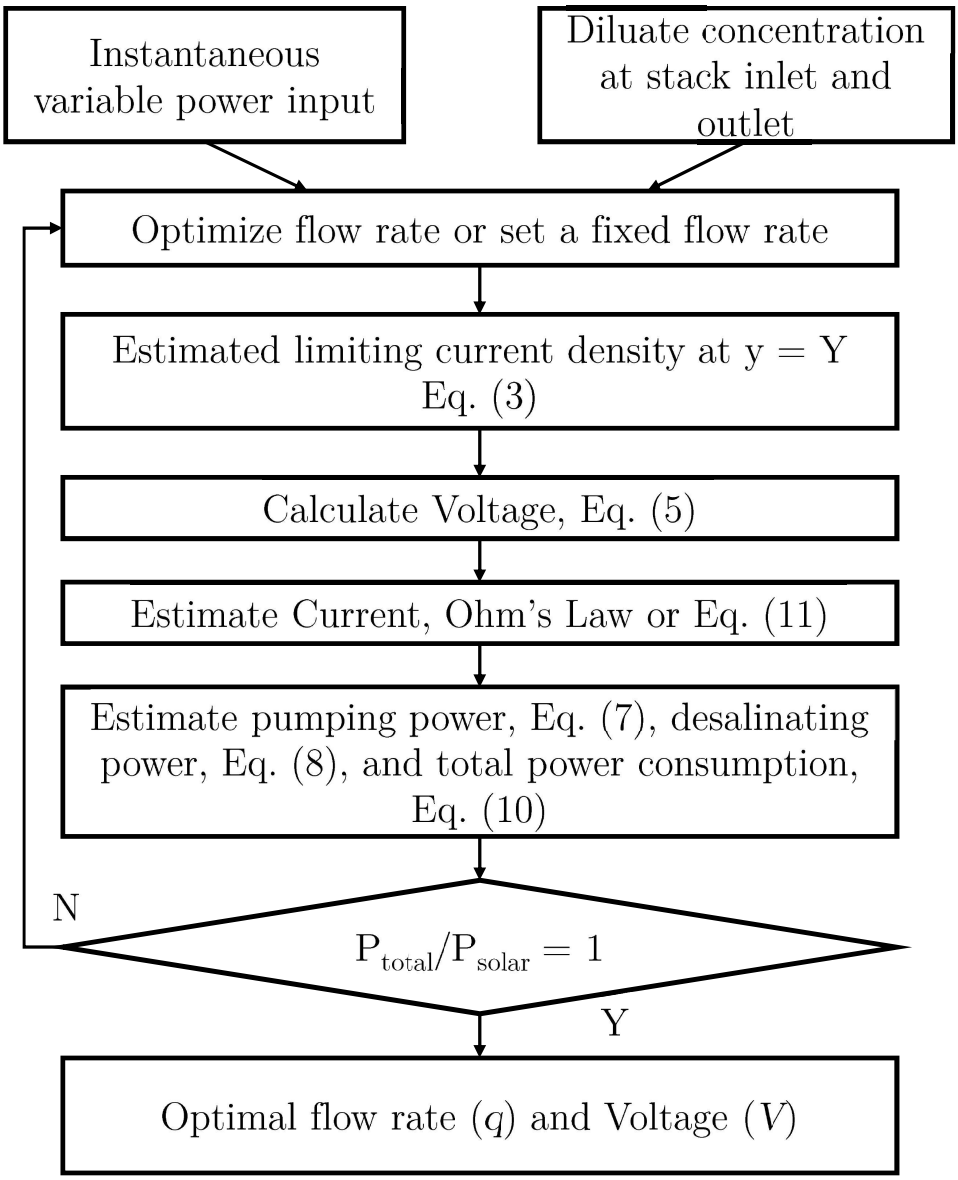


Figure 2-3: Flow Chart of the Model-Based Controller for Time-Variant ED Operation [32]

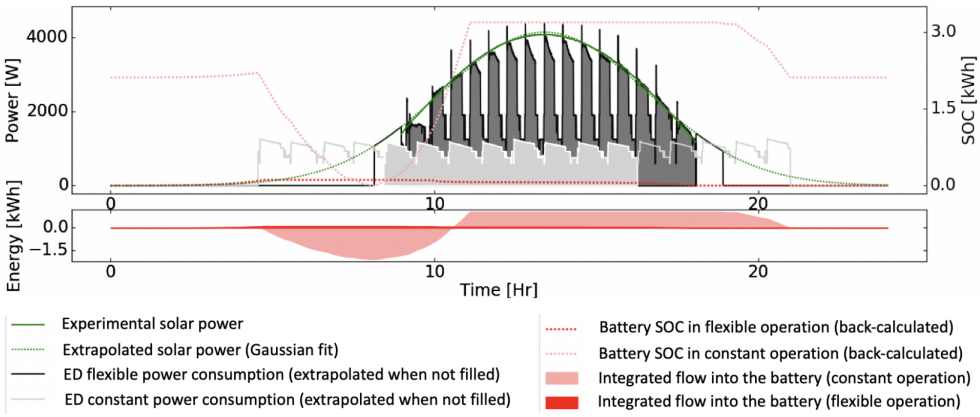


Figure 2-4: Experimental results from two daily tests on a full-scale flexible operation PV-EDR pilot system in New Mexico: time-variant and constant operation were run on 03/17/2019 and 03/11/2019, respectively. Raw power profiles are extrapolated to a complete and common full day. The difference between generated and consumed power over time is analyzed to derive the integrated flow into batteries, the minimum battery capacity required by a real installation to produce the desired water demand and the corresponding state of charge [37].

2.2.1 Batch Startup Control

Despite the demonstrated effectiveness of the optimal flow rate and voltage control method, experimental evidence showed that the model used for controlling the system was underperforming at the batch startup, causing a large power overshoot at the beginning of each batch, as shown in Figure **??** [37]. To compensate for this overshoot and avoid system failure, power storage is required, which would, in turn, increase capital costs.

This power overshoot is the result of the underlying system dynamics changing rapidly at the start of the batch. The voltage applied creates a transient within the stack conductivity, with a characteristic time shorter than the control time step. Furthermore, the conductivity probes are located beyond the outlet of the stack, within piping, creating a time delay between the conductivity transient and measurement. These conductivity measurements are used to estimate the stack current, I_{ED}, characterizing the ED system through the full system model as validated in He et al. 2020 [31].

To avoid the power overshoot and subsequent system design consequences, a refinement of the model for controlling the flow rate and voltage has been introduced. Specifically, an investigation of the EDR dynamics demonstrated that 1) the dynamic behavior at the system startup is significantly different than at steady-state, and 2) voltage and flow rate have different transients that could be better exploited for minimizing the power overshoot. Therefore, stack current, which was previously estimated using conductivity measurements that were affected by both voltage and flow rate dynamics, is now estimated using the Ohm's law, decoupling the fast electric dynamics from the slower hydraulic one. Furthermore, the short time characteristic of the electric dynamics allows the stack resistance to be modeled through a quasi-static approximation, and thus we can estimate the stack current using only its past estimates and voltage measurements. Stack Current was previously estimated using Equation **??** with the measured conductivity,

$$I_{approx} = \frac{Q_d(C^b_{d,1} - C_{d,\bar{\imath}})\zeta F}{N\phi} \tag{2.11}$$

but is now estimated using Ohm's law, using the applied Voltage, V, and assuming the resistance of the stack does not change between steps.

The comparison between the batch startup power using the two methodologies for current estimation are shown in Figure 2-5. As shown, the current estimation strategy based on conductivity measurements would experience a peak power draw of over 5 kW, taking roughly 13 seconds to settle to the final value of 1 kW. The proposed method has no power overshoot at startup, in fact, underestimating power consumption for the roughly 13 seconds of settling time. The two settling times are comparable between the two strategies, but by eliminating the power overshoot at the beginning of the batch, we can eliminate the need for expensive energy storage.

2.3 Water Production Curve Generation

With the elimination of the batch startup power overshoot, the system can now be assumed to operate in steady state for water production throughout the entire

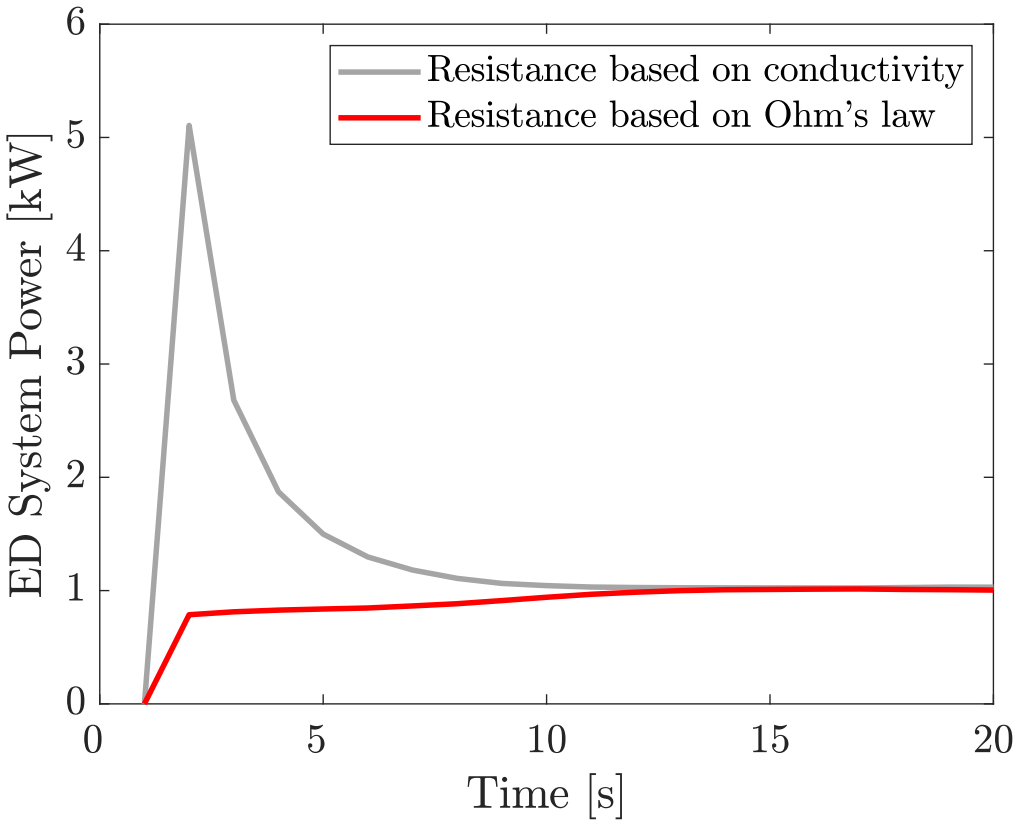


Figure 2-5: Simulated results of the previous versus proposed current estimation strategy simulated power consumption at the batch startup. The conductivity strategy estimated current through conductivity measurements, resulting in a large power overshoot especially at the beginning of the batch. The Ohm's Law strategy estimates current by determining the stack resistance through the previous current and voltage measurements, and assumes the stack resistance does not change between time steps. In both of these batches, we are simulating the time-variant control strategy using a fixed power input of 1 kW as opposed to solar power.

batch. Therefore, we can accurately simulate a fixed-volume batch of water produced for any amount of input power for any chosen system design using the previously described robust EDR system model [50]. By simulating batches with a range of input powers, it is possible to determine the difference in batch times associated with these batches simulated at different powers, and use this difference in time to calculate corresponding water production rates (expressed in m^3/h) and the associated SEC (kWh/m^3). An example of this curve is shown in Figure 2-6. By generating the SEC versus Water Production curve for any given EDR stack, insights regarding the

relationship between the SEC and water production become clear. In the first section of the curve, where the slope exceeds 1, more energy is used efficiently to increase water production. Where the slope flattens, significantly more energy is used for minor increases in water production rate. Furthermore, this curve shows insights on why energy storage could be useful to include in PV-EDR systems. With energy storage, the system could operate at a more efficient point on the Water Production Curve versus SEC, regardless of the power available. If there is excess energy, this could be stored for use later to operate either at night or during lower solar power periods. This operation strategy would maximize the water production for the total daily solar energy available as opposed to just maximizing desalination based on the instantaneous solar power available.

