Deep Reinforcement Learning based Optimization of an Island Energy-Water Microgrid System

Roozbeh Ghasemi[1], Gersi Doko[2], Marek Petrik[2], Martin Wosnik[3,4,5,*], Zhongming Lu[6], Diane L. Foster[3,7,8], Weiwei Mo[1,*]

[1]Department of Civil and Environmental Engineering, University of New Hampshire, NH, 03824, USA
[2]Department of Computer Science, University of New Hampshire, NH, 03824, USA
[3]Department of Mechanical Engineering, University of New Hampshire, NH, 03824, USA
[4]Center for Ocean Engineering, University of New Hampshire, NH, 03824, USA
[5]Atlantic Marine Energy Center (AMEC), University of New Hampshire, Durham, NH 03824, USA
[6]Division of Environment and Sustainability, Hong Kong University of Science and Technology, Clear Water Bay, Hong Kong, China
[7]School of Marine Science and Ocean Engineering, University of New Hampshire, NH, 03824, USA
[8]Shoals Marine Laboratory, University of New Hampshire, NH, 03824, USA

*Corresponding authors:
Weiwei Mo: Weiwei.Mo@unh.edu
Martin Wosnik: Martin.Wosnik@unh.edu;

Abstract

This study optimizes an island energy-water microgrid using reinforcement learning (RL) to schedule the water system as a virtual battery. Using the microgrid in the Shoals Marine Laboratory as a testbed, a dynamic model simulating physical water and energy interactions was integrated with an RL algorithm to improve economic cost, sustainability, and reliability. The RL scenario outperformed the status quo, leading to a 7.04% improvement in the overall score. It also demonstrated superior cost and reliability outcomes compared to a single-objective, heuristic management approach, though with a slight reduction in sustainability. The results show that the RL model strategically reserves battery storage for periods of peak renewable energy generation and extends water system operation to circumvent pumping constraints, effectively using the water system as a virtual battery. Increasing the desalination rate was found to further enhance performance across all metrics, while enlarging water tank capacity offered minimal advantages.

Keywords: island energy-water microgrid; deep reinforcement learning; dynamic process-based simulation; fitted value iteration; multi-objective optimization; sustainability and reliability

1. Introduction

Microgrids have grown in popularity over the past decade as an alternative or complementary energy solution for communities and organizations in remote or isolated areas, or those seeking greater energy resilience and reliability [1], [2]. These small-scale, localized energy systems can function autonomously or in conjunction with the centralized grid [3], often incorporating renewable sources like solar and wind to improve sustainability and strengthen energy security. Currently, there are about 687 operational microgrids in the U.S., with a total capacity of 4.4 GW [4], and this number continues to grow [5]. Nevertheless, microgrids often require substantial energy storage capacity, which poses financial challenges for many communities due to the high costs involved [6]. Moreover, energy storage systems, particularly batteries, are limited by their finite charge-discharge cycles, and the variability of renewable energy sources further accelerates battery degradation over time [7]. To address these challenges, integrating water systems as "virtual batteries" into microgrids has gained traction [8]. This innovative approach leverages water treatment and storage processes to balance energy supply and demand, helping to mitigate the intermittency of renewable sources like solar and wind [9]. It is particularly advantageous in energy-intensive water supply applications, such as desalination. Despite the significant benefits of integrated energy-water microgrids, their effective management requires careful consideration of both energy and water supply and demand patterns to ensure their co-optimization.

As the integration of water systems as "virtual batteries" into microgrid designs is an emerging concept, it has not been widely adopted or deeply explored in previous microgrid studies. For those studies that do consider the interaction between energy and water systems, the physical interactions are often modeled using coarse time steps [10]–[19]. This can oversimplify the dynamics between energy and water supplies and demands, leading to an incomplete representation of the system's performance. Fine-grained time resolution is typically needed to fully capture the detailed balance of energy and water flows, especially in systems with variable renewable energy sources and fluctuating water demand. Furthermore, mixed integer linear and non-linear programming have been commonly used to co-optimize energy and water management strategies, typically focusing on objectives such as maximizing economic benefits, minimizing energy curtailment, or, less frequently, reducing environmental impacts like carbon footprint [11], [15]. Mixed integer programming (MIP) is a well-established mathematical optimization technique that handles both discrete integer and continuous variables, making it powerful for complex decision-making problems. However, it can become cumbersome when addressing the high non-linearities, multiple constraints, and significant uncertainties often present in microgrid systems, especially in relation to fluctuating water and energy demands [20]. Moreover, MIP is limited in handling non-stationarity, where parameters and constraints change over time, which often requires simplifications [20]. Due to these challenges, artificial intelligence (AI) methods have increasingly been explored as alternatives for optimizing microgrid systems, as AI techniques can better handle dynamic, uncertain, and complex systems without the need for as many simplifying assumptions [21]–[25]. Among these, reinforcement learning (RL) methods have been most widely applied to handle sequential decision-making problems, where an AI agent learns through trial-and-error interactions with the environment [26]. RL is especially useful in dynamic and uncertain settings, making it a powerful tool for real-time optimization and adaptive control in complex systems. Despite the popularity of RL in dynamic optimization problems, its application in the optimization of energy-water microgrids has been extremely limited. From the two studies that adopted RL methodologies in this research area, Xu et al. [20] applied Q-Learning to study the co-optimization of energy and water systems for an island microgrid and identified the optimal electric water heater operations that resulted in around 9% reduction in diesel consumption. Similarly, Kofinas et al. [27] used an

exploration/exploitation technique to optimize the battery usage by controlling the desalination system. They found the optimal water desalination scheduling that resulted in the highest water and energy availability. Both studies primarily focused on energy use or availability as performance metrics, while neglecting the microgrids' economic and environmental performances.

To address these knowledge gaps, we developed a process-based energy-water microgrid simulation model, which is then combined with a fitted value iteration-based RL algorithm to identify optimal trajectory of microgrid management decisions considering economic cost, sustainability, and reliability objectives. Our process-based simulation and RL modeling framework was tested on an existing off-grid microgrid at the Shoals Marine Laboratory (SML) on the Appledore Island, Maine. Through this work, we seek to answer the following research questions: 1) How does the RL management strategy compare to the current management scenario and a single objective, heuristic management scenario? 2) How does the RL strategy's behavior compare to the current and heuristic management scenarios? and 3) How do changes in key water system components impact the optimization outcomes? The following sections outline the study area and data (Section 2), the development of the process-based energy-water microgrid simulation model, evaluation metrics, and the RL optimization algorithms (Section 3), key findings and implications (Section 4), and conclusions (Section 5).

2. Study Area and Data Description

Shoals Marine Laboratory (SML) located on Appledore Island, the largest of nine islands in the Isles of Shoals, is six nautical miles off the east US coast at the border of Maine and New Hampshire [28]. SML is fully operational for about three months each year, from mid-June to mid-September. A schematic of SML's energy and water systems is provided in Figure 1a). Ongoing investments in the microgrid have led to a 50% reduction in energy usage and a 2% decrease in diesel consumption compared to 2007 levels. The microgrid currently generates electricity through 233 solar panels with a total nominal capacity of 68 kW, a Bergey Excel-10 wind turbine with the peak output of 12.6 kW, and two 27-kW diesel generators (operated one at a time), all supported by an absorbed glass mat (AGM) battery bank with a nominal capacity of 300 kWh. To extend the battery bank's service life, it is managed with a state-of-charge limit of 70% to 100%, allowing for a usable storage capacity of 90 kWh. Realtime system operation data is available through a dashboard and stored on a server for future analysis. More details related to the SML's microgrid system can be found in Ghasemi et al. 2023 [28], [29].

