### **Decarbonizing Water Desalination by Optimizing Renewable Energy using Artificial Intelligence**



The increasing intensity and frequency of water scarcity, carbon emissions, and climate risks pose critical challenges necessitating increased uptake of and a paradigm shift to energyand climate-smart water desalination processes. Desalination, particularly through reverse osmosis (RO), is an energy-intensive process, predominantly (~99%) reliant on non-renewable fossil fuels. Additionally, the high energy demands and costs of conventional desalination methods pose economic challenges. This study employs metrics and a decision framework to enable and accelerate the energy efficiency, decarbonization, and cost-effectiveness of RO water desalination processes. Integration and optimization of advanced artificial intelligence (AI) algorithms, hybrid renewable energy (RE) sources, and storage systems are utilized to explore tradeoffs and identify paths to maximize energy efficiency, while minimizing cost, energy inputs, and carbon footprint (by at least 50%) using multi-source, multi-metric data, modeling and analysis of four (4) U.S. water desalination plants (Alameda County, Tampa Bay, San Antonio-SAWS, Kay-Bailey Hutchison). This robust data- and analytical-driven approach provides the necessary tools to desalination plants for a sustainable transition to 100% RE, and to local governments and utilities to expedite the approval processes and provide continued rebates and innovative financing programs.

An essential step was analyzing various RE sources, such as photovoltaic panels, wind turbines, concentrated solar power, geothermal units, and hydro turbines; in addition, we examine battery storage systems to address the intermittency challenges associated with solar and wind energy. More specifically, we collected and sanitized information on top-of-the-line RE systems, including formulas for power generation, cost, carbon emissions, and specs. The feasibility of these diverse RE systems was assessed using 5 years of actual operational data from desalination plants including water production, energy consumption, water quality, and 10 year historical hourly weather data, establishing optimization functions and constraints.

In this research, to obtain optimal sizing of each RE system to satisfy the desalination plants' energy demand, methods including predictive modeling for water production, energy consumption, and long-term weather forecasting have been used. The plant data is preprocessed to account for variables like weather changes, plant shutdowns, and maintenance using statistical and AI tools, which form the core of the approach. Various machine learning models, including SARIMA, Random Forest, XGBoost, and Gradient Boosting, are employed for forecasting. The research also estimates the energy production potential of each RE source, factoring in surplus or deficiency management through battery storage and utility grid interconnection. This strategy aligns RE generation with the anticipated energy demands of the plants, facilitating a comprehensive optimization strategy that includes sizing and load forecasting of water desalination plants through AI models.

The results show that among AI models, Gradient Boosting and an innovative average method of XGBoost have the best accuracy. The Root Mean Square Error (RMSE) for weather prediction varied across different variables and locations. For temperature, with a range between -10.0-46.1ºC, the RMSE values were 3.37ºC for Alameda, 4.67ºC for SAWS, 3.47ºC for Tampa, and 3.62ºC for Kay-Bailey. In terms of Direct Normal Irradiance (DNI), measured within a range of 0-1,049 W/m<sup>2</sup>, the RMSE values were 189 W/m<sup>2</sup> for Alameda, 234 W/m<sup>2</sup> for SAWS, 199 W/m<sup>2</sup> for Tampa, and 204 W/m<sup>2</sup> for Kay Bailey. For wind speed, spanning from 0-35.2 m/s, the RMSE values recorded were 1.72 m/s for Alameda, 2.48 m/s for SAWS, 2.79 m/s for Tampa, and 2.97 m/s for Kay-Bailey. For treated water flows, which range from 0-30 MGD, Alameda's RF model showed 1.49 MGD, SAWS's GB model 0.86 MGD, Tampa's GB model 1.17 MGD, and Kay Bailey's GB model 0.39 MGD. Finally, for energy consumption, Alameda's RF model recorded an RMSE of 59,000 kWh (range of 0-1,000,000 kWh), SAWS's RF model 711 kWh (range of 0-300,000 kWh), Tampa's GB model 6,751 kWh (same range), while Kay Bailey's GB model 10,185 kWh (range of 0-83,333 kWh).

By harnessing AI and RE, a scalable, sustainable solution to water scarcity and carbon emissions challenges are explored in terms of costs when using RE mix of 50%, 75% and 100%. At a 50% RE mix, it could halve the current annual CO2 output, equating to a reduction of 16 billion gallons or 16,000 metric tons of CO2, akin to the impact of planting 600,000 trees and a significant long-term cost reduction at the four plants. The methods explored and results described contribute to environmental conservation, and enhance economic sustainability in water management, highlighting plausible transitions and decision-making considerations for future desalination.