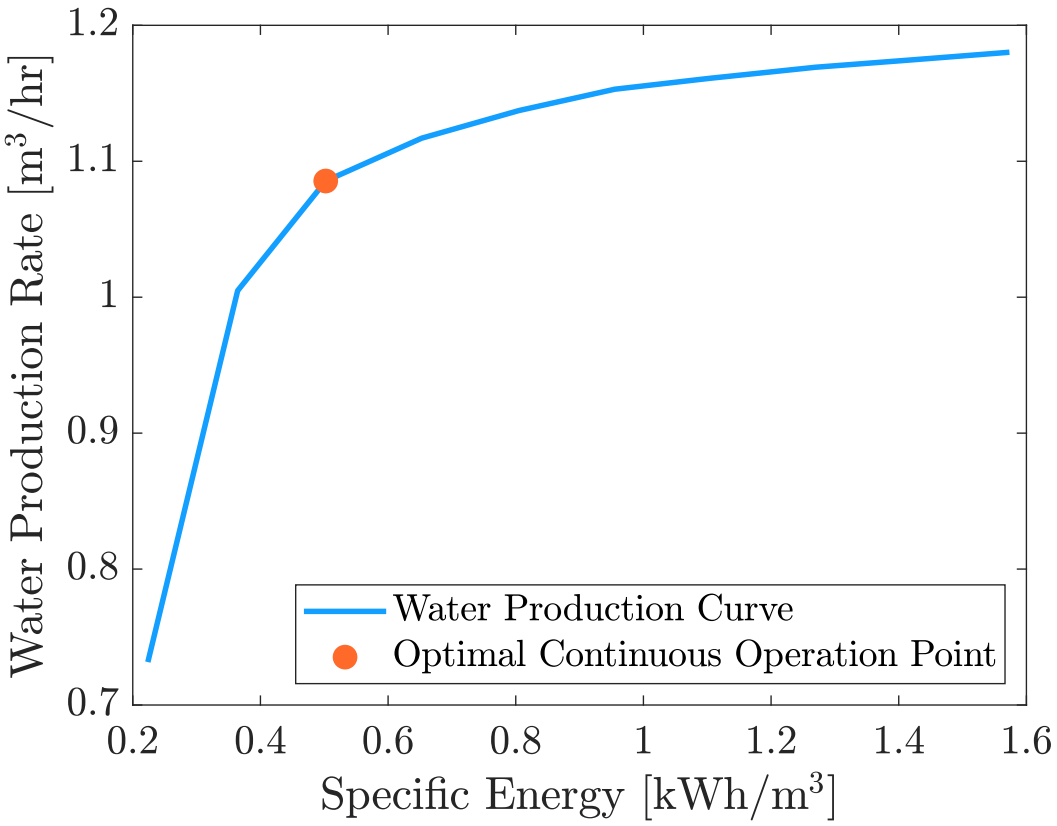


Figure 2-6: Simulated results of the water production rate versus SEC for a specific system design. The more water is produced, the lower the system efficiency. The red mark depicts the optimal operating point, which would be used to design and operate the system in traditional design theory. The system design used to generate this water production curve is a Suez V-20 ED stack with 100 cell pairs and a batch volume of 0.417 m^3.

Chapter 3

Control System Design

By leveraging the previous work in PV-EDR control theory described in Section 2, this section introduces an energy management controller. This controller utilizes the water production curve and weather forecasting to determine the optimal control action to reliably optimize water production that could, as discussed later, reduce costs.

3.1 Design Requirements of Control Theory

As described in Section 2, there is a need for a control strategy that can optimally determine the best utilization of solar power with the goal of reliably meeting the water demand every day at the lowest cost. Therefore, building off previous work, the focus of this additional module is to anticipate energy availability and plan water production accordingly. Moreover, the proposed control strategy needs to meet the following design requirements:

(1) *Flexible* – The control strategy should be functional on any PV-EDR system in any geography. Additionally, the control strategy must be flexible to any system configuration, minimizing need for tuning and calibration.

(2) *Low-cost implementation* – The control strategy must not have high capital cost due to expensive computing or sensing resources. The entire control architecture

should be executable on low-cost devices (e.g. a Raspberry Pi or Arduino), ensuring this strategy is applicable for emerging markets.

(3) *Optimal* – Control strategy should optimally determine the best energy management strategy between energy storage and energy use for desalination at any given time step, regardless of the technology used to store energy.

(4) *Predictive* – To be optimal, the energy management strategy requires foreknowledge of future available energy. The control algorithm should be able to anticipate energy availability at least a full day ahead, as this is the time frame of water production and consumption expectations.

We introduce a schematic of the novel control strategy in Figure 3-1. The high-level control system optimizes the energy flows of the system. To this end, a machine learning algorithm predicts future solar irradiance that is used to estimate the available solar power in the future. This module is the only part of the control system dependent on the geography, decoupling the remaining control strategy from where the system is used. Then, the power set for producing water is used as a reference for the low-level control system, which is based on the theory discussed in Section 2. The optimal flow rate and voltage that maximizes the desalination rate subject to the power set by the high-level controller are estimated. Finally, an event supervisor manages the machine operations for batch water production. In subsequent subsections, each piece of the control system is explained in detail and validated.

3.2 Power Forecasting Component of Control Strategy

Predictive capabilities are necessary to inform the energy management algorithm to find an optimal strategy [35]. Specifically, in this application the control system needs an estimate of the available future solar irradiance to mathematically determine the optimal power split between instant utilization and storage for future use.

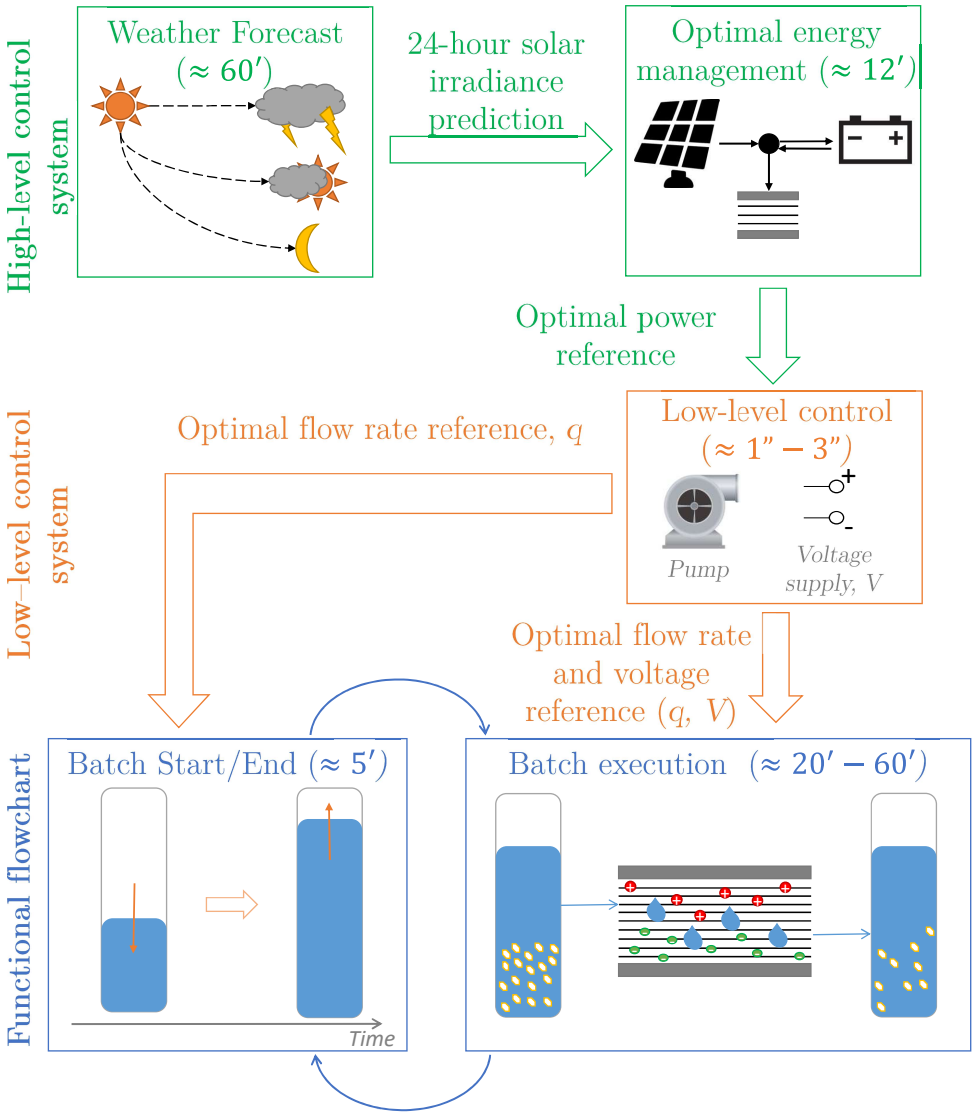


Figure 3-1: Flow chart of the controller's functions. Top (in green): High-level energy management system with the weather forecast predicted using machine learning and the predictive controller. Middle (orange): Low-level control system based on [37] and modifications discussed in Section 2. Bottom (blue): Machine operations, controlled by a discrete time event supervisor.

Fortunately, solar irradiance prediction is a well studied problem in the machine learning community. Indeed, there exists several data-driven algorithms to predict

weather data, differing in terms of sensors used, temporal resolution, and prediction horizon reviewed in Mellit et al. 2020 [40]. For successful implementation of any data-driven algorithm, high-resolution time-series weather data are needed. Therefore, access to appropriate weather datasets is a constraint for both the algorithm choice and implementation. The only publicly available weather data set for the geography examined in this study has an hourly temporal resolution. For this reason, it was possible to leverage only algorithms that could make hourly predictions. Based on the aforementioned design requirements, Mellit et al.'s algorithm, which is a Multi-Layer Perceptron Artificial Neural Network, proved to be the most suited prediction algorithm thanks to its high prediction performance and minimal sensing layout [39]. The algorithm is indeed capable of predicting hourly solar irradiance over a 24-hour prediction horizon using only a temperature and a pyranometer, enabling inexpensive predictions with an easy-to-install system setup.

As in any machine learning algorithm, historic data are used to train an algorithm to make future decisions autonomously. To this end, two phases are key for the algorithm success: learning and testing. During the former, data are leveraged to fit the algorithm in order to minimize the algorithm's prediction error, leaving the designer only the task to tune some calibration parameters (also known as hyperparameters) that can help improving the fit. This procedure can be automatically performed using K-fold cross-validation, a technique used to minimize the time and effort spent in manual calibration[1].

Once the learning phase is completed, testing is then performed to assess the performance of the algorithm against a set of data not used for learning. Testing results are evaluated through the coefficient of determination (R^2) and the Root Mean Squared Error (RMSE) between the predicted and measured data. The coefficient of determination indicates the proportion of the variance that is correctly predicted with the prediction algorithm, and is computed as [41]

$$R^2 = 1 - \frac{\sum_{i=1}^{n} (y_i - \hat{y}_i)^2}{\sum_{i=1}^{n} (y_i - \bar{y})^2}, \tag{3.1}$$

[1] In this application, K-fold cross-validation was used with K = 5.

in which n is the number of samples, y_i is i-th sample, $\hat{y}_i$ is the predicted value of the i-th sample, and $\bar{y}$ is the mean value of all the samples. Instead, RMSE is a statistical measure of the average prediction error and is calculated as

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}. \tag{3.2}$$

It is worth noting that the RMSE was also selected as loss function during the learning phase, as its squared nature weights outliers heavily, which is crucial for a successful training in sporadic low-irradiance events (*i.e.*, cloudy or rainy days) and are critical for the control strategy.

Both phases would require as much data as possible to better train the algorithm or test its performance, but a trade off is necessary. In this application, 75% of the data are used to train the algorithm and 25% are employed to test it's performance, following standard practice [26].

The algorithm was trained and tested against weather data from Medchal, India, using 20 years of data obtained through the free database built by Gelaro et al. 2017. Testing results showed an R^2 value of 0.914, indicating a successful training. In Mellit and Pavan 2010, the authors reported a coefficient of determination of 0.94 when tested against Italian weather, slightly higher than what was achieved for the Indian one. This drop in performance could be due to India's narrower temperature and irradiance swings; improvements could be possible using more training data.

The RMSE is instead analyzed based on the prediction hour, evaluating how accuracy changes further into future. As shown in Figure 3-2, the RMSE increases significantly every additional hour into the future, reaching a peak of 96 W/m^2 for the 15-th hour. Then, the error improves towards the end of the prediction window, likely due to the periodic nature of solar irradiance on a 24-hour scale.

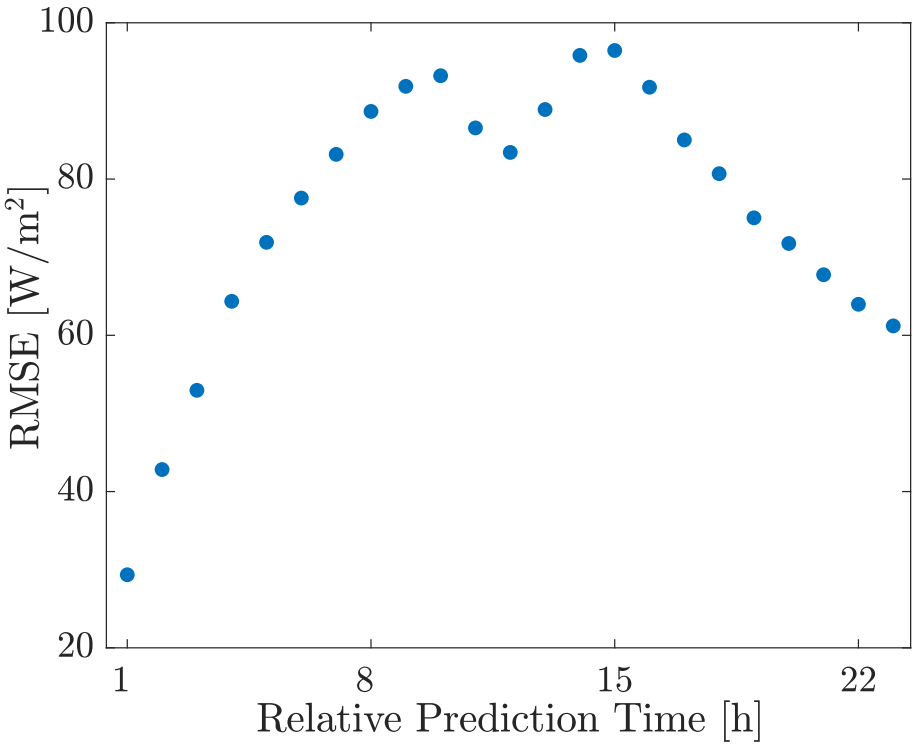


Figure 3-2: RMSE at different prediction hours. The solar irradiance prediction algorithm RMSE steadily increases as the prediction is further in the future. At hour 15, the RMSE is at a maximum. Accuracy improves towards the end of the prediction horizon, owing to the fact that irradiance is periodic on a 24-hour time scale.