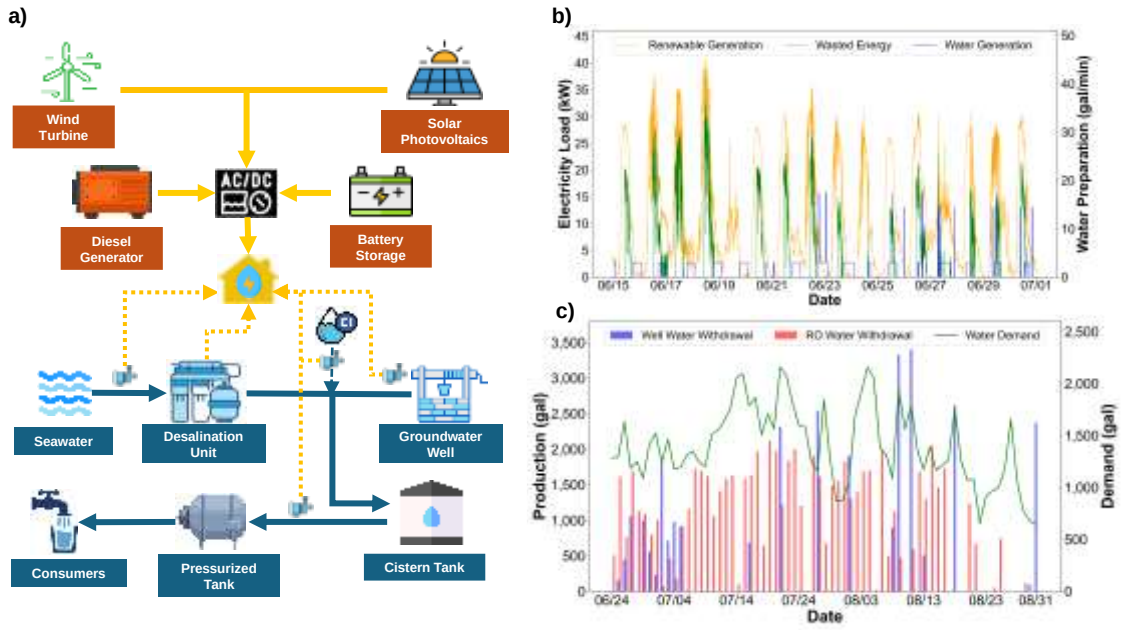

Figure 1. a) Schematic of the Shoals Marine Laboratory energy-water microgrid; b) the temporal relationship between the wasted renewable energy and water preparation under the status-quo; c) the relationships between groundwater, seawater withdrawal and water demand.

Water at SML is sourced from both groundwater and seawater desalination via reverse osmosis (RO). Groundwater is pumped at a rate of approximately 14.5 gallons per minute, based on pump capacity and hydraulic head, as verified by historical data. The desalination system draws seawater via an external pump and produces fresh water at a rate of 3-4 gallons per minute through a commercial-grade RO system [30]. Groundwater and desalinated seawater are blended and disinfected before storage. Two small peristaltic pumps regulate chlorine addition, maintaining a chlorine level of 2 ppm. Treated water is stored in a 14,000-gallon underground cistern tank. From there, water is transferred to a 5,000-gallon pressurized tank, which supplies water directly to users. The pressurized tank operates within a pressure range of 40 to 160 psi, corresponding to 20-80% of its total capacity. Water distribution is managed through automated controls that regulate transfers between the cistern and pressurized tank based on real-time pressure levels in the pressurized tank.

The water system's peak electricity load accounted for up to 43% of the microgrid's average instantaneous load. Given that groundwater is a less expensive operation, it is typically prioritized over seawater desalination. However, there is currently no formal protocol governing the use of these water sources. Instead, the SML manager relies on expert judgment to decide when to activate the groundwater or desalination systems and how long to operate them, based on well water levels, energy storage status, and weather conditions. Figure 1b) presents a 15-day sample of observed operational data, illustrating the significant amount of unused renewable energy due to limited storage capacity. The timing mismatch between peak energy generation and peak water demand suggests that surplus renewable energy could be better utilized for water production, highlighting the potential value of utilizing water system as a virtual battery to improve system performance.

We used data from 2022 to investigate the optimal management strategy for the energy-water microgrid. Minute-by-minute weather, electricity, and water data were sourced from SML's monitored data, while wind speed data were obtained from the nearby IOSN3 station. We pre-

processed the observed data to address any missing entries, which accounted for less than 1% of the total records for the selected year. For data gaps shorter than 10 minutes, linear interpolation was used, filling gaps based on the difference between the last recorded value and the first available value after the gap. For longer gaps, the missing data were assumed to follow the trend of the previous day. Figure 1c) shows the monitored daily drinking water production from both systems and the water demand at SML over the study period. The total water consumption during the period was 124,529 gallons, averaging 1,354 gallons per day.

3. Methods

In this section, we first introduce the process-based energy-water microgrid model (Section 3.1). Next, we outline the three performance metrics developed for evaluating the energy-water microgrid (Section 3.2). Lastly, we describe the RL algorithm used to optimize microgrid management by co-maximizing these performance metrics (Section 3.3).

3.1 Dynamic Process-Based, Integrated Energy-Water Microgrid Modeling

The process-based energy-water microgrid model was developed in Python V3.11.7 and operates on a minute-by-minute basis. Figure 2 represents a simplified schema of the model, which contains two main components: an energy sub-model and a water sub-model. The energy sub-model simulates wind, solar, and diesel energy supply, alongside energy demand and balance. Particularly, wind energy output is calculated based on prevailing wind speeds at each time step, while solar energy is estimated using instantaneous solar radiation and ambient temperature, following equations from Ren et al. (2020) [31] and Lilienthal (2005) [32]. Diesel serves as a backup power source, with its operation governed by predefined rules that consider battery storage levels, energy demand, and renewable energy availability. Any excess renewable energy is not utilized. Further details on the energy sub-model's operation can be found in Ghasemi et al [29].

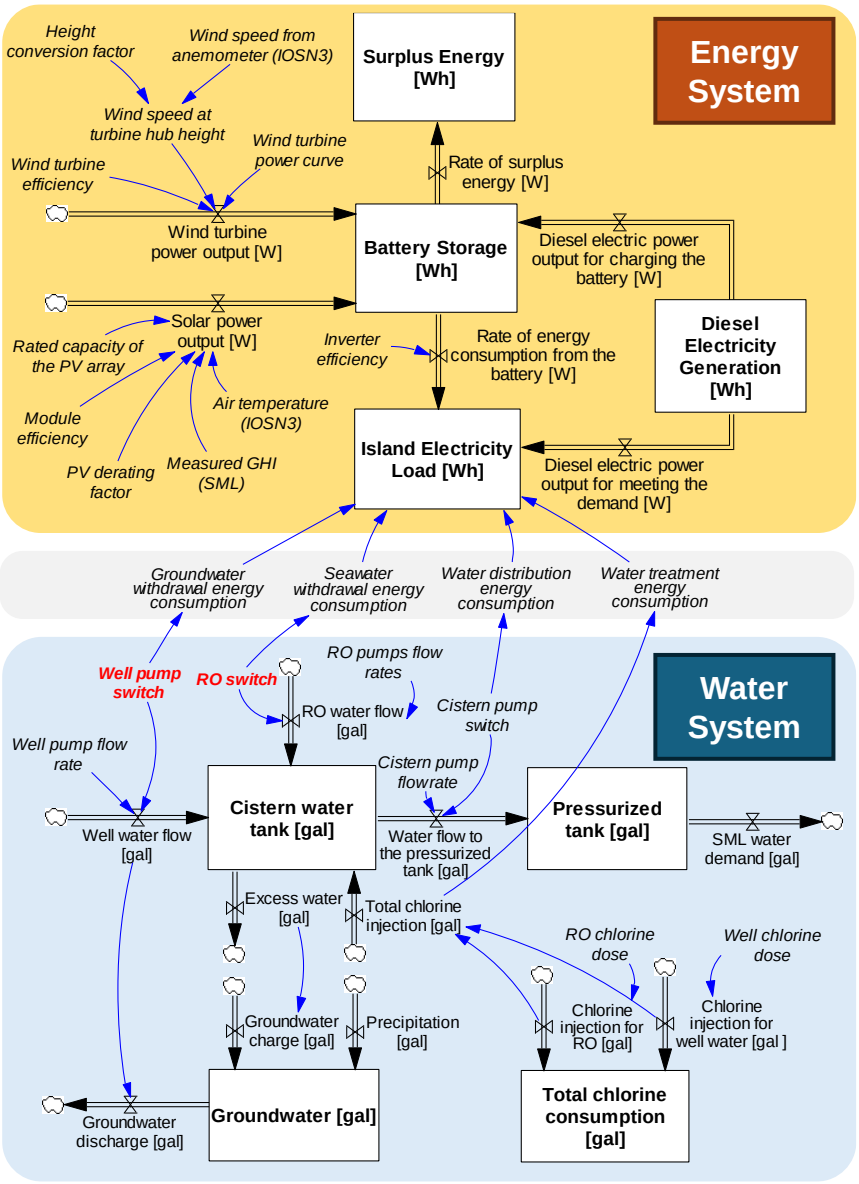

Figure 2. A schematic of the process-based energy-water microgrid simulation model highlighting the interactions between the energy and the water systems. Variables highlighted in red are the two main decision variables for the optimized control of the energy-water microgrid system.