# **EUROMED 2024 Desalination for Clean Water and Energy**  $-6 - 9$  May 2024 Four Seasons Hotel, Sharm El Sheikh, Eg

# Decarbonizing Water Desalination by Optimizing Renewable Energy using Artificial Intelligence

Om Sanan, Joshua Sperling, David Greene, Ross Greer





# Acknowledgement







**Dr. Joshua Sperling** National Renewable Energy Laboratory Manager – New Concepts Incubator, Urban Futures and the Energy-X Nexus Colorado

**David Greene** National Renewable Energy Laboratory Water, Energy, Climate, and Sustainability Engineer and Program Manager Colorado

**Ross Greer** UC – San Diego PhD Candidate - Electrical & Computer Engineering California



# Introduction

- Clean Water and AI Enthusiast and Researcher, Scarsdale High School Junior
- Intern, National Renewable Energy Labs (since Fall 2022) and Millwood WTP (since April 2021)
- Founder and CEO of Day Zero Water (non-profit); founded April 2018



Millwood WTP (GE Suez) in NY, 5<sup>th</sup> Grade Millwood WTP (Veolia) in NY, 11<sup>th</sup> Grade







# Vision

- **Vision** 
	- Resilient Water Supply Infrastructure Water Security and Energy Independence
	- Achieve Decarbonization and Reduce Cost in Water Desalination through Renewable Energy (RE) and Battery Storage proliferation
	- AI will help harbor the next wave of innovation in eco-friendly and efficient Desalination

## **Uniqueness**

- Creation of an automated Energy Management System that can:
	- Process significant amount of input data
	- Can automatically obtain data from users or APIs and run various modeling algorithms
	- Enables efficient decarbonization by using ML forecasting models & optimization algorithm to optimally align RE generation with energy demand on an hourly basis



# **Background**

### **Global Water Scarcity**

- According to U.N., 2 billion or 26% people lack access to clean water, which is expected to double by 2050
- Global Warming and Population growth are exacerbating the issue resulting in diminishing fresh water supply

### **Renewable Energy & Battery Storage**

- Abundant RE source; can replace 100% electricity
- RE cost has reduced ~80% over the last decade
- Currently powers only 1% of Desalination
- High initial CAPEX & Intermittent nature of RE
- Battery storage: Facilitates the efficient utilization of RE, enabling grid stabilization, peak shaving, and load shifting

## **Problem Statement**

- Desalination harnesses an abundant source of water, however, it is:
	- Expensive; energy cost  $= 40\%$
	- Results in significant carbon emission

## **Gap in Research / Problem Statement**

- Current studies use data that is either simulated or on a small scale
- Lack comprehensive assessment of multiple RE sources and optimization



E



# Research Overview

- **Objective:** Obtain the optimal system sizing of hybrid RE sources and battery storage using AI algorithms to decarbonize desalination
	- Multi-objective of minimizing cost and carbon footprint (50-100% RE)
	- Protects against energy and price volatility (energy independence)
- **Real-life, mid-large size desalination plants** are used to demonstrate effective decarbonization on a large scale
- **Systems Evaluated:** PV, CSP, Wind Energy, Hydro Energy, Battery Storage
- Methods include: (i) Sophisticated ML models to forecast water demand, energy consumption, and weather/environmental data; and (ii) Optimization algorithms to design an optimal RE and battery mix



# Summary of Four Desalination Facilities



**The 4 U.S desalination plants in the study use ~75 million kWh of electricity, which if converted to 100% RE can result in annual carbon emissions reduction of 32,000 metric tons, akin to benefit derived from planting 1.2 million trees** 

E

# Systematic Framework / Methodology





EUROMED 2024

C.

 $\equiv$ 

- 1)  $P_{\text{EC}}$  is the energy demand of desalination plant.
- 2)  $P_{\text{pv}} P_{\text{csp}} P_{\text{w}} P_{\text{hydro}} P_{\text{geo}}$  represent renewable energy generation from PV, CSP, WT, HT and GE, respectively.<br>3) N.,, N.,,, N.,,, N.,,,, N.,,,, N.,,,, N.,,,, N.,,, represent the number of PV arrays. CSP units
- $N_{\rm op}$   $N_{\rm cso}$ ,  $N_{\rm w0}$ ,  $N_{\rm had}$ ,  $N_{\rm but}$  N<sub>inv</sub> represent the number of PV arrays, CSP units, wind turbines, hydro turbines, geothermal power plants, battery banks and inverters, respectively.