3.3 High-level Optimal Energy Management Supervisor

A predictive controller was designed to optimize the energy flows of the systems using a control strategy known as Model Predictive Control (MPC) [23, 38]. MPC is a model-based control systems that, at every new measurement, numerically solves an optimization problem based on the current state of the system and a prediction of its future response, aiming to minimize a cost function. This operation is continuously repeated to mitigate modeling uncertainty and disturbances acting on the system, or to better react in case of unexpected events (*e.g.*, a system is temporarily turned off for maintenance). MPC is also a powerful control algorithm because it explicitly accounts for constraints of the system during the optimization, a feature particularly

important to account for actuator saturation or in applications with limited resources.

MPC is comprised of two main components: the nonlinear optimization problem and the numerical solver. In this application, the optimization problem is defined as follows

$$\min_{P_c,P_d,P_{ED}} J = \underbrace{\alpha\left(\overline{\text{SoC}} - \text{SoC}(T)\right)}_{J_b} + \underbrace{\beta\left(\overline{\nu} - \nu(T)\right)}_{J_\nu} \tag{3.3}$$

subject to

$$\begin{aligned} \dot{\text{SoC}}(t) &= \frac{1}{Q_b}\left(P_c(t) + P_d(t)\right) \\ \dot{\nu}(t) &= \frac{1}{\nu_t\left(1 + e^{-a\left(P_{ED}(t) - P^*_{ED}\right)}\right)} \sum_{j=0}^{4} b_j P_{ED}(t)^j \end{aligned} \tag{3.4a}$$

$$\begin{aligned} \frac{1}{\eta_b} P_c(t) \leq & \frac{1}{2}\left(P_{sol}(t) - P_{ED}(t)\right) + \\ & \frac{1}{2} P_{sol}(t) - P_{ED}(t) \\ \eta_b P_d(t) = & \frac{1}{2}\left(P_{sol}(t) - P_{ED}(t)\right) - \\ & \frac{1}{2} P_{sol}(t) - P_{ED}(t) \end{aligned} \tag{3.4b}$$

$$\begin{aligned} \underline{\text{SoC}} \leq & \text{SoC}(t) \leq \overline{\text{SoC}} \\ \underline{\nu} \leq & \nu(t) \leq \overline{\nu} \end{aligned} \tag{3.4c}$$

$$\begin{aligned} \underline{P}_c \leq & P_c(t) \leq \overline{P}_c \\ \underline{P}_d \leq & P_d(t) \leq \overline{P}_d \\ \underline{P}_{ED} \leq & P_{ED}(t) \leq \overline{P}_{ED}, \end{aligned} \tag{3.4d}$$

with $\text{SoC}(0) = 50\%$ and $\nu(0) = 0\%$ as initial condition.

The cost function in Eq. (3.3) is composed of two different terms: J_b, which expresses a cost related to the battery state of charge (SoC), and J_ν, which accounts for the final product tank level (ν). In this formulation, the controller aims to find the optimal selection of the power for the ED system (P_{ED}), for charging (P_c) and discharging (P_d) the battery in order to have the battery and the product tank full at

the end of the prediction horizon T, securing water production in case of subsequent low-irradiance days. Weights on the battery state of charge (α) and on the tank level (β) account for the battery aging cost and the daily cost of the plant, which make them dependent on the system design. Therefore, the controller is always aware of the actual cost of operating the system, and its optimal selection of the energy flows aims to find the least costly use of the system.

The optimization space is limited by a set of constraints, grouped in four main categories: constraints on the system dynamics, as listed in (3.4a); constraints on the power allocation, (3.4b); constraints on the state variables, (3.4c); and constraints on the control variables, (3.4d). The system dynamics equations account for the battery charging and discharging dynamics, and the water production. The battery state of charge dynamics is formulated as

$$\dot{\mathrm{SoC}}(t) = \frac{1}{Q_b}\left(P_c(t) + P_d(t)\right), \tag{3.5}$$

in which Q_b is the battery size. Instead, the dynamics of the water produced is modeled with a sigmoid σ and a fourth-order polynomial γ, as in

$$\dot{\nu}(t) = \underbrace{\frac{1}{\nu_t\left(1 + e^{-a\left(P_{ED}(t) - P^*_{ED}\right)}\right)}}_{\sigma} \underbrace{\sum_{j=0}^{4} b_j P_{ED}(t)^j}_{\gamma}, \tag{3.6}$$

in which ν_t is the nominal tank volume. The fourth order polynomial γ is used to fit the water production curve obtained for each configuration of the ED system, in which the curve is generated as discussed in Section 2.3; whereas the sigmoid σ is used to virtually turn off the stack when the power allocated for producing water is below a given threshold P^*_{ED}. Figure 3-3 shows the result of this modeling approach based on a water production curve taken as example. Thanks to this modeling strategy, the controller becomes aware of when the system needs to be turned off and how much water is produced when powered for any given power P_{ED}.

The equality and inequality constraints listed in (3.4b) are key to selecting the

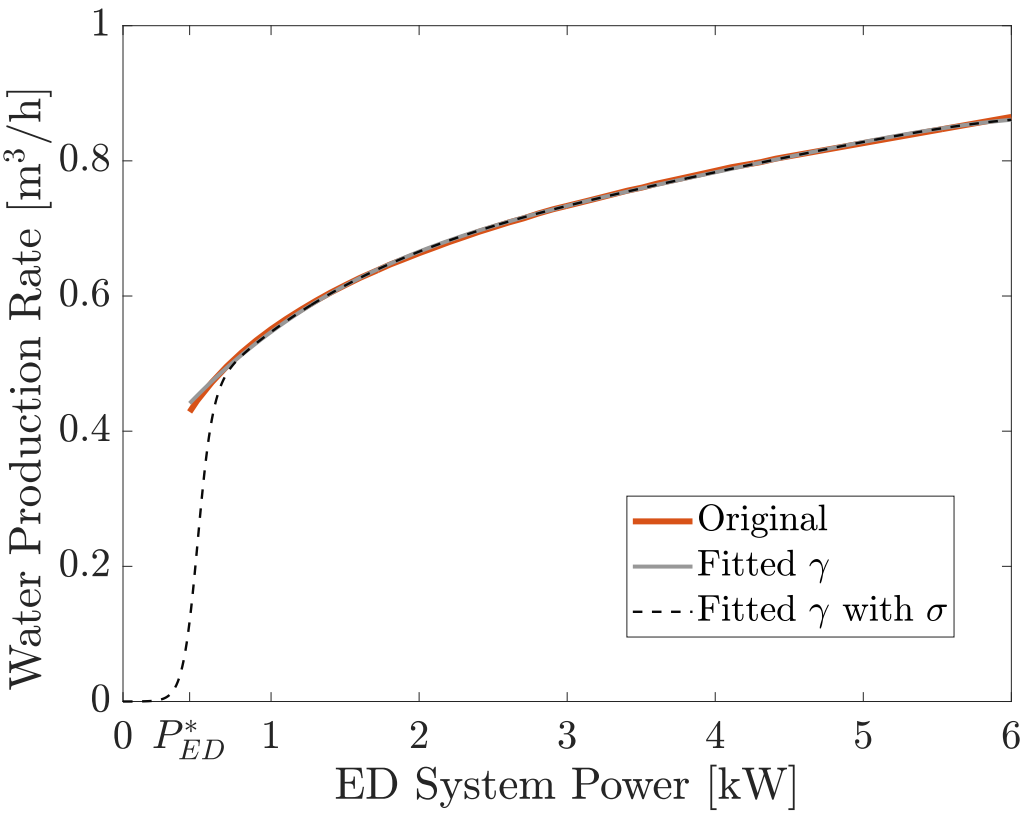


Figure 3-3: Water production rate versus ED system power. Starting from a water production curve, function γ (in gray) is fitted to have a the fourth-order polynomial function that relates power and product water. A sigmoid function σ (dashed) is then added to inform the controller that the system would be virtually powered off below a minimum power, denoted with P^*_{ED}.

charging and discharging power based on the difference between what produced by the PV panels (P_{sol}) and what used by the ED system for producing water, accounting for battery inefficiencies correctly (η_b). Specifically, the inequality equation prescribes the battery to be charged only when there is unused solar irradiance power, up to what available discounted by the battery losses. Therefore, the optimal controller may decide not charge the battery even if some solar power is available and remains unused (*e.g.*, the battery cannot be recharged when it is already full). Contrarily, the equality equation forces the battery to be discharged whenever the PV panels are producing less power than needed, exactly by what necessary.

The set of inequalities in (3.4c) and (3.4d) constrain the states and control variables within physical or technological limits. All the parameters used to defined these

constraints and, more broadly, in this optimal control problem formulation are listed in Table 3.1.

Parameter	Symbol	Value	Unit
State of charge cost	α	0.075	\$
Tank level cost	β	5.059	\$
Min ED system power	P^*_{ED}	0.45	kW
Sigmoid gradient	a	20	–
Min state of charge	$\underline{\text{SoC}}$	0	%
Max state of charge	$\overline{\text{SoC}}$	100	%
Min tank level	$\underline{\nu}$	0	%
Max tank level	$\overline{\nu}$	100	%
Min charging power	$\underline{P}_c$	0	kW
Max charging power	$\overline{P}_c$	4	kW
Min discharging power	$\underline{P}_d$	-4	kW
Max discharging power	$\overline{P}_d$	0	kW
Min ED system power	$\underline{P}_{ED}$	0	kW
Max ED system power	$\overline{P}_{ED}$	4.5	kW
Prediction horizon	T	24	h
Sampling time	T_s	12	min

Table 3.1: Parameters used in the optimization to define the cost function and the constraints. Any parameter not listed here depends on the system configuration, which is discussed in the case study section.

The second major component of an MPC is the numerical solver. To reduce the implementation costs, the entire nonlinear MPC has been designed and implemented in MATLAB Simulink using: 1) CasADi, an efficient, open-source tool for algorithmic differentiation [18]; and 2) GRAMPC, a nonlinear MPC framework based on an augmented Lagrangian formulation, which makes it suitable for efficient implementation on embedded hardware [24]. As a result, the implementation is numerically efficient, reducing the computational burden and enabling the use of an inexpensive, low-resource device such as a Raspberry Pi.

The MPC has been validated for this problem formulation in simulation. The goal of the validation process was to analyze how the system accommodates weather changes, proving the robustness of the control system. Figure 3-4 shows an overview of the controller response during a sequence of low- and high-irradiance days. The controller successfully meets the water demand each day, maximizing the water stored in

the tank (i.e., available to use), while limiting the use of the battery to minimize battery aging. As shown, in both high- and low-irradiance days, the controller manages the use of the battery to cope with weather variations. Specifically, in low-irradiance days, the available solar energy is combined with what is stored in the battery to produce enough water to meet the target water production. In high-irradiance days, the abundance of solar power is distributed between the ED system and the battery. Through the analysis of the time series in Figure 3-4, simulation results prove that the MPC manages the system efficiency and wisely. The system operates at a lower power than the total solar power available to operate at a lower SEC, producing water more efficiently over longer hours instead of operating the system inefficiently to increase water production and refill the tank more quickly.

3.4 Low-Level Voltage and Flow Rate Control

The combination of the solar irradiance prediction and the MPC optimize the use of the power for water production subject to the system design. Once the power allocated towards water production is determined, the low-level controller optimizes the voltage and flow rate according to the time-variant strategy described in Section 2, shown in Figure 2-3, and experimentally validated by [37]. This control strategy uses the power allocated by the high-level controller for finding the most optimal voltage and flow rate for the system to maximize the desalination rate. Maximizing the desalination rate maximizes the clean water production of the system. This control strategy will dynamically adjust the voltage every second, and the flow rate every three seconds, matching the speed of the empirically determined system dynamics.