The water sub-model simulates the processes and energy use associated with groundwater and seawater intake, treatment, and distribution. The energy uses for these unit processes are determined by the activation of their respective pumps and corresponding flow rates. When a pump is operating, its energy consumption is calculated as the product of its activation and power rating per minute [33]. To enhance the utilization of surplus renewable energy and conserve groundwater for system resilience, a groundwater recharge mechanism was added to the model. When both the cistern and pressurized tanks reach their maximum operational capacities and surplus renewable energy is available, the excess renewable energy is used to desalinate water, which is then diverted to recharge the groundwater well. To model changes in the groundwater well level, a “bathtub” model was developed as a simplified representation of the island’s watershed. This model conceptualizes the watershed as a rectangular storage tank (Figure 3a)). The tank is replenished by precipitation and groundwater recharge and depleted through groundwater withdrawal. The volume of water stored in the watershed can be calculated using Eq. 1.

$$\Delta H_t \times A_w = Groundwater\ recharge_t + Precipitation_t - Groundwater\ withdrawal_t \quad \text{Eq. 1}$$

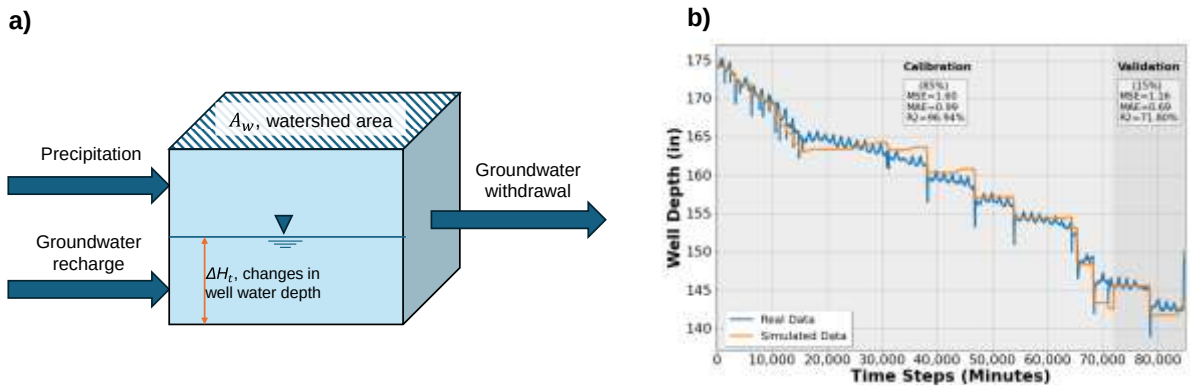

Figure 3. a) A conceptual island watershed stock and flow model adopted in this study. A_w is defined as the hypothetical catchment area of the watershed and ΔH_t is the instantaneous well water change. b) The observed and simulated well water level during both calibration and validation periods.

Based on observed groundwater level, groundwater withdrawal, and precipitation data from the island, we calibrated the watershed area, A_w, which is subsequnetly used for predicting groundwater depth. Groundwater level was monitored over a span of 59 days between June 25 and August 23, 2022. Approximately 50 days of this data were used for model calibration, while the remaining 9 days were used for validation. The hypothetical catchment area was varied systematically from 1 to 1,500 square feet, with increments of 1 square foot. The calibration accuracy was assessed using Mean Squared Error (MSE) and R^2 values. The optimal catchment area, A_w =1097 ft², was identified, resulting in an MSE of 1.6 in² and an R^2 of 96.94%. Validation over the remaining 15% of the data yielded an MSE of 1.16 in² and an R^2 of 71.80%. Figure 3b) illustrates the comparison between observed and simulated groundwater levels for both calibration and validation periods.

3.2 Energy-Water Microgrid Evaluation Metrics

We defined three criteria to evaluate the performance of the energy-water microgrid, calculated as cost, sustainability, and reliability scores.

Cost score represents the economic cost associated with the operation of the energy-water microgrid. It is calculated as a ratio between the instantaneous cost at each time step and the maximum possible cost in a single time step. The instantaneous cost includes the cost of diesel consumed in diesel generator, the opportunity cost related to the wasted renewables calculated as surplus energy, and the energy infrastructure depreciation cost. Eq. 1 shows the calculation of the cost score.

$$Cost\ Score_{a,t} = 1 - \frac{(E_{Grid,t}+E_{RO,a,t}+E_{Well,a,t})*f*P_{Diesel}*S_{Diesel,t}+E_{Waste,t}*f*P_{Diesel}+\sum_m C_{d_{m,a,t}}}{C_{Max}} \quad \text{Eq. 2}$$

where a is the index of the action taken in the time step t (a vector of well pump and RO switches); $E_{RO,a,t}$ is the energy consumption of the RO system at time t given the switch states defined by vector a. $E_{Well,a,t}$ is the energy consumption of the well pump at time t given the switch state a. $E_{Grid,t}$ is any other energy consumptions at SML at time t. P_{Diesel} represents the average diesel price over the study period, which is estimated to be $5.39 per gallon. $S_{Diesel,t}$ indicates whether the energy is supplied by diesel at time t, 0 indicating that

the diesel generator is off, and 1 indicating the diesel generator is on. $E_{Waste,t}$ is the amount of waste renewable energy at time *t*. f is the conversion factor, 0.0903 gallons of diesel/kWh of electricity [34]. The opportunity cost is included to minimize energy waste. We assumed that the unit cost of waste energy is the same as diesel price in our analysis. *m* is the index of each system component. $\sum_m C_{d_{m,a,t}}$ is the total depreciation cost of all system components *m*, as a function of well pump and RO operation switches a, including battery storage, PV panels, diesel generator, RO system including filters, and water pumps, calculated based on asset price of the parts (see the SI for asset price details). Depreciation was calculated using straight-line method with a salvage value of zero [35]. Depreciation only occurs when the unit infrastructure is in operation. Finally, C_{Max} is calculated based on the maximum cost in a single step based on asset prices of components and the observed economic cost data. The maximum cost utilized in this study is $0.97 per minute.

Sustainability score represents the sustainability of SML's energy-water microgrid using carbon emissions as a surrogate measure. It is calculated as a ratio between the instantaneous carbon emissions at each time step and the maximum possible carbon emissions in a single time step. Instantaneous carbon emissions include emissions associated with diesel consumed in diesel generator, the opportunity emissions related to the waste energy, and the carbon emissions associated with infrastructure construction and maintenance. Eq. 2 presents the numerical calculation of the sustainability score.

$$Sustainability\ Score_{a,t} = 1 - \frac{(E_{Grid,t}+E_{RO,a,t}+E_{Well,a,t})*f*CF_{Diesel}*S_{Diesel}+E_{Waste}*f*CF_{Diesel}+\sum_m CF_{m,a,t}}{CF_{Max}} \quad \text{Eq. 3}$$

where CF_{Diesel} is the carbon emissions associated with unit diesel consumption, 1.19 kg CO_2e/kWh. $\sum_m CF_{m,a,t}$ is the carbon emissions of all system components *m*, as a function of well pump and RO operation switches a (see SI for carbon emission details of each microgrid component). Finally, CF_{Max} is the maximum possible carbon emission value in a single step. We estimated the carbon footprint of each system component using the SimaPro 9.5.0.2 with IPCC 2021 GWP100 method. Details of the carbon emission estimations can be found in the SI. The maximum sustainability value utilized in this study is 2.49 kg CO_2e per minute.

Reliability score represents the reliability of SML's freshwater reserve using the groundwater well depth as a surrogate metric. It is calculated as a ratio between the well's instantaneous water availability and the maximum water availability. Instantaneous water availability is estimated as the difference between the current and the lowest possible water level at each time step. The maximum water availability is estimated as the difference between the highest possible and the lowest possible water levels. Eq. 4 captures the calculation of the reliability score.

$$Reliability\ Score_{a,t} = \frac{H_{Well,t}-H_{Low}}{H_{High}-H_{Low}} \quad \text{Eq. 4}$$

where H_{Low} and H_{High} are the maximum and minimum possible well water level measured from the well bottom, set at 100 inches and 250 inches respectively based on historical observed data. $H_{Well,t}$ is the instantaneous well water level in each time step, which is an output of the simulation model.

3.3 Reinforcement Learning

3.3.1 Dynamic Programming

Reinforcement Learning (RL) is a method where a machine learns the best action through interaction with its environment [36]. Here the RL algorithm trains a virtual agent, which interacts with the simulation model described in Section 3.1 and makes management decisions about the energy-water microgrid to achieve the highest possible reward. The reward is the immediate feedback of the system that indicates the immediate benefit resulting from the agent's action [36]. A simple reward can be defined as a composite score based on the cost, sustainability, and reliability scores described in Section 3.2 using Eq. 5.

$$\begin{cases} r(s,a)_t = c \times Cost\ Score_{a,t} + s \times Sustainability\ Score_{a,t} + r \times Reliability\ Score_{a,t} - \lambda \\ c + s + r = 1 \end{cases} \quad \text{Eq. 5}$$

where $c, s, and\ r$ are the weighting factors that can be defined based on the stakeholders' interests. In this study, we first defined the three scores as equal weighted, and then conducted sensitivity studies on the weighting. λ is defined as a penalty term to penalize the agent for violating the constraints. The agent receives penalties based on the severity of each constraint violation, such as insufficient water supply, over-topping or overusing the well, running the water system while the energy system relies on diesel generation, and draining the battery when renewable energy is inadequate. The penalty term also provides feedback to discourage rapid switch changes that occur in less than 5 minutes to avoid water pump and desalination system damage.