# Data Collection

## ● **Desalination Plant Data**

• Hourly/Daily/Monthly 5-year info on energy use, water flow, and quality metrics such as TDS, Turbidity, temperature, pH

## ● **Weather / Environmental Data**

● Hourly 20-year metrics including solar irradiance, wind speed, temperature, streamflow rate and hydraulic head from National Solar Radiation Database (NSRDB) and USGS WaterWatch

## ● **Levelized Cost of Energy (LCOE) & Carbon Emissions of RE**

● NREL Annual Technology Baseline (ATB)

## ● **High-end RE System Specs**

• From various leading vendors of RE systems



厚

# Sample of Data Collected





F

# RE & Battery Formulas; LCOE &  $CO<sub>2</sub>$  Emissions

F

**EURO** 





# RE and Battery Assumptions



F

# ML Modeling – Forecasting Water Demand, Energy Consumption and Weather/Environmental Data

- **Forecast of Daily Treated Water Flows and Energy Consumption** 
	- Data Used: Rainfall, Hours in Operation, Raw Water Flows, Backwash Water Usage, Peak Demand, Avg Turbidity, Max Turbidity.
	- **Data Forecast: Water Flows, Energy Consumption.**
	- **Models Used: SARIMA, Random Forest, XGBoost, Gradient Boosting, Ensemble.**
- **Forecast of Weather Data (Climate & Hydro)**
	- Why: To determine RE generation potential
	- Data Used: Temperature, Solar Irradiance, Humidity, Wind Speed, Discharge
	- **Data Forecast: Solar Irradiance, Wind Speed, Temperature, Discharge**
	- **Models Used: XGBoost, LSTM, Linear/Polynomeal Regression.**



厚

# AI Modeling – Forecasting Weather Data and Energy Consumption/Water Flows

- Daily or monthly data sourced from desalination plants
- **Preprocessing and Feature Engineering:** Log transformations, lag creation, train/test split
- **Model Training**

Ę

- Regression based models: Random Forest, Gradient Boosting, SARIMA (seasonal)
- Hyperparameter tuning through GridSearchCV (cross-validation) tuned *n\_estimators, learning rate, max\_depth, min\_samples split, min\_samples\_leaf, bootstrap, max\_features, subsample*
- **Model Forecast**
	- Used each model to obtain test set predictions
	- Ensemble approach averages model predictions to mitigate individual model biases
- **Evaluation (for Tampa)**
	- Energy Consumption (range 0-300,000 kWh) RMSE (GB) = 3,853 kWh; MSE = 1,107 kWh
	- Treated Water (range  $0-10$  MGD) RMSE (RF) = 1.26 MGD; MSE = 0.85 MGD
- **Further Forecast and Application**
- Used tuned Gradient Boosting model for projecting 5-year daily treated water output and **EUROMED 2024 energy consumption** 15

## Forecast Treated Water & Energy Consumption (Tampa Bay)

### Total Treated Water (MGD) Daily Energy Consumption (kWh)





Ę

# AI Model Pipeline - Hydro Data

- 5-year 15-minute spaced readings for Discharge (or Streamflow Rate) and Hydraulic Head
- **Feature Engineering:** Trend Analysis (moving averages), Seasonality Detection, Lag Features
- Lots of visuals done to analyze data trends and find the best modeling approach.
- **Hybrid Modeling Approach**
	- Linear/Polynomial Regression for Trend and Seasonality
	- XGBoost for Residuals: Used XGBoost on residuals from the linear regression model.
	- Combination of Forecasts: Combined predictions from both models to generate final forecasts
- **Model Evaluation**
	- Discharge (range 0-2000 m<sup>3</sup>/s): RMSE = 47.76 m<sup>3</sup>/s, MAE = 17.07 m<sup>3</sup>/s
- **Forecasting:** Forecast future values for the next five years using the trained hybrid model
- **Incorporating Spikes:** Analyzed data for spikes and incorporated this information into the forecast, aiming to improve the model's forecast accuracy for outliers



厚

## Forecast Streamflow (Tampa Bay - Bullfrog Creek)

Streamflow  $(m^3/s)$ 





F

# AI Model Pipeline - Climate Data

- Integrated a complex forecasting process into a Streamlit application
- **Data Preparation and Preprocessing** 
	- Fetched 20 year hourly climate data using NSRDB API years 1998 to 2021
	- Train-Test Split (Train: 1998-2016; Test: 2017-2021)
- **Feature Engineering:** Lag Features, Rolling Window Statistics, Differencing
- **Model Training and Validation**

Ę

- Parameter Tuning: Used random search to tune n\_estimators, learning\_rate, max\_depth, min\_child\_weight, gamma, subsample, colsample\_bytree, reg\_alpha, reg\_lambda
- XGBoost: Employs gradient boosting decision trees for predictive accuracy in complex datasets
- Prophet: Handles daily observations with strong seasonal effects, ideal for time-series forecasting
- Average Method: Simple approach, forecasts future values as the average of past observations
- Autoregressive Validation: Utilizes past data points as inputs to forecast future values
- **Forecasting: Iterative Forecast + Optional Noise Injection**
- **Evaluation** (Average Model for Tampa)
	- Wind Speed (range 0-30 m/s): RMSE =  $2.79$  m/s, MAE =  $2.19$  m/s
	- Temperature (range  $0-35$  °C): RMSE =  $3.47$ °C, MAE =  $2.39$ °C