3.5 Plant Level Operations

The last part of our control strategy manages plant level operations to produce water in batches. This piece of the control strategy executes a sequence of actions to automatically (1) fill the concentrate and diluate tanks with feed water, (2) move

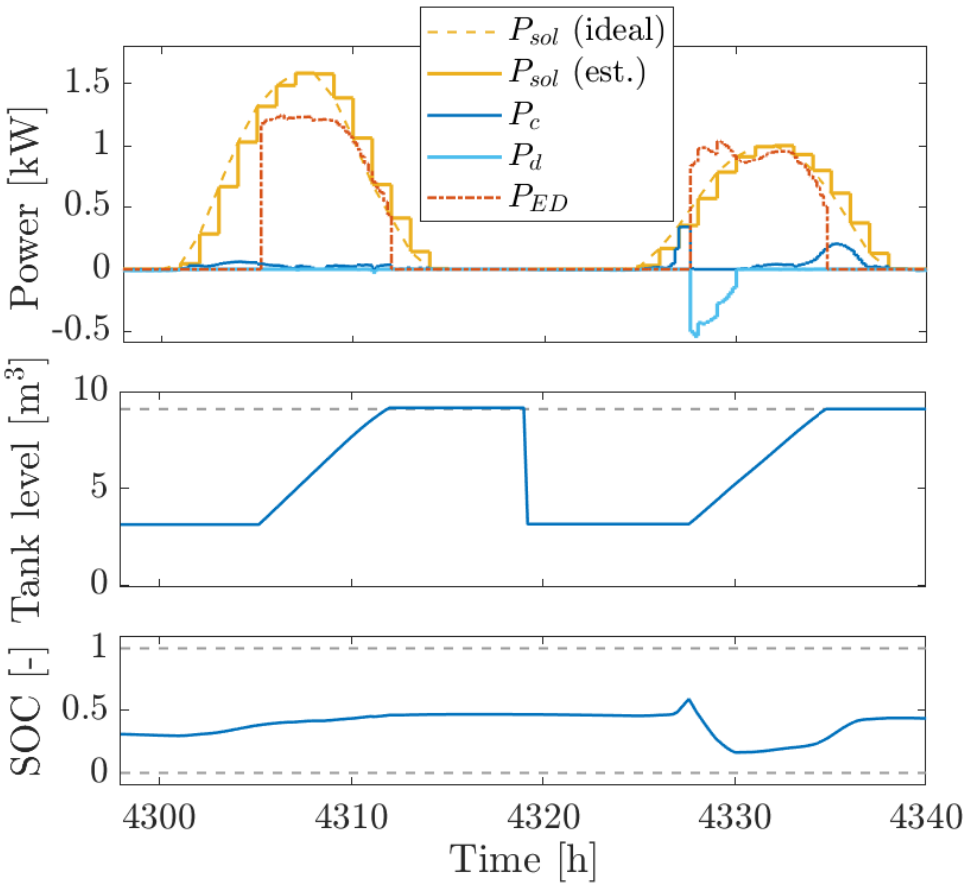


Figure 3-4: An overview of the response of the controller in a series of low- and high-irradiance days. Top plot: power use for ED system and battery pack, demonstrating that the controller optimally allocates power to produce water efficiently and minimize battery aging. In either day, the controller tries to operate at its most efficient regime (1 kW), reducing ED system power during the afternoon when the tank gets more and more filled. Middle plot: amount of water stored in the tank, reaching maximum capacity before water is consumed at night. Bottom plot: the analysis of SOC demonstrates that the energy management maintains the SOC in a narrow range, minimizing battery aging.

the desalinated diluate tank to a product storage tank, (3) and move the contents of the concentrate tank to an evaporation pond or other brine management. All of these operations are controlled by event interrupts through a series of system sensors, including tank level indicators, conductivity probes, and assorted valves. Whenever necessary, this instrumentation can interrupt the higher level controllers to stop water production and perform the key plant actions to switch between batches.

Chapter 4

PV-EDR System Design and Optimization

In Section 3 we have described a control strategy that includes an optimization to determine the best split of power between storage and desalination at every time step. In this section, we demonstrate how the control strategy can be used to design a PV-EDR system using our parametric theory of ED operations. Although this section looks at one specific design case, the methodology is applicable to any PV-EDR system, and could be used to design systems of different scales and applications. In contrast to the optimization executed by the MPC, which is focused on allocating power over time, the optimization introduced in this section is used to determine optimal system parameters. The system parameters (i.e. the solar panel area, battery size, number of cell pairs, batch size, and storage tank size) are selected through a simulation-based optimization with the objective of minimizing the overall system lifetime cost. Specifically, a year-long simulation is performed to assess the performance and cost of the controlled system over a year for any chosen system design. For each simulation, the expected cost and daily water demand are calculated and used to estimate the cost of the system over the entire lifetime. The architecture of the system design optimization and how it interfaces with the controller is shown graphically in Figure 4-1.

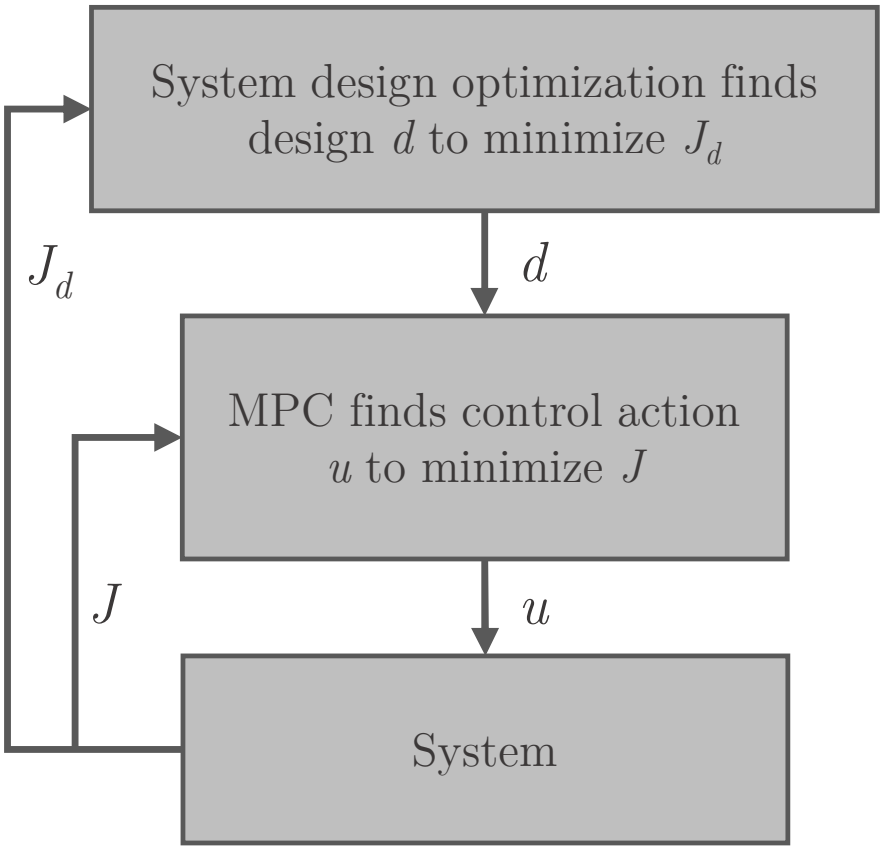


Figure 4-1: System design and MPC optimization loops. In the middle block, for a specific design configuration, the MPC finds the most optimal control action based on the response of the system (bottom block), computed through the objective function J as described in Section 3. In the top block, the system design optimization block tries to find the most optimal design that reduces cost, estimated simulating the controlled system over a year.

4.1 System Optimization Overview

The system optimization aims to determine the lowest cost system with a high reliability for a specific design case, assuming to control the system with the control theory described in Section 3. The system design optimization consists of four steps (Figure 4-2): (1) Define the optimization assumptions, which include the location, starting and ending salinity, water consumption profile, and operator schedule; (2) model the expected response of the system based on the system design, including the desalination performance, energy consumption, energy and water buffer utilization, and

water production; (3) perform a year-long simulation and calculate the performance metrics, which in this case are LCOW and reliability; (4) iterate until converged, which in this case is when the cost function does not change significantly between iterations. During step (1), the designer is actively involved in defining the initial conditions and system parameters. Instead, the optimization routine automatically performs steps (2)-(4).

Central for the success of the optimization is the definition of the cost function. In this application, we formulated the cost function for the system design optimization as the weighted sum of the LCOW and the difference between the simulated reliability $R_{\%}$ and its target value $\bar{R}_{\%}$, as in

$$J_d = \text{LCOW} + \rho \cdot \max(0, \bar{R}_{\%} - R_{\%}), \tag{4.1}$$

in which the maximum function is added for not rewarding reliability values higher than the desired one. Additionally, LCOW is defined as:

$$\text{LCOW} = \frac{P + \sum_{k=1}^{L} \frac{OM}{1+i^k}}{\sum_{k=1}^{L} W(k)}, \tag{4.2}$$

in which P is the plant capital cost; L is the system lifetime; i is the interest rate; OM is the annual operation and maintenance cost $W(k)$ is the amount of water produced at a specific year k, which can be obtained integrating γ as in (3.6) for a whole year. Moreover, reliability is defined as the percentage of time of a year in which production is met, which can be mathematically formulated as:

$$R_{\%} = \frac{100}{T_y} \int_0^{T_y} \frac{\max\left(0, \nu(t) - \nu_d(t)\right)}{\nu(t) - \nu_d(t)} \text{dt}, \tag{4.3}$$

where T_y represents the time duration of a year, whereas $\nu(t)$ is the tank level and $\nu_d(t)$ is the water demand at time t.

This reliability formulation is particularly suited for a time variant system. The reliability of the system is determined by whether or not the water demand at that time instant can be met. This formulation is then applicable to any water consumption

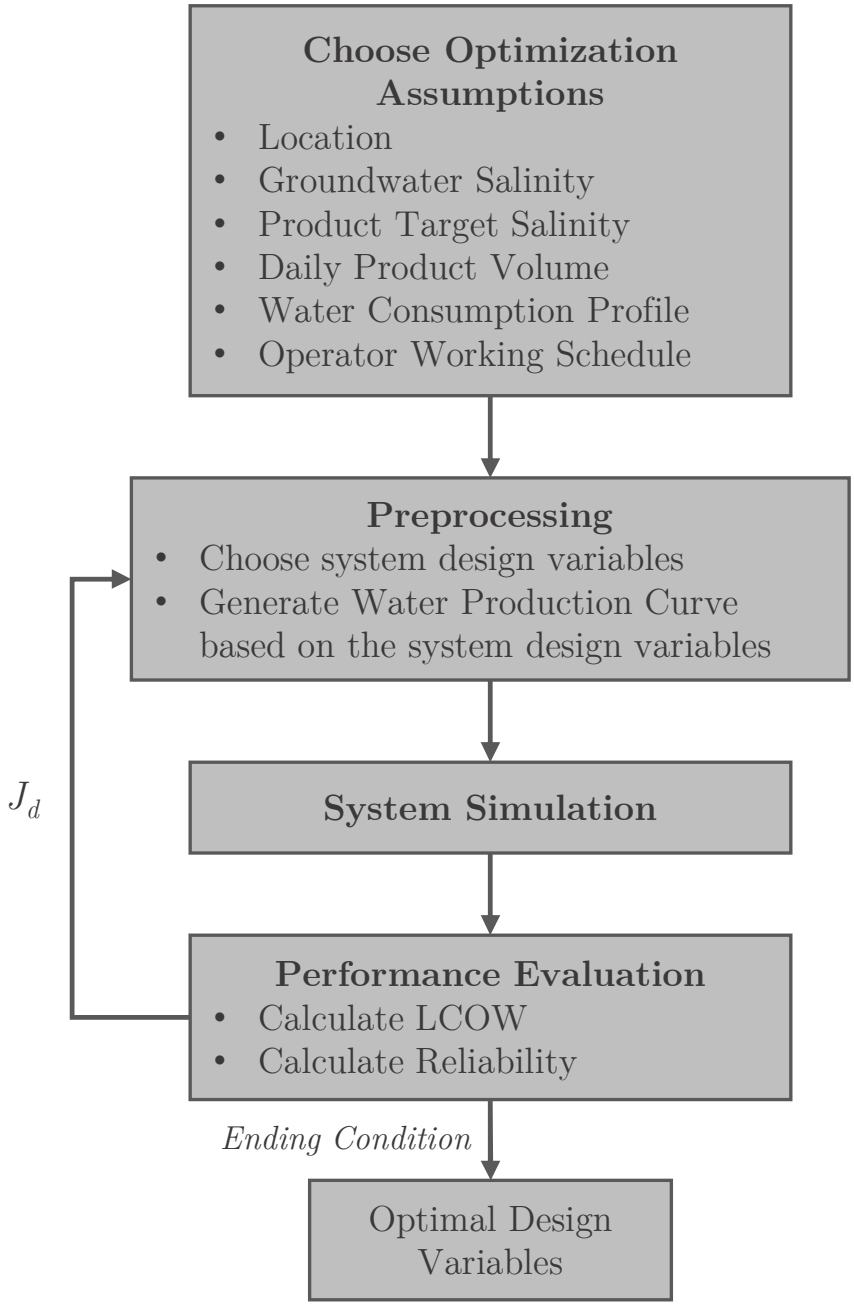


Figure 4-2: Flow chart of the system design optimization. The optimization assumptions are chosen and remain constant for the entire process. Then, design variables are changed, a water production curve is generated, and a year-long simulation is performed using the control theory described in Section 3. Based on the response of the system, costs and reliability are calculated and used to update the design variables. The optimization is repeated until the ending condition is met. When the optimization converges, the routine is terminated and the optimal design variables are outputted.

profile, where different volumes of water are demanded at different times throughout the day.

It is worth remarking on the role of the external penalty, ρ, used as the relative weighting between the importance of reliability and cost. This parameter is an empirically chosen and in this application is equal to 10. As a rule of thumb, the parameter can be tuned based on the designer experience or assuming a virtual cost that would be paid for any percentage of reliability not met (e.g. cost of ordering a truck of water, people discontent). However, if not meeting the water demand is of secondary importance, ρ should be reduced to enable finding systems at lower cost, as opposed to more reliable ones.

Lastly, the ending condition for the optimization algorithm checks when the cost function has converged, or when the objective function changes less than a small value chosen by the designer. Once converged, the output of the system optimization is the set of design variables for the optimal system.