In complex systems such as the microgrid system of study, actions impact not only the immediate time step but also the future time steps [36]. Therefore, our RL algorithm seeks to maximize both immediate and future rewards. We use "Value" to indicate the discounted sum of immediate and future rewards, which can be calculated using Eq. 6.

$$V_t(s_t;\ a_t, \dots,\ a_{T-1}) = \sum_{k=t}^{T-1} \gamma^{k-t}\, r_k(s_k, a_k) + r_T(s_T), \quad s \in S, and\ t \in \{1, .., T\} \quad \text{Eq. 6}$$

where $V_t(s_t;\ a_t, \dots,\ a_{T-1})$ is the Value at time t, and $r_k(s_k, a_k)$ is the immediate reward at time step k, iterating from time steps *t* to *T-1*, under system state s_k and management action a_k [36]. γ is the discount factor that calculates the present value of the future rewards. To ensure a sufficiently forward-looking assessment of rewards, we established a γ value of 0.9993. This allows the agent to factor in consequences up to one day ahead. In this work, the agent makes decisions in each minute in 92 days of simulation, which equals a total of 132,480 steps.

The management decisions made by the agent are whether to activate or deactivate the RO and the groundwater intake systems. Thus, the action space comprises four finite options: activating both systems, deactivating both systems, and activating one while deactivating the other. These decisions are made based on the "state" of the energy-water microgrid of study. In RL, a state refers to a specific configuration or situation that the agent perceives in its environment at a given point of time [36]. In this study, system state is captured through 12 variables/state features. Six of these state features are exogenous variables, including solar irradiation, wind speed, temperature, precipitation, water demand, and energy demand. The remaining six state features are endogenous, including diesel generator status, cistern pump switch, battery state of charge, pressure tank level, cistern tank level, and well depth. Table 1 details all 12 state features and their respective ranges.

Table 1. Energy water microgrid model's actions and state features

Type		Denotation	Variables	Range	Scores involved	Notes
Decision Variables		**RO**	RO Switch	0 or 1	All three scores	0 considers as "Off" and 1 considered as "On"
		Well	Well Switch	0 or 1	All three scores	0 considers as "Off" and 1 considered as "On"
States	**Endogenous**	**t**	Time step	$0 \leq \mathbb{N} \leq 132{,}479$	All three scores	92 days of operation in minute resolution
		Φ_1	Generator status	0 or 1	Cost and Sustainability Scores	0 considers as "Off" and 1 considered as "On"
		Φ_2	Cistern pump status	0 or 1	Cost and Sustainability Scores	0 considers as "Off" and 1 considered as "On"
		Φ_3	Battery state of charge	$70 \leq \mathbb{R} \leq 100$	Cost and Sustainability Scores	The range is determined based on the current operation rule of the SML
		Φ_4	Pressure tank level	$0 \leq \mathbb{R} \leq 5{,}000$	Reliability Score	In the unit of gallons
		Φ_5	Cistern tank level	$0 \leq \mathbb{R} \leq 1{,}4000$	Reliability Score	In the unit of gallons
		Φ_6	Well depth	$0 \leq \mathbb{R}$	Reliability Score	In the unit of inches
	Exogenous	Φ_7	Solar GHI	$0 \leq \mathbb{R}$	Cost and Sustainability Scores	In the unit of W/s^2
		Φ_8	Wind speed	$0 \leq \mathbb{R}$	Cost and Sustainability Scores	In the unit of m/s
		Φ_9	Ambient temperature	$\mathbb{R}$	Cost and Sustainability Scores	In the unit of degree Celsius
		Φ_{10}	Energy demand	$0 \leq \mathbb{R}$	Cost and Sustainability Scores	Instantaneous electricity load in the unit of kW
		Φ_{11}	Water demand	$0 \leq \mathbb{R}$	Reliability Score	Instantaneous water demand in the unit of gallons
		Φ_{12}	Precipitation	$0 \leq \mathbb{R}$	Reliability Score	In the unit of inches

3.3.2 Fitted Value Iteration

In tabular RL, the training process typically involves presenting all feasible states and actions to an agent, and then determining the exact optimal action for each state based on its highest achievable Value. Given the large number of time steps (132,480 steps) and system states (combinations of 12 states) and a small action space (4 actions) involved in this study, we have opted to utilize a modified RL method called Fitted Value Iteration (FVI) to achieve the optimization purpose. FVI is particularly suited for the present study because it can predict the Value of sampled states in a continuous state space over a foreseeable time horizon by exhausting a small action space. FVI is comprised of two steps: 1) agent training and 2) calculating the optimal control sequence.

a) Agent training

To train the agent, FVI starts from the last time step and goes backward to estimate the Values. In each time step, the agent provides a sample size *n* for each state feature $\widehat{\Phi_i}$, $i = 1, \dots, 12$. The maximum Value out of the four possible actions for each sample is calculated using Eq. 7.

$$\hat{V}^*{}_t(\hat{\mathrm{s}}_t) = \max_{a \in A}\left[r_t(a, \hat{\mathrm{s}}_t) + \gamma\, \tilde{V}^*{}_{t+1}\left(f(a, \hat{\mathrm{s}}_t)\right)\right]$$
$$\hat{\mathrm{s}} = \left[\widehat{\Phi_1}, \dots, \widehat{\Phi_{12}}\right], \hat{\mathrm{s}}_{t+1} = f(a, \hat{\mathrm{s}}_t), \gamma \epsilon [0,1] \quad \text{Eq. 7}$$

where $\hat{V}^*{}_t(\hat{\mathrm{s}}_t)$ is the highest achievable Value of the sampled state $\widehat{s_t}$ at time step *t*; $\widehat{s_t}$ is a vector of one sampled state comprised of 12 state features of $\widehat{\Phi_1}$ through $\widehat{\Phi_{12}}$; a is the action within the action space *A*; *f* is the transition function that transits the state $\hat{\mathrm{s}}$ from time t to *t*+1 considering action *a*; and γ is the discount factor that calculates the present Value of the

future rewards. The algorithm uses a predictive model generated in the previous step, $\tilde{V}^*{}_{t+1}(f(a,\hat{s_t}))$, to estimate the future rewards (more specifically the Value for the next time step). The transition function $f(a,\hat{s_t})$ is a part of the process-based simulation model that takes one state (a vector of state features $\hat{s}$) and action a at any given time step and defines the next state. For time T, agent's achievement is associated with the instantaneous reward as there is no future rewards in the last time step. This sampling process significantly reduces computing time as it prevents the need for an exhaustive search.

Using the samples selected from each time step, the FVI algorithm fits a model that predicts the maximum Value based on the current system states, $\tilde{V}^*{}_{t+1}(f(a,\hat{s}))$. The algorithm originally took a linear regression method and stores the coefficients using Eq. 8.

$$\tilde{\omega}_{t=T}, \tilde{\beta}_{t=T} \leftarrow \arg\min_{\tilde{\omega}_t,\tilde{\beta}_t} \sum_{i=1}^{n} (\widehat{V}^*{}_{t=T}(\hat{s}_i) - (\hat{s}_i{}^T \tilde{\omega}_t + \tilde{\beta}_t))^2 \qquad \text{Eq. 8}$$

where $\tilde{\omega}_{t=T}$ and $\tilde{\beta}_{t=T}$ are the coefficients of the linear regression at time step T and $\hat{s}_i{}^T$ is the transposition of the i^{th} sampled state vector $\hat{s}_i$. The output of this step is T number of predictive models that each of them predicts the maximum achievable future rewards ($\tilde{V}^*{}_t$) associated with any state.