**EUROMED 2024** GHI (range 0-1200 W/m<sup>2</sup>): RMSE = 198.64 W/m<sup>2</sup>, MAE = 110.82 W/m<sup>2</sup>

# Forecast GHI, Temperature, Wind Speed (Tampa Bay)

• XGBoost Average Method Model

F

• Test data used to validate the accuracy of the training model



# Results: RE Generation Forecast & Error Rates

• Forecast future years' temperature, wind speed and GHI (solar irradiance) were used to calculate renewable energy generation





# RE Gen. by Source & Power Consumption (Annual)

Assumes 30,000 PV Modules (Blue) or 10 Wind Turbines (Orange) or 40 units/acres of CSP (Green) to satisfy/oversize energy needs of desalination plant (Red)





This chart represents the projected annual aggregate values of each renewable energy system, each scaled by an estimaed multiplier to match the forecasted energy demands of each plant (PF).

# RE Gen. by Source & Power Consumption (Monthly)

PV and CSP produce more consistent energy throughout the year, followed by wind energy which is lower in the summer months





 $\!=$ 

This chart represents the projected aggregate values of the average monthly sum of energy genration/demand (from 4/1/23 to 3/31/26) for each renewable energy system, each scaled by an estimated multiplier to match the forecasted energy demands of each plant (PF).

# RE Gen. by Source & Power Consumption (Hourly)

During the day, PV and CSP have the greatest volatility, followed by wind energy\*





Ę

This chart represents the aggregate values of the average hourly sum every day of energy genration/demand (data spanning from 4/1/23 to 3/31/26) for each renewable energy system, each scaled by an estimaed multiplier to match the forecasted energy demands of each plant (PF).

# Optimization Algorithms – RE & Battery Systems

### **Scope**

F

- Develop a Streamlit or Flask interface for renewable energy system sizing for desalination
- Focus on Tampa site initially with future expansion in mind
- Output optimal units of PV, wind, CSP, hydro, battery banks, inverters to minimize costs and emissions

### **Methodology**

- Minimize multi-objective loss function (L1, L2) while subject to a set of constraints in order to find optimal values for a set of decision variables -> aims to be practical and flexible for different sites
- Use ML models for accurate forecasting in desalination processes
- Evaluate photovoltaic, wind, CSP, hydro for cost and environmental efficiency
- Optimize using algorithms (NN, PSO-GWO) based on forecasted energy data

### **Technical Requirements**

- Enable selection among renewable sources for user-specific optimization
- Integrate battery storage for energy management
- Allow users to set constraints like renewable energy percentage, storage usage, and space limitations

### **Optimization Details**

- Objective: Efficient system sizing to meet desalination power needs
- Focus on minimizing LCOE and CO2 emissions
- Include constraints on renewable mix, battery depth of discharge, and plant capacity



# Scenarios and Optimization Constraints

## **Scenarios**

厚

- 1. Min 50% RE (max energy purchases from the utility = 49.9%) and battery
- 2. 100% hybrid RE (no energy purchases from utility)

## **Constraints**

1. 
$$
N_{pv}
$$
,  $N_{w}$ ,  $N_{csp}$ ,  $N_{hyd}$ ,  $N_{geo}$ ,  $N_{batt}$ ,  $N_{inv} \ge 0$ 

2. Renewable Fraction  $(RF) \ge 1.0$  or 0.5 (based on scenario 1 or 2)

 $RF = (P_{PV} + P_{csp} + P_{w} + P_{hvd} + P_{geo}) / P_{FC}$ 

- 3. DEPTH OF DISCHARGE (DOD) OF BATTERY STORAGE = 0.8
- 4. RO PLANT CAPACITY LIMITS

 $P_{\text{FCmin}} \leq P_{\text{FC}} \leq P_{\text{FCmax}}$ 





# **Limitations**

- Forecasted water/energy/weather data will not be 100% accurate
- We did not account for plant downtime in model training for energy consumption and water flows
- This study takes into account PV, Wind, CSP, Hydro Energy and Battery Storage; future Studies can include Geothermal Energy and Hydrogen Storage



# **Conclusion**

- **Key Achievements**
	- Established a robust framework and formulas for optimizing renewable energy integration in desalination
	- Water production, energy consumption, and weather forecasting using AI models to calculate RE generation, and optimization algorithims for optimal RE mix
	- Utilized real-world data to offer tangible insights and actionable recommendations

## ● **Future Study**

- Improve accuracy of AI models to forecast weather, energy consumption, and water flows
- Perform sensitivity analysis for optimization mix of RE system and battery storage
- Automate our Energy Management system so it can be implemented at any desalination plant or other industries by inputing some basic information



 $\equiv$ 



# Thank You!