The optimization algorithm used for this study is the pattern search method in the MATLAB 2020a Global Optimization Package. Pattern search uses an approximate gradient method to find an optimum for the cost function, searching for the best overall direction for the individual design variables [14].

4.2 Case Study - PV-EDR System for Rural Indian Village

To demonstrate this control theory in a real-world case study, we investigated and optimized a village-scale water system for Medchal, India. This location requires desalination of groundwater prior to human consumption due to the high salinity. As a point of comparison, typical drinking water salinities range from 100-500 ppm [44]. Traditionally, desalination in this area is primarily performed through on-grid RO systems. However, the work discussed in Wright and Winter V [49] showed that rural Indian communities are well suited for the implementation of PV-EDR technology due

to the wide availability of solar power and relatively low groundwater salinity. Design requirements for an Indian community have previously been determined by Wright and Winter V [49], and the key requirements for the present study are summarized in Table 4.1.

Design Requirement	Value	Unit
Feed Water Salinity	1200	uS/cm
Daily Water Volume	6	m^3
Plant Lifetime	10	years
System Recovery	70	%
Average Solar Power Efficiency	15	%
Maximum Pumping Efficiency	60	%
Power Supply Efficiency	92	%
Operator Working Schedule	8	h/day
Target System Reliability	100	%
Water Consumption Profile	Step at end of the day	-

Table 4.1: Design requirements for a PV-EDR System Optimization for a rural community in Medchal, India

An important assumption made is on the operator work hours. Specifically, it is assumed to have on site a full-time operator who works for 8 hours each day. As the system can only be operable when an operator is present, the system only operates for the same 8 hour period each day. For this reason, it was assumed that this 8 hour window corresponds with highest irradiance hours of a day, regardless of the normal working day in Medchal.

It was also assumed that water is consumed all at once, at night. Although this may not accurately reflect the actual water consumption profile, the step consumption represents a conservative design case, where water storage is needed for all the water produced during the day. In absence of knowledge of the typical water consumption profile of Medchal, this assumption is a necessarily conservative choice.

4.3 Optimization Parameters

The system optimization requirds the selection of several parameters, including: the free system design variables, the MPC time step, and the ending condition.

The design variables for this system optimization were: the solar panel size, the battery size, the number of cell pairs, the batch size, and the storage tank size. Most of the design variables are related to the storage of both water and power to maximize the system reliability. It would be possible to add additional flexibility to the design optimization by optimizing the ED stack size, or adding further capability to increase number of hydraulic or electrical stages of the ED system. However, for this case study, this additional flexibility is not necessary because the target daily water production is small enough to be easily achieved with a single Suez V-20 stack, the stack chosen in this investigation [1].

The time between MPC updates was empirically chosen to be 12 minutes for this application, trading off computational complexity with solution accuracy. As solar irradiance data for the target location are available only on a a hourly time scale, solar power was assumed constant for the entire hour. Therefore, updating the MPC more frequently and increasing the solution granularity do not change the performance of the controller significantly.

The optimization was terminated when the cost function varied between iterations by less than 0.001. This represents a very small change in either reliability (as this is scaled by ρ) or a small change in the LCOW.

4.4 Cost Inputs for System Optimization

Accurate cost information is necessary for a meaningful system design optimization. Cost can be dependent on geography, especially for common electrical components and shipping costs, making it critical to determine cost information that is both accurate and relevant. In this case study, we primarily obtained cost information from a database of wholesale costs provided to us through an Indian partner, Tata Projects,

to determine geographically relevant cost information for the fixed components of the EDR system [12]. These costs have been further validated by previous EDR pilots in India [31]. However, where bulk pricing from Tata Projects was not available for fixed components, we obtained prices for comparable equipment from manufacturers in the US. These costs are not wholesale. More details on the fixed cost components can be found in Appendix A.

To determine cost estimates for each of the design variables, we obtained multiple quotes for different sizes of each design variable component from several suppliers. Using the results of previous system PV-EDR optimizations, we estimated the relevant range of sizes for each design variable of this case study, and received quotes for components throughout the relevant range, ensuring the cost estimates are appropriate extrapolations for this system size [19]. To ensure numerical continuity in the design space, we used a linear fit function to determine a linear equation relating size of the component to cost. This technique enables the algorithm to search over a continuous design space because the determined design variable cost functions are continuous. More details on the fitting process and results can be found in Appendix A.

In addition to the fixed components and the design variables, there are components dependent on the size of design variables and the simulated system operation. For instance, the pumps, the associated pumping hardware, and the Direct Current (DC) power supply are all sized and costed after the year-long simulation. This approach gives the controller the freedom to determine the optimal power utilization regardless of the hardware limitations. Then, upon completion of the year-long simulation, the maximum pumping power and electrode power are used to determine the respective component sizes and cost. Linear fit functions were used to estimate the cost of the dependent system components, and are derived through the same process as the cost relationship functions for the independent design variables. Further details on the quotes of these components and their fit are provided in Appendix A.

There are also operational and maintenance costs of the PV-EDR system included in the LCOW calculation, including costs that are geographically specific as well. Op-

erational costs account for labor. In Medchal, the cost for operator time is $0.37/hour [37]. PV-EDR maintenance costs have been previously determined for Medchal, India, and are summarized in Appendix A [37]. Replacement costs for system components are determined from the manufacturer's lifetime expectation of each component and are only included if the component lifetime is expected to be less than the plant's one. Instead, replacement expectancy for the battery is estimated by computing the battery cycling experienced during the simulated year of operation and projecting it over the entire plant lifetime. The battery is therefore considered to be replaced only if the total estimated cycling is greater than the rated number of cycles. More details on the replacement cost for the battery and other hardware components can be found in the Appendix A.

The LCOW of an on-grid RO system is used as a point of comparison for the PV-EDR systems in this analysis. To determine the LCOW for an on-grid RO system, cost information for commercially available RO systems is needed. The authors of [37] previously determined the cost of an RO system that is currently available for sale and serving the Medchal, India desalination market. Moreover, the work in [37] included cost estimates of operation, maintenance, and component replacement. The same cost information are used in this study, as summarized in Appendix A.

4.5 Sensitivity Study

Sensitivity studies were performed to demonstrate the flexibility of this design tool as well as explore the optimal design's sensitivity to the design requirements and assumptions from Table 4.1. This sensitivity study demonstrates the flexibility of the design tool through varying several design assumptions. Varying these design assumptions does not change the control theory (from Section 2 or the methodology of the system optimizations. Moreover, the optimal results from the varied assumptions indicate the sensitivity to those design requirements.

A constraint of the geography of interest is the operator working schedule. Water systems can be seen as employment opportunities for village members, and thus an

operator would always be present when the system is operational [12]. Moreover, village members are less accustomed to remote monitoring of equipment in rural India, and thus would not be comfortable leaving equipment running without an operator present [12]. Therefore, it is crucial to evaluate the sensitivity of this design requirement to understand how the operator working schedule impacts the system design and, as a consequence, the LCOW. Starting from the traditional 8-hour work day, we also investigated a 12-hour work day with two operators shifting every 6 hours, and the growing practice of remote monitoring. In the latter case, an operator intervention is required only in case of failures and for routine maintenance. This allows a company to split one operator across many systems. From now on, we shall refer to this practice as flexible operation strategy.

Another key part of the validation of this design tool is to quantify the sensitivity of the optimal design to the solar irradiance prediction algorithm. Holding the rest of the design requirements constant, we run a system optimization with the known solar irradiance and predicted solar irradiance and compare the results. In both cases, the known and predicted solar irradiance data for the next 24 hours is supplied to the MPC. The known solar irradiance is the measured hourly solar irradiance for one year. The predicted solar irradiance data changes each hour as the prediction improves, and the MPC then changes its behavior based on the updated weather prediction. This comparison serves to validate how predicting weather data affects the design and whether the performance of the algorithm chosen increased costs significantly.

An additional assumption made in the optimization initialization regards the consumption profile of the water produced. In this sensitivity analysis, we also assume that water is consumed as a Gaussian function during the day because it seems feasible for an off-grid community to consume water during daylight hours. However, the authors of [37] assumed water is consumed all at once at night, as a step function at the very end of the day. We compare the two water consumption profiles to determine the importance of this assumption, and the cost or reliability implications of misunderstanding this assumption.

Chapter 5

Design Case Study Results

In this section, we examine the results of the case study defined in Section 4. We compare the results from this system optimization to the state of the art PV-EDR system designs previously determined in LeHenaff 2019. Then, we introduce the results from the sensitivity analysis. Specifically, we examine the way the operator working hours, the water consumption profile, and the prediction of solar irradiance effect the optimal result.

5.1 System Design Optimization Results

First, we aim to determine whether there are significant cost savings with the novel controls strategy, quantifying the benefit of the presented design method. As a point of comparison, LeHenaff [37] demonstrated an optimal direct drive system design that achieves a LCOW of \$2.06/m^3, with a reported reliability over 99%. To understand whether the control system and design methodology proposed in this paper improves upon prior work, we analyze the results in Figure 5-1a, which shows the most optimal LCOW for different reliability values in the explored design space. The lowest cost system that achieves 100% reliability has an LCOW of 1.91 \$/m^3. This reduction in LCOW from prior work translates to savings of over \$3200 over the lifetime of the plant. Moreover, Figure 5-1a demonstrates that there are many designs that have 100% reliability, but those designs could be significantly more expensive than

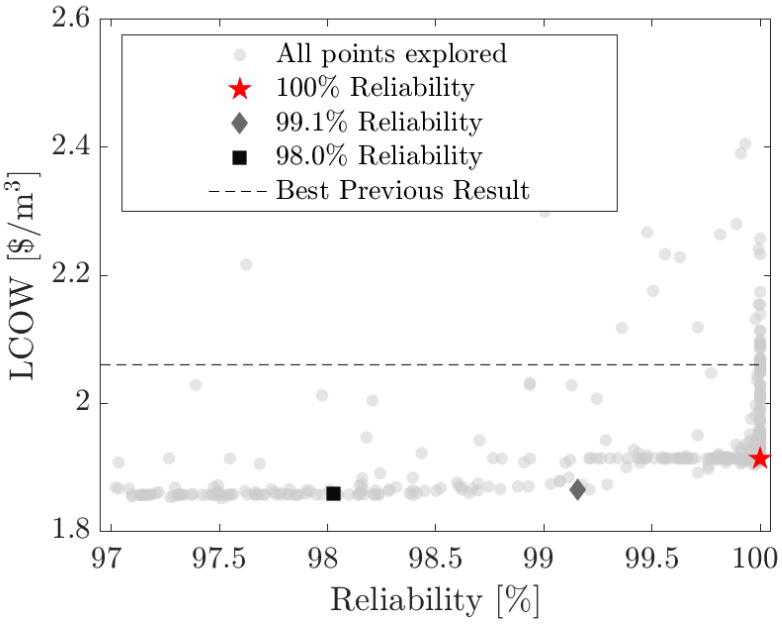
the optimal design achieved with our parametric optimization. Indeed, some designs range over nearly \$0.35 of LCOW, which equates to over \$7,500 over the lifetime of the plant. Therefore, a system level optimization is key to ensure finding the system configuration at lowest cost.

Another notable result from Figure 5-1a is the small difference in LCOW between the 100% reliable system and less reliable systems. For instance, the difference between the 100% reliable system and the 98% reliable system is \$0.06 per m^3. This result reinforces that there is not significant financial benefit to designing a system that has a reliability lower than 100%. Therefore, targeting 100% reliability for system designs remained the objective of the optimization for this case study.

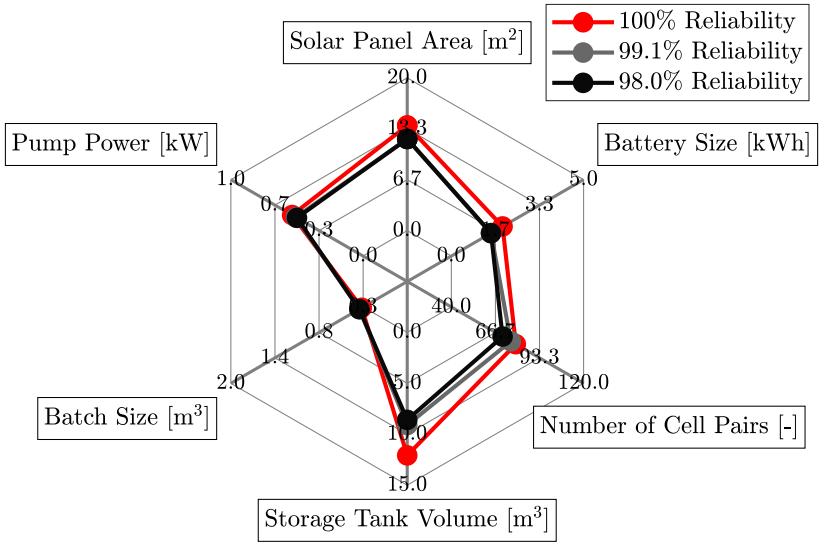
An overview of the design variables for the optimal systems at three different reliabilities is shown in Figure 5-1b. Notably, all three system designs are very similar: they have nearly identical batch sizes, whereas minor differences are found for all the other variables, which are generally smaller than than the 100% reliable design. The design variables which changed most substantially to achieve 100% reliability are the battery size and the storage tank size. The significant increase in additional storage capacity is necessary to overcome multiple days of low irradiance. The design variables that varied the least are the pump size and the number of cell pairs. The system converges on similar results for the cell pairs and the pump size regardless of the reliability, indicating that these are cost-optimal and do not impact reliability as significantly as the design variables related to water and energy storage.