Due to the high computational burden of the training step, we made a few simplifications. We trained the predictive model based on a large sample size, n = 2.6 million, while sampling uniformly for all time steps (T=132,480) for 20 times (n = 20 × 132,480=2.6 million) to include diversity in time dependent state features, for instance, to include both days and nights. The sampling strategy is explained in Section 3.3.3 These samples were used to train one comprehensive predictive model $\tilde{r}^*(\hat{s})$. We used an Artificial Neural Network (ANN) model with three layers (32, 64 and 32 neurons, respectively) to train a model predicting maximum achievable reward $\tilde{r}^*$ based on the sampled states $\hat{s}$. A weight decay of 0.01 was considered to lower the chance of overfitting (see the SI for details). We ran the training process 100 times considering each time with a new set of sampled states of n = 2.6 million and selected the best performing ANN model that gives the best composite score in Eq. 5. This trained model was then used in all the time steps with a time discount factor as indicated in Eq. 9.

$$\tilde{V}_t{}^*(\hat{s}) = \tilde{r}^*(\hat{s}) \times \frac{1-\gamma^{(T-t+1)}}{1-\gamma} \qquad \text{Eq. 9}$$

where $\tilde{V}_t{}^*(\hat{s})$ is the time discounted maximum Value. $\tilde{r}^*(\hat{s})$ is the predicted maximum reward for each sample based on system state using the ANN model.

b) Calculate the optimal control sequence

Once the predictive model for the maximum Value in each time step is determined, the algorithm starts from initial time t=0 to find the best action that results in the optimal Value. The system states associated with the initial time step was obtained from the observed data at 06/01/2022 00:00. The agent then utilizes the observed exogenous state values, and the process-based model simulated endogenous states to determine the optimal control sequence for the next time step. In each time step, the agent finds the optimal action (a_t^*) considering both immediate and future rewards using Eq. 10. The algorithm is executed for all ANN models trained in 100 iterations to obtain the optimal control sequences.

$$\begin{cases} a_t^* = \arg max_{a\in A}[r_t(a,\mathrm{S_t}) + \gamma\, \tilde{V}_{t+1}^*(f(a,S_t))], & t = 0,\dots,T-1 \\ a_T^* = \arg max_{a\in A}[r_T(a,\mathrm{S_T})], & t = T \end{cases} \qquad \text{Eq. 10}$$

For each ANN model, the final output of the algorithm is a sequence of optimal management actions associated with each time step *t*, referred to as the optimal control sequence: $a_1^*, a_2^*, a_3^*, \dots, a_T^*$. Once the optimal control sequence was obtained, we compared the outcomes of the RL with the status-quo of system operation in the study period.

3.3.3 Stochasticity of the State Features

To enhance the predictive capability and account for stochastic variability in the FVI model beyond historical climate patterns, we developed a separate algorithm for generating exogenous states. We utilized five years of historical climate data to generate stochastic values for wind speed, global horizontal irradiance (GHI), and air temperature. This data was reformatted into 92-day segments, with each day featuring five distinct 24-hour weather patterns. For GHI and temperature, we first identify the specific day over the 92-day period, and then randomly select one out of five patterns associated with the day. Wind speed and precipitation, however, are sampled randomly from the entire pool of 92×5 days of observed data. The availability of historical water and energy demand data is limited. To address this, we assume that these demands vary based on the time of day and are influenced by previous and subsequent demand values. We model the probability of water and energy demand occurring as a normal distribution for the 30 minutes before and after each time step. For each time step, we calculate the mean and standard deviation of demands within a 1-hour window surrounding the time step. Using these parameters, we generate random demand values from a normal distribution. Endogenous states, on the other hand, are sampled randomly using a uniform function within the specified ranges outlined in Table 1.

4 Results and Discussion

4.1 Performance Comparison of the RL, Status-Quo, and Heuristic Scenarios

Figure 4 compares the performance of the RL scenario with the status quo and a heuristic scenario. The status quo reflects the current manual operations of SML, while the heuristic scenario represents a single-objective approach common in microgrid management, focused solely on maximizing renewable energy utilization. In the heuristic scenario, the water system is activated whenever surplus renewable energy is available. Both groundwater and seawater desalination systems are used to fill the cistern and pressurized tanks. Once these tanks reach capacity, only seawater desalination is employed for recharging the groundwater well.

Compared to the status quo, the RL scenario enhanced all three scores, achieving approximately 0.60% improvement in cost, 0.44% in sustainability, and 20.08% in reliability scores, resulting in a 7.04% improvement in the overall weighted total score. These improvements translate to a $771.03 saving in economic cost, a 1.45 Mg reduction of CO_2e emissions, and an average increase of 28 inches in groundwater level over the 92-day study period. Our results indicate that the current operations have attained high performance levels in terms of sustainability and cost, leaving limited room for further refinement. The most significant improvements are observed in the reliability score. From an energy consumption perspective, the RL scenario achieved a 1.64% reduction in diesel consumption and a 38.84% reduction in wasted renewable energy. This translates to a diesel saving of 56 kWh and an increased renewable energy utilization of 1,608 kWh - equivalent to a virtual battery capacity of 17.5 kWh per day, or 23.3% of the current storage capacity. It is worth noting that the RL scenario results in higher total energy consumption compared to the status quo. This increase arises from the RL algorithm's deliberate use of groundwater recharge to improve system reliability, a strategy not utilized in the *status quo* operations. This approach balances increased energy consumption with significant gains in reliability and reduced wastage, demonstrating the RL's effectiveness in enhancing overall system performance.

Compared to the heuristic scenario, the RL scenario presents a 10% increase in the weighted total score. Specifically, the RL approach enhances the cost score by 0.16%, translating to a saving of $205.61, and achieves a substantial 30.08% improvement in reliability, increasing the groundwater level by 42.11 inches. The RL scenario also outperforms the heuristic scenario in renewable energy utilization by 33% (1,248 kWh). Nevertheless, the RL scenario results in a 0.18% lower sustainability score as compared to the heuristic policy (594 kg of CO_2e), as well as a 16.86% higher diesel consumption (484 kWh). This suggests that while a single-objective policy focused solely on renewable energy utilization can still achieve significant diesel reductions, it does so at the expense of groundwater reliability. Notably, the large diesel reduction in the heuristic policy does not correspond to an equally significant sustainability score improvement, mainly due to additional carbon emission factors considered in the sustainability score. This underscores the limitations of single-objective, use-phase-focused optimizations, which may overlook the broader life cycle impacts of the solutions being evaluated, resulting in sub-optimization.

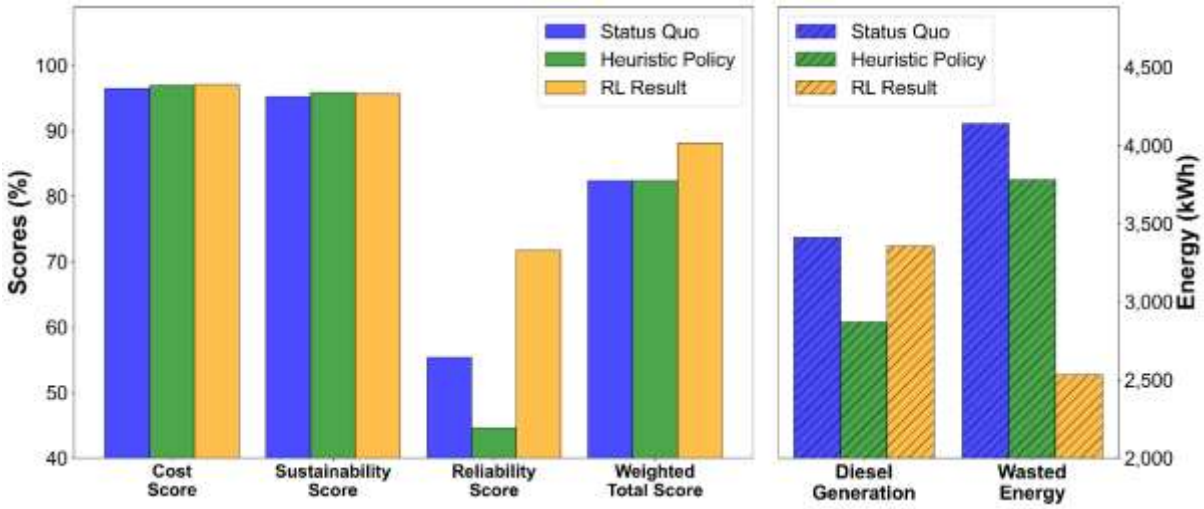

Figure 4. Comparison of the RL optimized energy-water microgrid management solution and the status quo. Yellow bars indicate the RL-optimized solution, while the blue bars indicate the status quo.

4.2 Behavior Comparison of the RL, Status-Quo, and Heuristic Scenarios

Figure 5 compares the temporal trends in cistern tank water levels, groundwater well levels, cumulative diesel energy consumption, cumulative wasted energy, and daily renewable energy utilization across the RL, status quo, and heuristic scenarios. The RL scenario consistently maintains the cistern tank at its maximum capacity (Fig. 5a)), contrasting with the status quo and heuristic approaches. This indicates that the RL scenario prioritizes maximizing the use of water tanks and groundwater recharge to optimize renewable energy utilization, with groundwater recharge occurring only after both the cistern and pressurized tanks reach full capacity. Notably, the status quo approach intentionally depletes the cistern tank toward the end of the season (approximately the final two weeks), a trend not observed in the RL and heuristic scenarios.

The groundwater level is better preserved in the RL scenario compared to both the status quo and the heuristic scenario (Figure 5b)). During periods of higher water demand, particularly from mid-July to mid-August, the likelihood of groundwater recharge under the RL scenario decreases, resulting in a reduced groundwater preservation benefit during this time. The heuristic approach, in particular, underperforms in maintaining groundwater levels, even

relative to the status quo. This is because groundwater is more frequently withdrawn to fill water tanks in order to utilize excess renewable energy.