5.2 Sensitivity of Operator Cost

To understand the impact of the assumption around the operator working day, we performed additional optimizations comparing the optimal designs for the operator working a 12-hour day, an 8-hour day, and the flexible operation strategy. All other assumptions for these additional optimizations match the design requirements in Table 4.1, and target reliability $\bar{R}_{\%}$ is set to 100%. For the optimal system design in each case, we analyzed the LCOW composition (Figure 5-2). Unsurprisingly, the primary



(a) Optimization's Pareto front.



(b) Design variables for systems 100, 99.1, and 98.0% reliable.

Figure 5-1: Analysis of the optimization results. a) LCOW versus Reliability plot for optimization results above 97% reliability. LCOW ranges within \$0.06 for the top 3% reliable systems. b) Diagram of the design variables for different reliability values. As shown, design configurations differ mainly in terms of storage, both tank and battery, and cell pairs.

difference in LCOW across the three PV-EDR design cases were the operator cost.

Although the LCOW without the operator cost is comparable for the three PV-EDR design cases, there are some differences in the optimal design variables. For instance, the 8-hour operator design has larger solar panels than the 12-hour or flexible operator design, as the system must capture all the usable energy in a shorter period of time. Thus, a larger solar panel area is necessary to meet the daily water production in fewer operation hours. Furthermore, the ED stack for the flexible operator design has fewer cell pairs (70) than the 12-hour or 8-hour operator designs (79), as the flexible operator system can operate over a longer period of time. For this reason, the ED stack can be smaller and energy can be stored in the battery during the day and use it at night, achieving the same volume of water produced as the time-constrained systems with more membrane area.

Additionally, the flexible operator design has the smallest storage tank of the three designs (9 m^3), with the 12-hour operator and 8-hour operator designs having storage tanks 33% larger. The smaller storage tank in the optimal design indicates that the flexible operator design can more easily reach the target volume each day, while the other design cases need a larger storage tank to achieve 100% reliability.

Finally, the pump size for the flexible operator design is smaller (0.45 kW) than the 8-hour and 12-hour designs (0.54 kW and 0.48 kW respectively). This smaller pump can again be explained by the ability of the flexible system to reshape more power than the other two time-constrained system designs. Therefore, the pump can be smaller due to the ability of the system to run more batches at a lower power consumption over a longer period of time. Contrarily, the 12- and 8-hour designs need progressively larger pumps to deal with the constraint on operation time. The battery size and batch size changed little in the different optimal results, indicating that these vary less with the constraint on operator time.

Figure 5-2 also compares these three optimal designs to the market leader in small scale desalination systems, on-grid RO [49]. As shown, the *fixed cost* is the most dominant for the RO system. The *fixed cost* is the combination of controls, hardware, and brine management needed for both the PV-EDR systems and the RO

systems. Brine management is included explicitly as a cost even though brine is not necessarily treated properly in all geographies, including in India. However, due to the impact improper brine management can have on the salinity of the groundwater, it is increasingly apparent that the need for desalination includes the need to properly treat and dispose of brine. Particularly significant are brine management costs, as RO systems have a significantly lower recovery than ED – 40% for RO as compared to the 70% assumed for PV-EDR [49]. Another major difference between the PV-EDR system LCOWs and the RO LCOW is the recurring energy cost for RO systems. This is an operating expense for RO systems that is not needed for a PV-powered desalination system.

This analysis also demonstrates that this system design optimization method achieves a PV-EDR system comparable in lifetime cost to the market leading technology. The optimized PV-EDR system with an operator for 8 hours each day is at price parity with the same sized RO system. Furthermore, if the design context allows for a more flexible operator approach, the optimal system design would have a significantly reduced LCOW as compared to standard practices.

5.2.1 Known versus Predicted Weather Design Results

To test the effectiveness of the weather prediction algorithm and its impact on the system design, the optimization was performed using both the known weather and the predicted weather to directly compare the differences in optimal outcomes. For this optimization, the operator schedule awasdjusted to the flexible operation because the flexible operation schedule is most susceptible to weather variation, as the system has to decouple power generation and consumption entirely. For this reason, a flexible operator schedule enables the system to operate longer hours if necessary with respect to a manually operated system, in which the system is operated the same 8 or 12 hours every day. The results from this analysis determine the efficacy of using the chosen Machine Learning algorithm (from Section 3.2) in the PV-EDR design process.

The LCOW for the predicted weather is \$1.4992/m^3, \$0.036/m^3 more than its counterpart obtained with known weather, which increases costs by only \$790 over

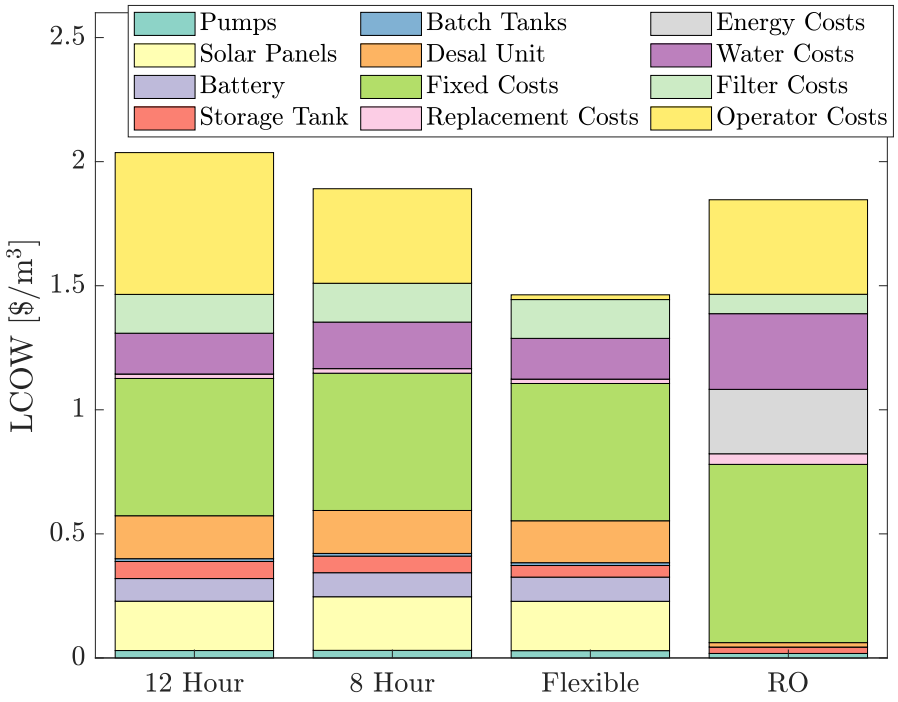


Figure 5-2: LCOW of optimal system designs for different operator conditions compared to on-grid RO systems, the market leader. The optimized PV-EDR system is at price parity with the on-grid RO system with the same operator constraint. A PV-EDR system with a flexible operator strategy has a significantly lower LCOW.

the entire lifetime of the system. To better contextualize this difference, Figure 5-3 shows the comparison of the design variables for the optimal design for the predicted and known weather. When weather is predicted, the optimal design has smaller solar panels, a larger storage tank, and more cell pairs than the known weather optimal design. These differences are likely due to the overestimated future power with respect to using real data. By predicting more solar irradiance than actually available, the predicted weather design has smaller solar panels. Moreover, excess storage capacity is needed to compensate the inaccurate prediction. Additionally, if the system predicts it will get more power, the system optimization selected a larger number of cell pairs such that the system operates more efficiently at the average operating power.

To assess the significance of these design differences, we estimated a reliability of 99.7% over a year-long simulation using the system design based on predicted weather with the known weather data set. Although the system design using predicted weather

does not perform perfectly on the known weather, the reliability of any built system will not likely be 100% due to dynamics not captured by the system model used for this optimization. Therefore, it can be considered acceptable losing less than 0.3% of reliability in this design study due to inaccurate weather prediction.

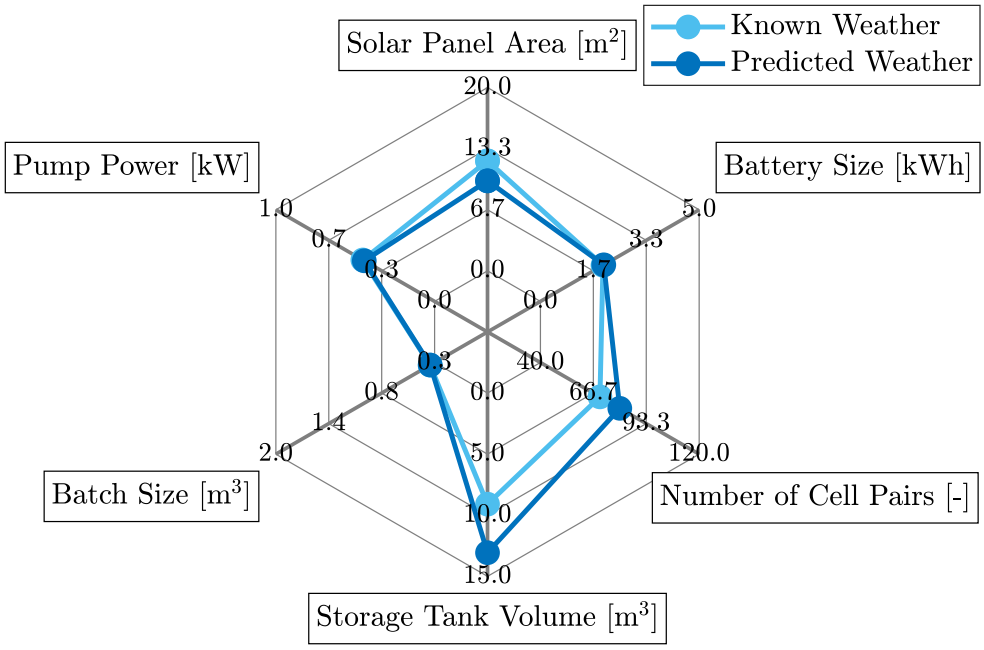


Figure 5-3: Diagram showing the differences in design variables between the known weather and predict weather optimal results.

5.3 Results Related to Water Consumption Profiles

In previous studies, the water consumption profile was modeled with a step function at midnight, where the daily water demand is removed from the storage tank all at once [37]. However, this assumption is unlikely to represent how water is consumed adequately.

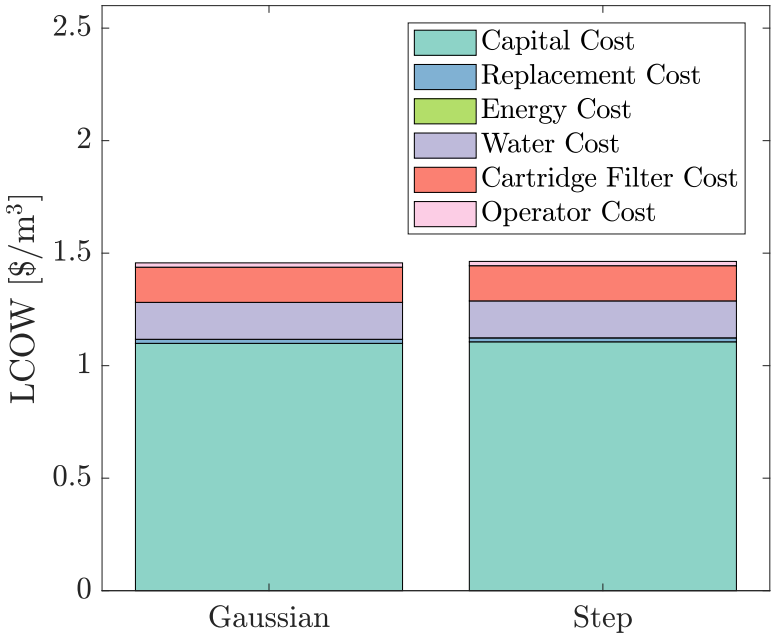
For this analysis, we used a flexible operator scheme to fully exploit the performance of the controlled system with no temporal constraints. The results of this analysis for the two optimally designed systems and their LCOWs are shown in Figure

5-4.

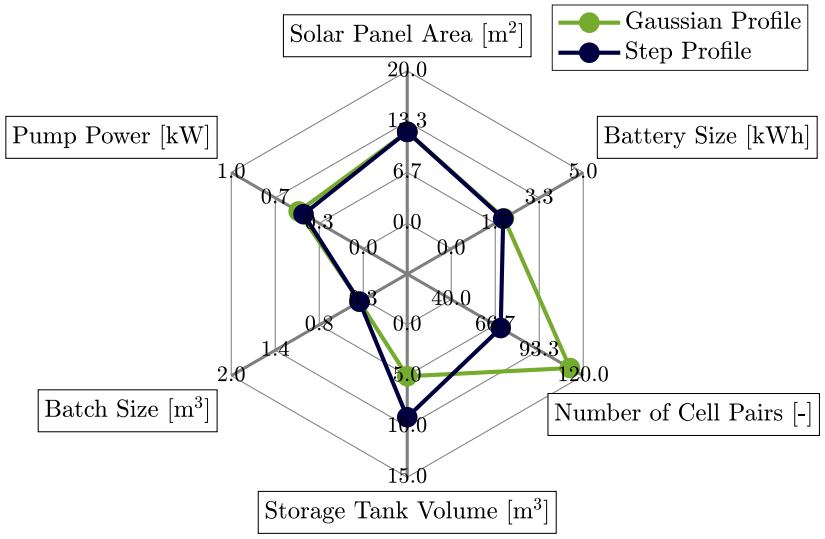
As illustrated in Figure 5-4a, the LCOW of the optimal system designs obtained using the different water consumption profiles are very close (difference of \$0.0065). Not surprisingly, the LCOW estimated using the Gaussian profile is slightly less than its counterpart for the step profile: Due to the large amount of water consumed instantaneously, the design for the step profile requires a larger storage tank that can hold more product water (Figure 5-4b). Conversely, in the continuous (Gaussian) consumption case, water is continuously removed from the storage tank, resulting in a smaller product tank that can store only the water needed during the hours in which the system is powered off.

Moreover, the two optimal systems differ significantly in the number of cell pairs, with the fewer pairs required for the step consumption case (70) than the Gaussian one (112). To explain this difference, it is paramount to consider how the energy management system operates: In fact, the controller aims to have the tank full at the end of the prediction horizon. Therefore, when water is continuously consumed, the controller tries to match production rate to the consumption one. Contrarily, in the case of the step consumption, the controller does not try to match a consumption rate because it happens all at once. Thus, production is stretched out over a longer period of time, requiring a significantly smaller stack.

Another way to understand the importance of knowing the water consumption pattern is to use a system designed for one water consumption pattern and simulate another water consumption pattern with that design, and determine the reliability. If the optimal system design for the Gaussian water consumption is used in a context where the step water consumption profile is actually executed, the system reliability (assuming reliability on a daily time scale) is zero. The tank cannot store the total quantity of water needed each day as it is only 5 cubic meters. Therefore, the system only ever produces 5 cubic meters of water to fill the tank. Alternatively, if the optimal design with an assumed step water consumption is operated in a context where the water consumption profile is Gaussian, the reliability of the system is 99.9% reliable. The step water consumption design is more robust to the Gaussian water consumption



(a) LCOW for different water consumption profiles.



(b) Design variables for different water consumption profiles.

Figure 5-4: Comparison of the system design for a step and a Gaussian water consumption profile. The analysis of LCOW (a) shows minimal differences ($0.0065), meaning that LCOW is not sensitive with respect to the different consumption dynamics. However, the consumption profiles influence the system design (b), requiring a larger storage tank for a step profile, whereas a larger number of cell pairs is necessary when water is consumed continuously.

profile.

These system design differences highlight the importance of understanding the local context when performing system optimizations. By knowing the water consumption profile, in either case, we can find a system with a nearly equal LCOW and high reliability. When this information is not available, designers would have to make an assumption about the water consumption profile. As these results indicate, different assumptions lead to major design differences, which could reflect in a suboptimal system design in case the assumption is inaccurate.

Chapter 6

Discussion

This study set out to determine how to best design and control a low-cost, high-reliability PV-EDR system. Using the prior work from [37], we sought to utilize the insight that using more energy increases the SEC of water production, and therefore there is a more efficient method of operation through minimal energy buffers to maintain a higher reliability than with a direct-drive system as well as minimize buffer costs. In order to operate at a lower SEC, it is paramount that the system has some foreknowledge of the power available over the course of the day to plan power utilization. A hierarchical control algorithm was designed, combining together several modules. A machine learning algorithm was employed to inform the system's power budget for utilization between the system's competing interests: charging the battery and filling up the tank. Then, an energy management control strategy uses optimal control theory to make decisions based on both the power predicted and the current state of the plant, aiming to achieve a desired objective: having the battery and the tank full at the end of the predicted horizon.

Leveraging this novel control theory, we optimized the system architecture, simulating system operations over a year to calculate the reliability and system cost. The system optimization converges on the lowest-cost, highest-reliability system design, subject to the inputs including geography and cost data. The optimal system design for the case study in rural India using this control theory has an LCOW reduction of \$0.15 with 100% reliability as compared to the optimal system design of [37].

This control strategy is applicable beyond optimally designing and controlling PV-EDR systems. The control strategy demonstrated in this paper can handle time-varying power inputs, and make the most optimal use of power over time to achieve predetermined goals. For instance, if power was varying in cost throughout the day in a predetermined way, this strategy could be modified to minimize power cost for the system as a whole subject to the grid power costing schedule. Indeed, an optimal, time-variant control strategy can not only help optimize power utilization from variable power loads but also improve performance of systems powered with renewable energy or variable-cost power sources. This tool can be particularly helpful in alleviating the concern of adequate reliability for using renewable energy as a primary energy source. This concern is amplified when the power is used for basic human needs like water production.

Moreover, the design tool using this novel controls system for PV-EDR systems can be generalized for any design case. Although we have only demonstrated this tool for one particular case study, the power of this design tool is within the inherent and proven flexibility to work with different geographies, water salinities, daily water production rates, water consumption profiles, operator working norms, system size scales, opening potentially new research directions. Moreover, all design requirements listed in the 4.1 are inputs to this design tool, and could be modified for another design case.

Through the sensitivity analysis performed, it is clear how critical it is to understand each design requirement for determining the optimal system design. In the case of water consumption profiles, the system design optimization tool was able to find systems for both the step water consumption profile and the continuous water consumption profile that were comparable in LCOW. However, the system designs were significantly different. Therefore, knowing the water consumption patterns is an important input for using this system optimization tool to find a cost optimal system design. Additionally, the operator cost is the single largest recurring cost of the system. Understanding this cost for the design case, and any limitations in terms of hours that operator is allowed to work, is critical for determining the LCOW of

the system. In other cases, this operator cost will be even more critical to understand and account for in any system design as India has relatively low labor costs compared to other geographies like the United States or European Union Nations.

This work demonstrates that the choice of the control system, and the assumptions made in how that system will operate, affects the system design and performance. This MPC control strategy has changed the optimal system design as compared to previous studies performed with similar design requirements. In Bain et al. [19] where a rules-based control strategy was implemented on a direct-drive system, the optimal system design had oversized solar panels to account for uncertainty. However, this control strategy does not oversize solar panels due to having the foreknowledge of the upcoming solar irradiance, using the battery storage to compensate for the prediction uncertainty. Moreover, the choices of water consumption profile and operator cost have implications on the LCOW and system design. We have demonstrated these differences through the sensitivity analysis.

This work also highlights the importance of the definition of reliability. In this case study, reliability corresponds to meeting the water demand at every time instant. Historically, increasing reliability meant producing more water; in truth, it should mean producing water at the right moment, or when water is needed. Indeed, the optimal designs obtained using this design tool produce less water than LeHenaff, but achieve a higher reliability. This efficient production helps to reduce brine management costs further as compared to previous works.

This design tool can be modified to change the objective to maximize water production or minimize SEC, depending on the interest of the system designer. Moreover, the control system objective could even change during the plant lifetime, subject to the same system design. Defining the objective for the control system is a designer's choice, and can have significant effects on the system's design and operation.

These results show the promise of PV-EDR technology, achieving price parity with RO. However, there are further cost savings that could be anticipated, especially as renewable energy becomes more affordable. In Bloomberg New Energy Finance: New Market Outlook, it is anticipated that solar panel prices will reduce by 34% by 2030

[33]. This trend would result in a $\$0.07/m^3$ decrease in LCOW for our optimal solution from section 5.1. From similar market reports, battery costs are anticipated to drop by 50% by 2023, especially due to the prevalence of Electrical Vehicles increasing the research and development of more cost effective batteries. An aggressive cut on battery costs would reduce the LCOW of the optimal design by $\$0.07/m^3$ [33]. As investments in improving the affordability of renewable energy continue, it is likely that PV-EDR could become the clearly more cost-effective option compared to on-grid RO.

6.1 Limitations of This Work

As pointed out in Section 5.2.1, access to more precise solar irradiance data may improve the performance of the control system and reduce cost. With only hourly weather data precision, we cannot use many more sophisticated solar irradiance prediction algorithms to improve our weather prediction accuracy. Moreover, using data predicted or collected with large step sizes may also reduce the performance of the system as it is impossible to optimally estimate the power actually available within the step size. As a consequence, updating the energy management more frequently than what currently done (i.e., 12 minutes) would not provide any benefit without a finer prediction.

This work is highly sensitive to cost analysis. Due to the market for stock materials and components, bulk pricing is typically very favorable for companies that can leverage scale economies. The cost reported in Appendix A for RO represents a system that is produced in the thousands of units, and thus the company is able to obtain favorable bulk pricing for all components. However, the costs summarized in Appendix A are not all wholesale prices and do not necessarily reflect the cost of producing a PV-EDR system at scale. Therefore, we could be overestimating costs for some system components. However, the design method presented herein is applicable to any variation in system variables.

We aim to experimentally validate this controller and system design optimization

tool by building the optimal controller and the optimal system design for a chosen geography, and test the system at that chosen geography. We further hope to explore the effects of a varying water consumption profile through our experimental validation to understand how robust the system design is to variable groundwater composition. By validating this tool through a pilot-scale experimental setup, we can evaluate any uncertainty in this design methodology. We hope to prove that this strategy is optimal for designing PV-EDR systems.

Chapter 7

Conclusions

In this paper, we introduced a novel control strategy that leveraged predictive capabilities to optimally operate PV-EDR systems, trading off water production and SEC in order to achieve a lower cost system through the minimization of system buffers. To demonstrate the utility and value of this new control theory, we performed system design optimizations using this control strategy. Through these system optimizations, we demonstrated that PV-EDR systems sized using this theory represent a significant decrease in LCOW at maximum reliability as compared to previous system designs. Through a sensitivity study, we also demonstrated both the flexibility of this design tool and the importance of design assumptions, as they greatly effect the system design and reliability. Furthermore, we compared an optimized PV-EDR system to the market leader, on-grid RO, and show that the two systems are equal in LCOW over their system lifetimes. We further elucidated the benefits of this controls strategy beyond that of PV-EDR systems, as this is broadly a strategy for optimal energy management for varying power loads.

This theory uses model-based approaches to lower the cost of renewable energy desalination systems. As groundwater is continuing to grow more saline, especially in areas where grid power is not reliable, low-cost, high-reliability renewable energy desalination solutions are needed to meet the drinking water demand. The design tool presented in this work enables designers to optimize system designs based on providing low-cost, high-reliability PV-EDR systems. Furthermore, this design tool

is generalizable to any design context, making it relevant to system designers of any size, scale, or application.

Appendix A

Detailed Cost Information for LCOW Calculation

Table A.1 reports the detailed cost information for the fixed components used to design a PV-ED system. These components include the hardware and controls needed to run a PV-ED system in batch mode, and are not changing with the size of the system. This is included as part of the capital cost in the calculation of the LCOW for all design cases explored in this study.

Table A.2 contains the cost information for the operation costs for PV-EDR systems. This cost is calculated annually, and then discounted based on interest over the course of the plant lifetime to estimate the operation cost of the plant over the lifetime.

Instead, Table A.3 reports the lifetime of each component, which is used to calculate the number of replacement parts for the designed system. It is worth to point out that because the battery ages, among many factors, based on its actual use, which is estimated in terms of cycles [20]. Thus, to estimate the number of replacements, battery aging is computed over one year is projected for the entire lifetime of the system, as in

$$\vartheta = \frac{L}{2Q_b} \int_0^{T_y} P_c(t) + P_d(t) dt, \tag{A.1}$$

in which L is the lifetime of the system (expressed in years), whereas P_c and P_d are

Component	Quantity	Cost[$]		Ref.
		Unit	Total	
Click PLC Controller	1	198	198	[2]
Click PLC Power Supply	1	41.50	41.50	[2]
Click PLC Analog Input I/O Block	11	95	1045	[2]
Click PLC Digital Input I/O Block	1	34.50	34.50	[2]
Click PLC Analog Output I/O Block	1	129	129	[2]
Raspberry Pi	1	35	35	[7]
Enclosure and Associated Hardware	1	775	775	[2]
Pressure Transmitter	6	75	451	[12]
Flow Transmitter	2	95.90	192	[12]
Conductivity Transmitter	4	42	168	[12]
Float Level Switch	6	5	30	[12]
Current Transducer	4	123	492	[2]
Voltage Transducer	4	186	744	[2]
Solar Pyranometer	1	249.99	249.99	[8]
Automated Valves	18	76	1368	[12]
Frame and Mounting	1	850	850	[12]
Cables and Wiring	1	550	550	[12]
Piping	1	200	200	[12]
Circuit Breakers	1	64.66	64.66	[12]
Acid Pump	1	77	77	[12]
Acid Tank	1	160	160	[5]
Brine Evaporation Pond[1]	2.57 m^3	1651	4245.42	[21]
Total	–	–	**8548.07**	–

Table A.1: Cost information for fixed system components.

the battery charging and discharging power respectively. Therefore, the LCOW does not account for a predefined number of replacements of the battery, but it is subject to an estimate based on the actual use.

Table A.4 shows the variable cost equations for both the independent design variables and the dependent design variables. The only dependent design variable not shown is the DC power supply. The DC power supply cost is determined by the maximum power applied to the stack over the simulation, following the cost model reported in Table A.5.