A closer look at cumulative diesel generation (Figure 5c) and wasted renewable energy (Figure 5d) reveals that the RL scenario prioritizes minimizing wasted renewable energy over reducing diesel consumption to gain a higher overall reward. Figure 5e) further examines the RL model's behavior on a typical day, highlighting its strategy to achieve this goal. Compared to the status quo and heuristic scenarios, the RL scenario is the last to reach battery curtailment, as indicated by the first drop in Figure 5e). This suggests that the RL model strategically uses the water system earlier in the day to "store" excess renewable energy, ensuring that the water pumping rate does not become a limiting factor for renewable energy utilization. In contrast, the heuristic scenario prioritizes charging the battery first, using the water system as a "virtual battery" only after the battery is fully charged. However, once the battery reaches capacity during peak renewable energy generation, the amount of excess energy that can be utilized by the water system is constrained by its pumping capacity, which explains the heuristic scenario’s overall lower renewable energy utilization as compared to the RL scenario. The status quo, on the other hand, operates on a water-demand-driven model without actively charging the battery or water tanks to maximize renewable energy use. As such, battery curtailment is reached later in the day compared to the heuristic scenario. Furthermore, the RO system activation in the status quo is delayed due to manual operation, as evidenced by the plateau following the initial curtailment in the blue line, further limiting excess renewable energy utilization.

This analysis highlights the RL model's strategic decision-making, which reserves battery storage for later in the day when renewable energy generation peaks. By allowing the water system to run for longer periods, the RL approach effectively circumvents the pumping capacity limitations, optimizing the water system as a virtual battery. These actions demonstrate how a carefully calculated, risk-taking approach can lead to greater overall system benefits.

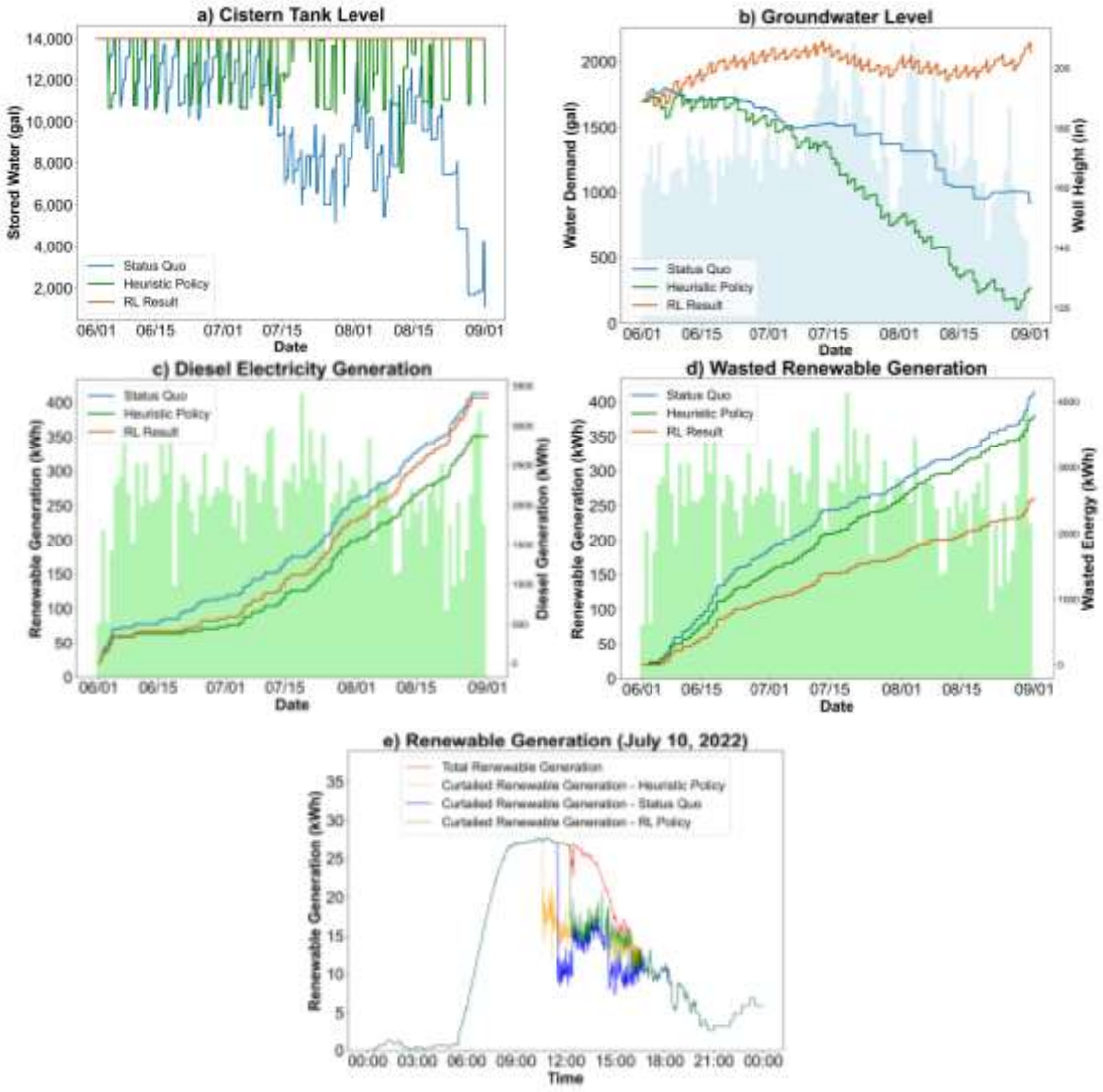

Figure 5. Comparison of the temporal trends in a) cistern tank water level, b) groundwater well water level, c) cumulative diesel energy consumption, d) cumulative wasted energy, and e) daily pattern of renewable energy utilization across RL, status quo, and heuristic solutions. The blue background bars in b) show the daily water demand throughout the simulation, while the green background bars in c) and d) show the daily renewable energy generation.

4.3 The Influence of Tank Size and Desalination Rate on the RL Solution

Figure 6 presents the three performance scores - cost, sustainability, and reliability - as well as the weighted total scores under various cistern tank sizes and desalination rates, to assess how system design impacts the outcomes of the RL solutions. Three reverse osmosis (RO) desalination rates were examined (2, 3, and 4 gallons per minute), using energy consumption rates obtained from field measurements [33]. The cistern tank capacity was also varied, ranging from 6,000 to 20,000 gallons in increments of 2,000 gallons. Overall, increasing the desalination rate consistently improves all three performance scores. A higher desalination rate enables greater utilization of excess renewable energy, reduces diesel consumption, and enhances groundwater recharge due to more frequent well recharging. These benefits outweigh the potential "costs" associated with operating a larger desalination system. Increasing cistern tank size, on the other hand, will slightly reduce cost and sustainability scores because of the higher utilization of well pump to maintain the cistern tank at its maximum level. It has a slight benefit to reliability score as groundwater is used less with a larger tank size.

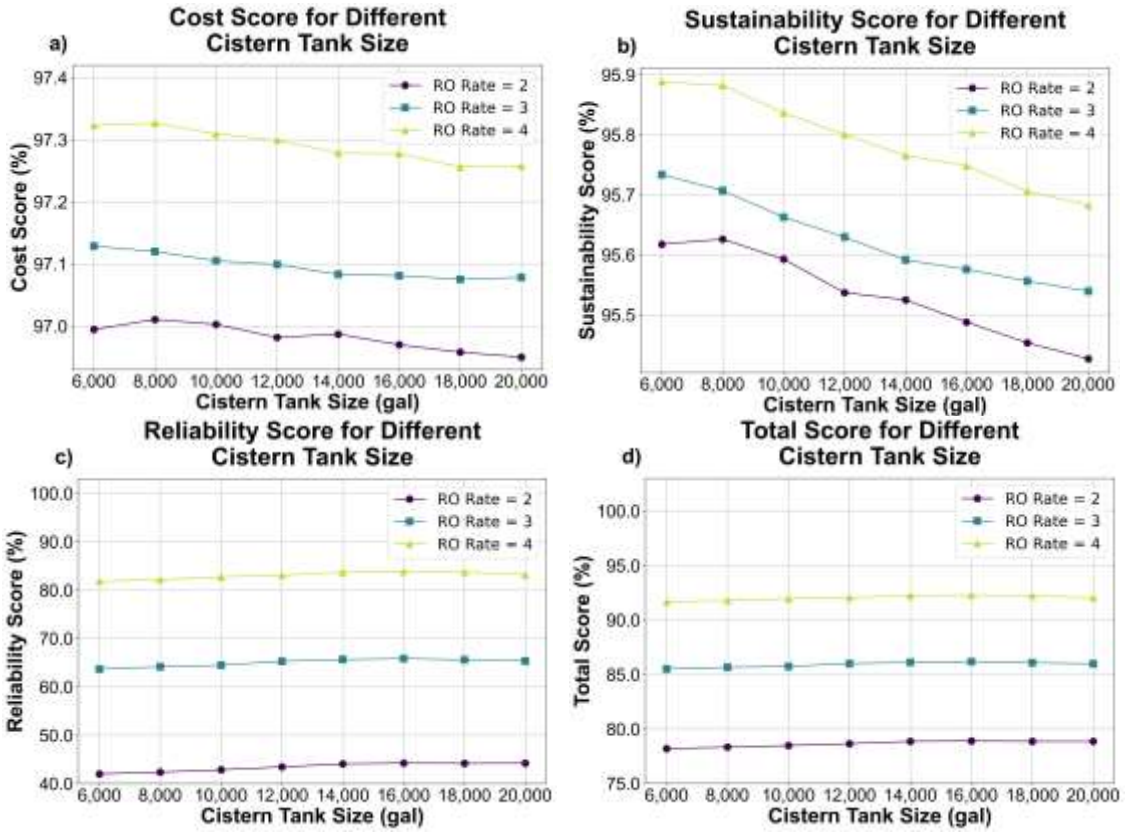

Figure 6. The influence of cistern tank size and desalination rate on the performance of the RL-optimized solutions

4.4 The Influence of Performance Metric Weightings on the RL-Optimized Solution

To explore the impact of performance metric weightings on the RL-optimized solutions, we systematically varied each score's weight from 0 to 1 (0, 1/3, and 1) while keeping the other weights equal, ensuring the sum of the weights remained 1. Each variation was tested 100 times with different samples, and the model with the highest total score for each scenario was selected. Figure 10 illustrates the changes in the three performance scores as their respective weights are adjusted. Overall, the reliability score is the most sensitive to changes in weighting, while the cost and sustainability scores show relative stability with only minor improvements. This suggests the microgrid has limited room for further improvement in terms of cost and sustainability. For example, increasing the cost score weight from 0 to 1 results in a marginal increase in the score, from 97.07% to 97.11%, leading to cost savings between $771.0 and $822.4 over one season compared to the status quo. Similarly, changing the sustainability score weight from 0 to 1 improves the sustainability score from 95.53% to 95.76%, equating to carbon savings of 1,121.6 to 1,880.3 kg CO2e per season. The reliability score shows a more pronounced and consistent improvement with increasing weight, rising from 55.85% to 78.27%, which corresponds to an increase in the average groundwater level from 6 to 37 inches. These findings highlight that while the cost and sustainability scores are relatively robust to changes in their weightings, the reliability score is more responsive to variations, indicating that weighting adjustments significantly affect groundwater preservation performance.

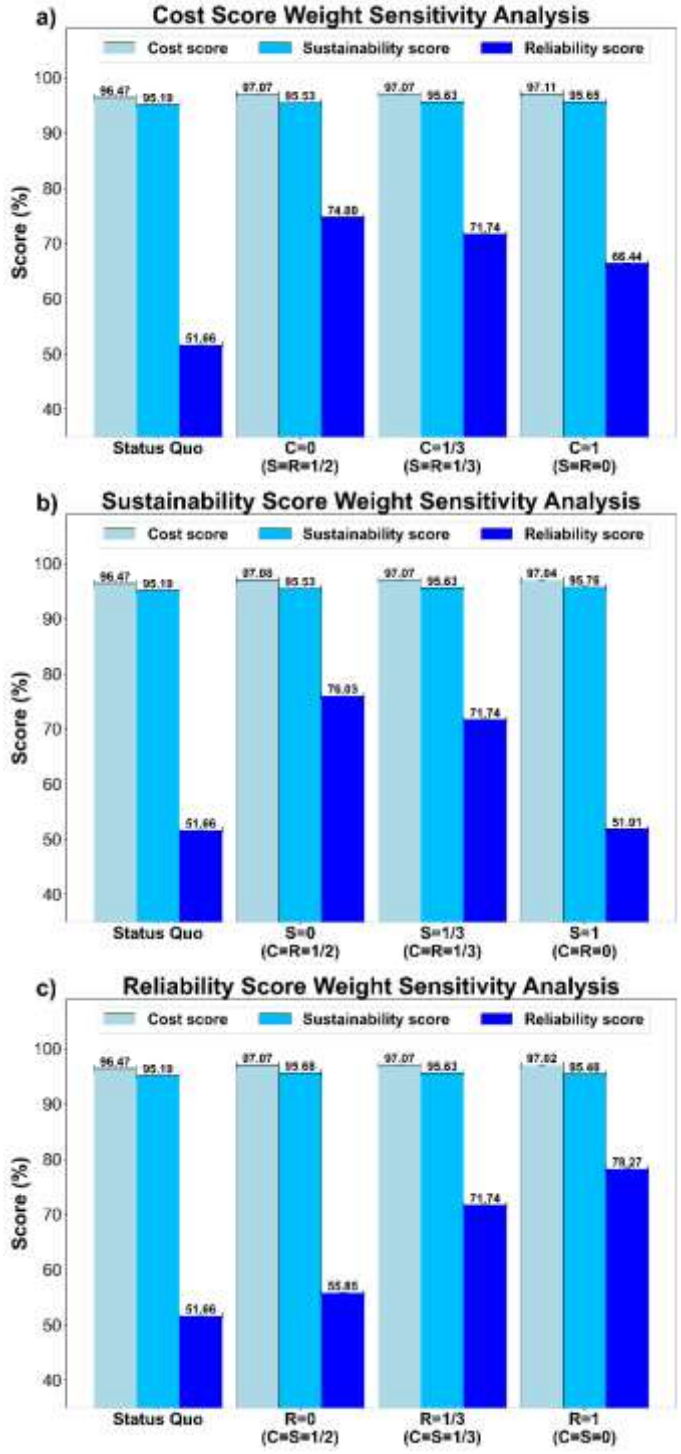

Figure 7. Sensitivity analysis results showing the influence of the weightings of the performance metrics on the RL outcomes.

5. Conclusions

This study investigated the cost, sustainability and reliability of an island energy-water microgrid through optimized scheduling of the water system to function as a virtual battery. A process-based dynamic model was developed to simulate the physical water and energy interactions, which was then combined with a reinforcement learning (RL) model for system optimization. Our findings reveal that the RL scenario outperformed the status quo, achieving improvements of approximately 0.60% in cost, 0.44% in sustainability, and a significant 20.08% in reliability, leading to a 7.04% increase in the overall weighted total score.

Additionally, the RL scenario demonstrated superior cost and reliability performances compared to a single-objective, heuristic scenario, albeit with a slight trade-off in sustainability. This highlights the inherent limitations of single-objective, use-phase-focused optimizations, which often fail to account for broader life cycle impacts, resulting in suboptimal solutions. A deeper analysis of model behaviors shows that the RL model makes strategic decisions by reserving battery storage for times of peak renewable energy generation later in the day. By operating the water system for extended periods, the RL approach effectively bypasses pumping capacity constraints, treating the water system as a virtual battery. These riskier strategies, which human operators might be reluctant to pursue, demonstrate how a well-calculated, risk-taking model can deliver superior system-wide benefits. Furthermore, our analysis suggests that increasing the desalination rate could lead to even greater improvements across all three performance metrics, whereas enlarging the water tank size yields only marginal gains.

The findings from this study highlight the critical role of optimized energy utilization, strategic planning, and tailored solutions in enhancing the efficiency, reliability, and sustainability of integrated energy-water microgrid systems. By demonstrating the long-term savings achievable through the optimal use of water as a virtual battery, these results can guide decision-makers in balancing trade-offs and benefits to improve system design, resource allocation, and emergency preparedness. The successful deployment of RL in this study offers a flexible model for similar microgrids in small islands or remote communities, promoting the adoption of AI-driven automation and fostering widespread system performance improvements. Future research could expand the RL model to incorporate energy system decisions and actions, such as diesel generator switching, alongside water system scheduling. This broader decision space would create a more comprehensive framework for optimizing the management of water-energy microgrids.

Acknowledgement

The authors would like to acknowledge the support of the US Economic Development Administration (EDA) through an Industry Challenge grant (ED20HDQ0200033) and the US National Science Foundation (NSF) under CBET award (#CBET-1706143). Any opinion, findings, conclusions, or recommendations expressed in the materials are those of the authors and do not necessarily reflect the view of the EDA or NSF. This work would not have been possible without the Shoals Marine Laboratory team. We gratefully acknowledge the work of Ross Hansen, Tyler Garzo, and David Buck in providing system information and data access. This is contribution # 207 to the Shoals Marine Laboratory.

References

[1] X. Zhou, T. Guo, and Y. Ma, "An overview on microgrid technology," *2015 IEEE Int. Conf. Mechatronics Autom. ICMA 2015*, pp. 76–81, Sep. 2015, doi: 10.1109/ICMA.2015.7237460.

[2] R. E. Giachetti, D. L. V. Bossuyt, W. W. Anderson, and G. Oriti, "Resilience and Cost Trade Space for Microgrids on Islands," *IEEE Syst. J.*, vol. 16, no. 3, pp. 3939–3949, Sep. 2022, doi: 10.1109/JSYST.2021.3103831.