Table A.6 lists all the capital costs for the RO system; whereas, the operation costs of RO are the same as the PV-EDR system, as previously presented in Table A.2. The maintenance schedule for RO equipment is summarized in Table A.7. The

Operation	Volume	Unit Cost[$]	Ref.
Water Cost	$\frac{1}{RR}$	0.14/m^3	[12]
Acid Cost	$\frac{1-RR}{RR}$	0.03/m^3	[12]
Evaporation Pond Maintenance	$\frac{1-RR}{RR}$	8.26/m^3	[21]
Filter Cost	–	36.96/mo	[12]
Operator Cost	–	0.37/h	[12]
Energy Cost	–	0.168/h	[12]

Table A.2: Cost information for operating costs for both PV-EDR systems and RO systems. The recovery ratio used in this analysis is 0.7 for PV-EDR. The recovery ratio used for RO systems is 0.4.

Component	Lifetime	Ref.
Pump	5 [years]	[4]
ED Membranes	10 [years]	[1]
Battery	6000 [cycles]	[3]
Storage Tanks	25 [years]	[5]
Solar Panels	10 [years]	[9]
Fixed Hardware	10 [years]	[2, 12, 8]

Table A.3: Components lifetime of an PV-EDR system.

LCOW for the RO system is calculated using the same LCOW equation 4.2. Since RO is assumed on grid, the reliability of the RO system is assumed to be 100%.

Optimization	Design Variable	Unit	Variable	Cost Model [$]	R^2	Range	Ref.
Direct	Solar Panels	m^2	s	$214.43 + s \cdot 1781.5$	0.997	5 - 50	[10]
	Battery Size	kWh	b	$1222.1 + b \cdot 464.16$	0.9814	0.5 - 45	[10]
	Batch Tanks	m^3	t	$73.148 + t \cdot 198.88$	0.9849	0.2 - 2	[5]
	Storage Tank	m^3	l	$-320.48 + l \cdot 147.82$	0.9749	2 - 20	[5]
	Cell Pairs	–	N	$20 \cdot N$	–	1 - 150	[11] (estimated)
Indirect	Pump	kW	p	$191.32 + p \cdot 133.26$	0.9861	1 - 3	[4]
	VFD	kW	v	$113.29 + v \cdot 59.862$	0.9966	0.5 - 4	[2, 13]

Table A.4: Variable cost fit model used for LCOW calculations in optimization. The first part of this table is used for the design variables. The second part of the table provides lists the devices that are indirectly optimized based on the outcome of the simulations.

Power range [kW]	**Cost** [$]
< 0.75	1700
0.75 - 1.5	2030
1.5 - 2.4	2909
2.4 - 3.3	3809
3.3 - 5.2	4612

Table A.5: DC Power Supply cost analysis for different power ranges [6].

Component	**No.**	**Cost [$]**		**Ref.**
		Unit	**Total**	
RO Membrane	1	388.5	388.5	[12]
Pump	1	396.9	396.9	[12]
Hardware	1	571.91	571.91	[12]
Storage Tank	1	566.44	566.44	[5]
Controls	1	299.92	299.92	[12]
Evaporation Pond	9 m^3	1651	14859	[21]

Table A.6: RO system costs.

Component	**Lifetime [years]**	**Ref.**
Pump	5	[12]
RO Membranes	3	[12]
Storage Tanks	25	[5]
Fixed Hardware	10	[12]

Table A.7: Components lifetime of a RO system.

Bibliography

[1] Ionics* v-20 stack brochure.

[2] Automation direct site catalog. URL https://www.automationdirect.com.

[3] Byd energy storage products. URL https://en.byd.com/wp-content/uploads/2017/06/b-box-eur-spec.pdf.

[4] Cnp pumps. URL https://www.cnppump.com/Water_treatment.html.

[5] Nto plastics site catalog. URL https://www.ntotank.com/cone-bottom-tanks.

[6] Tdk lambda programmable and high voltage products. URL https://www.genesysdcstore.com/collections/genesystm.

[7] URL https://www.canakit.com/raspberry-pi-4-2gb.html?cid=usd&src=raspberrypi.

[8] Apogee instruments pyranometer, . URL https://www.apogeeinstruments.com/sp-215-ss-amplified-0-5-volt-pyranometer/.

[9] Unbound solar black monocrystalline home pv specification sheet, . URL https://documents.unboundsolar.com/media/9434325__black_mono_1590187542.pdf?_ga=2.169481574.675583282.1619364761-1170437207.1619364761.

[10] Unbounded solar site catalog, . URL https://unboundsolar.com/.

[11] Quote received for v-20 stack from suez.

[12] Interviews from tata projects, ltd.

[13] Precision electric site catalog. URL https://www.precision-elec.com/shop/acs550-u1-038a-4/?gclid=Cj0KCQjw-O35BRDVARIs\AJU5mQXYQj9f1tQS23W4PpLCEbeJGgPdyzbsctJTWKJrnCXpUMjB1niss9oaAtIuEALw_wcB.

[14] Global optimization toolbox, matlab, 2020.

[15] A. Abulibdeh, T. Al-Awadhi, N. Al Nasiri, A. Al-Buloshi, and M. Abdelghani. Spatiotemporal mapping of groundwater salinity in al-batinah, oman. *Groundwater for Sustainable Development*, 12, 2021.

[16] M.R. Adiga, S. Adhikary, P. Narayanan, W. Harkare, S. Gomkale, and K. Govindan. Performance analysis of photovoltaic electrodialysis desalination plant at tanote in thar desert. *Desalination*, 67:59–66, 1987.

[17] S. Aminfard, F.T. Davidson, and M.E. Webber. Multi-layered spatial methodology for assessing the technical and economic viability of using renewable energy to power brackish groundwater desalination. *Desalination*, 450:12–20, 2019.

[18] J. A. E. Andersson, J. Gillis, G. Horn, J. B. Rawlings, and M. Diehl. CasADi – A software framework for nonlinear optimization and optimal control. *Mathematical Programming Computation*, 11(1):1–36, 2019. doi: 10.1007/s12532-018-0139-4.

[19] D. W. Bain, S.M. Watson, N.C. Wright, S.R. Shah, T. Buonassisi, D. Ramanujan, I.M. Peters, and A. G. Winter V. Optimization and design of a low-cost, village-scale, photovoltaic-powered, electrodialysis reversal desalination system for rural india. *Desalination*, 452:265–278, 2018.

[20] A. Barré, B. Deguilhem, S. Grolleau, M. Gérard, F. Suard, and D. Riu. A review on lithium-ion battery ageing mechanisms and estimations for automotive applications. *Journal of Power Sources*, 241:680–689, 2013.

[21] O. Barron, A. Raisat, G. Hodgson, D. Smith, E. Qureshi, D. McFarlane, E. Campos, and D. Zarzo. Feasibility assessment of desalination application in australian traditional agriculture. *Desalination*, 364:33–45, 2015.

[22] H. Birch and R. Weaver. Gwi market focus deck: Desalination reuse markets july 2020, 2020. URL https://gwi-private-storage.s3.amazonaws.com/production/reports/224/original/GWI_DesalData_Desalination___Reuse_Markets_July_2020.pdf?AWSAccessKeyId=AKIA3U2IRTUXITWDBSGJ&Expires=1619363034&Signature=5w3Gz3QTbd4x6rgWkIsLKUy%2F%2FSk%3D.

[23] E. F. Camacho and C. B. Alba. *Model predictive control*. Springer science & business media, 2013.

[24] T. Englert, A. Völz, F. Mesmer, S. Rhein, and K. Graichen. A software framework for embedded nonlinear model predictive control using a gradient-based augmented lagrangian approach (grampc). *Optimization and Engineering*, 20 (3):769–809, 2019.

[25] E. Farbod. Hybrid renewable energy systems for desalination. *Applied Water Science*, 10, 2020.

[26] J. Friedman, T. Hastie, and R. Tibshirani. *The elements of statistical learning*, volume 1. Springer series in statistics New York, 2001.

[27] R. Gelaro, W. McCarty, M.J. Suárez, R. Todling, A. Molod, L. Takacs, C.A. Randles, A. Darmenov, M.G. Bosilovich, R. Reichle, K. Wargan, L. Coy, R. Cullather, C. Draper, S. Akella, V. Buchard, A. Conaty, A. M. da Silva,

W. Gu, G. Kim, R. Koster, R. Lucchesi, D. Merkova, J. E. Nielsen, G. Partyka, S. Pawson, W. Putman, M. Rienecker, S.D. Schubert, M. Sienkiewicz, and B. Zhao. The modern-era retrospective analysis for research and applications, version 2 (merra-2). *Journal of Climate*, 30:5419–5454, 2017. doi: https://doi.org/10.1175/JCLI-D-16-0758.1.

[28] A. Gonzalez. Assessment of pilot-scale water purification module with electrodialysis technology and solar energy. *Applied Energy*, 206:1643–1652, 2017.

[29] L. Gurreri, A. Tamburini, A. Cipollina, and G. Micale. Electrodialysis applications in wastewater treatment for environmental protection and resources recovery: A systematic review on progress and perspectives. *Membranes*, 10, 2020.

[30] J. A. Hansen, B. C. Jurgens, and M. S. Fram. Quantifying anthropogenic contributions to century-scale groundwater salinity changes, san joaquin valley, califonria, usa. *Science of the Total Environment*, 642, 2018.

[31] W. He, S. Amrose, N.C. Wright, T. Buonassisi, I. M. Peters, and A.G. Winter V. Field demonstration of a cost-optimized solar powered electrodialysis reversal desalination system in rural india. *Desalination*, 476, 2020.

[32] W. He, S.R. Shah, A. LeHenaff, and A.G. Winter V. Performance analysis of time-variant voltage- and flow-controlled electrodialysis. 2021.

[33] S. Henbest, M. Kimmel, J. Callens, T. Brandily, M. Annex, J. Attwood, M. Bartels, I. Berryman, K. Bhavnagri, J. Chase, R. Chatterson, V. Cuming, D. Doherty, J. Frith, Gandolfo. A., T. Harries, A. Jain, D. Kang, Y. Liu, S. Lu, C. McKerracher, S. Sanghera, I. Savut, Y. Sekine, A. Wilton, L. Sutter, B. Vickers, and M. Ichinose. Bloomberg new energy finance report: New energy outlook. Technical report, 2020.

[34] K. Kikuchi, S. Takahashi, S. Kubota, S. Wakamatsu, Y. Eguchi, Y. Ikenaga, O. Kuroda, N. Sohma, S. Itoh, and K. Nishinoiri. An electrodialysis sea water desalination system system powered by photovoltaic cells. *Desalination*, 65:161–169, 1987.

[35] D. E. Kirk. *Optimal control theory: an introduction*. Courier Corporation, 2004.

[36] J. Kucera. *Desalination: water from water*. Wiley, New York, 2019.

[37] Anne-Claire LeHenaff. Time-variant solar-powered electrodialysis reversal desalination for affordable off-grid clean water supply. Master's project, Massachusetts Institute of Technology, Mechanical Engineering Department, 2019. This is a full MASTERSTHESIS entry.

[38] W. S. Levine, L. Grüne, R. Goebel, S. V Rakovic, A. Mesbah, I. Kolmanovsky, S. Di Cairano, D. A. Allan, J. B. Rawlings, and M. A. Sehr. Handbook of model predictive control. 2018.

[39] A. Mellit and A. M. Pavan. A 24-h forecast of solar irradiance using artificial neural network: Application for performance prediction of a grid-connected pv plant at trieste, italy. *Solar Energy*, 84(5):807–821, 2010. doi: 10.1016/j.solener.2010.02.006.

[40] A. Mellit, A. M. Alessandro Massi Pavan, E. Ogliari, S. Leva, and V. Lughi. Advanced methods for photovoltaic output power forecasting: A review. *Applied Sciences*, 30, 2020.

[41] D. C. Montgomery and G. C. Runger. *Applied statistics and probability for engineers*. John Wiley & Sons, 2010.

[42] P. Murray. *Electrodialysis and Electrodialysis Reversal - Manual of Water Supply Practices*. M38 (1st edition) edition, 1995.

[43] Central Ground Water Board Ministry of Water Resources RD and GR Government of India. *Ground Water Quality in Shallow Aquifers in India*. Faridabad, 2018. URL http://cgwb.gov.in/WQ/Ground%20Water%20Book-F.pdf.

[44] World Health Organization. *Guidelines for drinking-water quality, 4th edition, incorporating the 1st addendum*. World Health Organization - Licence: CC BY-NC-SA 3.0 IGO, 2017.

[45] WWAP (United Nations World Water Assessment Programme). he united nations world water development report 2015: Water for a sustainable world. Technical report, 2015.

[46] S.R. Shah, S. L. Walter, and A.G. Winter V. Using feed-forward voltage-control to increase the ion removal rate during batch electrodialysis desalination of brackish water. *Desalination*, 457:62–74, 2019.

[47] R. Wassef and H. Schüttrumpf. Impact of sea-level rise on groundwater salinity at the development area western delta, egypt. *Groundwater for Sustainable Development*, 2-3:85–103, 2016.

[48] P. D. Wellea, J. Medellín-Azuara, J. H. Viers, and M. S. Mauter. Economic and policy drivers of agricultural water desalination in california's central valley. *Agricultural Water Management*, 194:192–203, 2017.

[49] N. C. Wright and A. G. Winter V. Justification for community-scale photovoltaic-powered electrodialysis desalination systems for inland rural villages in india. *Desalination*, 352:82–91, 2014.

[50] N.C. Wright, S.R. Shah, S.E. Amrose, and A.G. Winter V. A robust model of brackish water electrodialysis desalination with experimental comparison at different size scales. *Desalination*, 443:27–43, 2018.