[3] DOE, "The Role of Microgrids in Helping to Advance the Nation's Energy System," *Department of Energy*, 2022.

[4] DOE, "Combined Heat & Power and Microgrid Installation Databases," *U.S. Department of Energy*, Dec. 2022.

[5] DOE, "U.S. Department of Energy Combined Heat & Power and Microgrid Installation Databases," *U.S. Department of Energy*, 2024.

[6] R. Georgious, R. Refaat, J. Garcia, and A. A. Daoud, "Review on Energy Storage Systems in Microgrids," *Electron. 2021, Vol. 10, Page 2134*, vol. 10, no. 17, p. 2134, Sep. 2021, doi: 10.3390/ELECTRONICS10172134.

[7] M. Lee, J. Park, S. I. Na, H. S. Choi, B. S. Bu, and J. Kim, "An Analysis of Battery Degradation in the Integrated Energy Storage System with Solar Photovoltaic Generation," *Electron. 2020, Vol. 9, Page 701*, vol. 9, no. 4, p. 701, Apr. 2020, doi: 10.3390/ELECTRONICS9040701.

[8] S. P. Nandanoori, I. Chakraborty, T. Ramachandran, and S. Kundu, "Identification and Validation of Virtual Battery Model for Heterogeneous Devices," *IEEE Power Energy Soc. Gen. Meet.*, vol. 2019-August, Aug. 2019, doi: 10.1109/PESGM40551.2019.8973978.

[9] D. Manolakos, G. Papadakis, D. Papantonis, and S. Kyritsis, "A stand-alone photovoltaic power system for remote villages using pumped water energy storage," *Energy*, vol. 29, no. 1, pp. 57–69, Jan. 2004, doi: 10.1016/J.ENERGY.2003.08.008.

[10] M. T. Mito, X. Ma, H. Albuflasa, and P. A. Davies, "Modular operation of renewable energy-driven reverse osmosis using neural networks for wind speed prediction and scheduling," *Desalination*, vol. 567, p. 116950, Dec. 2023, doi: 10.1016/J.DESAL.2023.116950.

[11] N. Rezaei, Y. Pezhmani, and R. P. Mohammadiani, "Optimal stochastic self-scheduling of a water-energy virtual power plant considering data clustering and multiple storage systems," *J. Energy Storage*, vol. 65, p. 107366, Aug. 2023, doi: 10.1016/J.EST.2023.107366.

[12] B. Zhou *et al.*, "Multi-time scale optimal scheduling model for active distribution grid with desalination loads considering uncertainty of demand response," *Desalination*, vol. 517, p. 115262, Dec. 2021, doi: 10.1016/J.DESAL.2021.115262.

[13] Q. Guo, T. Guo, Q. Tian, and S. Nojavan, "Optimal robust scheduling of energy-water nexus system using robust optimization technique," *Comput. Chem. Eng.*, vol. 155, p. 107542, Dec. 2021, doi: 10.1016/J.COMPCHEMENG.2021.107542.

[14] S. A. Putri, F. Moazeni, and J. Khazaei, "Predictive control of interlinked water-energy microgrids," *Appl. Energy*, vol. 347, p. 121455, Oct. 2023, doi: 10.1016/J.APENERGY.2023.121455.

[15] S. Xiao, Q. Guan, W. Zhang, and L. Wu, "Optimal scheduling of the combined power and desalination system," *Energy Reports*, vol. 8, pp. 661–669, Nov. 2022, doi: 10.1016/J.EGYR.2021.11.253.

[16] I. Karakitsios, A. Dimeas, and N. Hatziargyriou, "Optimal Management of the Desalination System Demand in Non-Interconnected Islands," *Energies 2020, Vol. 13, Page 4021*, vol. 13, no. 15, p. 4021, Aug. 2020, doi: 10.3390/EN13154021.

[17] M. Elsir, A. S. Al-Sumaiti, M. S. El Moursi, and A. T. Al-Awami, "Coordinating the

day-ahead operation scheduling for demand response and water desalination plants in smart grid," *Appl. Energy*, vol. 335, p. 120770, Apr. 2023, doi: 10.1016/J.APENERGY.2023.120770.
[18] F. Moazeni, J. Khazaei, and A. Asrari, "Step towards Energy-Water Smart Microgrids; Buildings Thermal Energy and Water Demand Management Embedded in Economic Dispatch," *IEEE Trans. Smart Grid*, vol. 12, no. 5, pp. 3680–3691, Sep. 2021, doi: 10.1109/TSG.2021.3068053.
[19] F. Moazeni and J. Khazaei, "Optimal design and operation of an islanded water-energy network including a combined electrodialysis-reverse osmosis desalination unit," *Renew. Energy*, vol. 167, pp. 395–408, Apr. 2021, doi: 10.1016/J.RENENE.2020.11.097.
[20] J. Xu, H. Mahmood, H. Xiao, E. Anderlini, and M. Abusara, "Electric Water Heaters Management via Reinforcement Learning with Time-Delay in Isolated Microgrids," *IEEE Access*, vol. 9, pp. 132569–132579, 2021, doi: 10.1109/ACCESS.2021.3112817.
[21] T. Wu and J. Wang, "Artificial intelligence for operation and control: The case of microgrids," *Electr. J.*, vol. 34, no. 1, p. 106890, Jan. 2021, doi: 10.1016/J.TEJ.2020.106890.
[22] M. L. T. Zulu, R. P. Carpanen, and R. Tiako, "A Comprehensive Review: Study of Artificial Intelligence Optimization Technique Applications in a Hybrid Microgrid at Times of Fault Outbreaks," *Energies 2023, Vol. 16, Page 1786*, vol. 16, no. 4, p. 1786, Feb. 2023, doi: 10.3390/EN16041786.
[23] L. Lv, Z. Wu, L. Zhang, B. B. Gupta, and Z. Tian, "An Edge-AI Based Forecasting Approach for Improving Smart Microgrid Efficiency," *IEEE Trans. Ind. Informatics*, vol. 18, no. 11, pp. 7946–7954, Nov. 2022, doi: 10.1109/TII.2022.3163137.
[24] A. Joshi, S. Capezza, A. Alhaji, and M. Y. Chow, "Survey on AI and Machine Learning Techniques for Microgrid Energy Management Systems," *IEEE/CAA J. Autom. Sin.*, vol. 10, no. 7, pp. 1513–1529, Jul. 2023, doi: 10.1109/JAS.2023.123657.
[25] E. Mohammadi, M. Alizadeh, M. Asgarimoghaddam, X. Wang, and M. G. Simoes, "A Review on Application of Artificial Intelligence Techniques in Microgrids," *IEEE J. Emerg. Sel. Top. Ind. Electron.*, vol. 3, no. 4, pp. 878–890, Aug. 2022, doi: 10.1109/JESTIE.2022.3198504.
[26] L. P. Kaelbling, M. L. Littman, and A. W. Moore, "Reinforcement Learning: A Survey," *J. Artif. Intell. Res.*, vol. 4, pp. 237–285, May 1996, doi: 10.1613/JAIR.301.
[27] P. Kofinas, G. Vouros, and A. I. Dounis, "Energy management in solar microgrid via reinforcement learning," *ACM Int. Conf. Proceeding Ser.*, vol. 18-20-May-2016, May 2016, doi: 10.1145/2903220.2903257.
[28] SML, "Appledore Island," 2019.
[29] R. Ghasemi, M. Wosnik, D. L. Foster, and W. Mo, "Multi-Objective Decision-Making for an Island Microgrid in the Gulf of Maine," *Sustain.*, vol. 15, no. 18, p. 13900, Sep. 2023, doi: 10.3390/SU151813900/S1.
[30] Lifestream, "Lifestream Watersystems - SW6," 2016.
[31] M. Ren, C. R. Mitchell, and W. Mo, "Dynamic life cycle economic and environmental assessment of residential solar photovoltaic systems," *Sci. Total Environ.*, vol. 722, Jun. 2020, doi: 10.1016/j.scitotenv.2020.137932.
[32] P. Lilienthal, "HOMER® micropower optimization model," National Renewable Energy Lab.(NREL), Golden, CO (United States), 2005.
[33] B. L. Elzweig and Z. Robarts, "Sustainable Engineering Internship 2021," 2021.
[34] L. Balkin, A. D'Orlando, S. Jakositz, and E. Khanna, "Sustainable Engineering Internship 2017," 2017.
[35] C. de Faro, "Economic life of equipments and depreciation policies," 1998.
[36] R. Sutton and A. Barto, *Reinforcement learning: An introduction*. 2018.

[37] X. Luo, X. Chang, and X. Ban, “Regression and classification using extreme learning machine based on L1-norm and L2-norm,” *Neurocomputing*, vol. 174, pp. 179–186, Jan. 2016, doi: 10.1016/J.NEUCOM.2015.03.112.