Operational optimization and control in the water and wastewater sector for energy demand management through advancements in data analysis and visualization.

By

ERIN NICOLE MUSABANDESU

DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Civil and Environmental Engineering

in the

OFFICE OF GRADUATE STUDIES

of the

UNIVERSITY OF CALIFORNIA

DAVIS

Approved:

Frank Loge, Chair

Jay Lund

Jonathan Herman

Committee in Charge

2022

Abstract:

Shifting energy consumption into times when renewable energy is abundant can increase renewable integration and reduce greenhouse gas emissions. As large energy users with flexibility in their operations, water and wastewater utilities can participate in this type of energy demand management. However, the primary goal of the water sector is to provide clean, safe, and adequate drinking water and protect human health and the environment by treating wastewater to appropriate standards. Water or wastewater treatment and delivery systems are both complex, with their own operational challenges; optimizing energy use must not interfere with the system's primary goals. In this dissertation, I present research focused on exploring techniques that allow water and wastewater utilities to perform energy demand management while still prioritizing key operation performance metrics such as water quality or system reliability. I also leverage data analytics, machine learning, and visualization to allow users, including water and wastewater utility optimizers, to better characterize and improve optimization problem formulations and methods.

The first investigation focuses on testing load shifting strategies at a full-scale wastewater treatment plant to participate as a demand resource. During these test events, the facility shifted energy load by modifying select operations with little to no impact on water quality. A cost-benefit analysis showed that the facility achieved cost savings of up to 4.8% by participating in the proxy demand response program, which allows users to bid on the California Wholesale energy market. From this case study, we identified two main barriers to wastewater utilities participating as a demand resource: the difficulty of correctly timing energy reductions to demand response periods and the inability of the standard baseline methodology to measure demand reduction for this complex system.

In the second research effort, we present a new optimization problem formulation for the water distribution system pump optimization problem: secondary time-based controls. This proposed

pump control uses a hierarchal structure to allow utilities to prioritize maintaining sufficient water reserves while responding to time-based energy incentives. The formulation was tested on three case studies and compared to two baseline pump control decision variable representations. A Monte Carlo sensitivity analysis was also conducted to determine the robustness of the control structures to uncertain water demands. Results show that this formulation similarly or further reduces energy costs than the two benchmark decision variable formulations without reducing average storage or violating operational constraints. In addition, the secondary time-based controls more consistently maintain water reserves and prevent constraint violations with uncertain demands, allowing water utilities to more comfortably manage energy demand to support renewable energy growth.

In the final study, we developed a visual analytic framework that characterizes the optimization fitness landscape to help users improve the optimization problem formulation and search efficiency. We also present a corresponding optimization method that guides the optimization using similar interpretable machine learning methods. Both the framework and optimization method were designed for use in the water distribution pump operation problem, but many of the analytic and visualization techniques could be applied to a broad range of complex optimization applications. We used the framework to examine the differences in fitness landscape between two different decision variable problem formulations on a benchmark water distribution system. We then tested the performance of several existing heuristic optimization methods either with or without guidance from the visual analytic framework and compared them to the proposed optimization method. The existing methods informed by the visual analytic framework and the new optimization method showed improvement over standard optimization.

Table of Contents

Acknowledgments

I would like to thank my advisor Frank who has inspired and guided me throughout this process. I will always be grateful for your encouragement and support.

To my partner and best friend, George, thank you for always being there, especially when I needed it the most. I would have struggled without you.

To Yiran, Takanori, and Professor Ma, thank you for helping me forge new paths. Our collaboration sparked innovation and added depth to my research.

To Jon, thank you for always being available to provide advice. Your insights and feedback helped me so much along the way.

To Robert, thank you for being a great sounding board, co-conspirator, and always having my back.

To Amanda and our million walks together, your friendship was a much-needed breath of fresh air.

To Liam, thanks for keeping me excited about research and all its potential.

To my family, you have been my biggest cheerleaders, friends, and a respite through it all. Thank you for your love and encouragement.

Thank you also to all of our collaborators and supporters at the Laguna Wastewater Treatment Plant in Santa Rosa and the Moulton Niguel Water District. Your commitment to expanding energy management in the water sector provided me with great opportunities to conduct applied research.

This research was possible thanks to financial support from the California Energy Commission and the California State Water Resources Control Board. The content of this research is solely the responsibility of the author and does not necessarily represent the official views of the supporting agencies.

1 Chapter 1: Introduction

With growing concerns over global warming and climate change, renewable energy generation is projected to increase significantly (International Energy Agency, 2019). However, rises in solar and wind generation have caused new institutional (Verzijlbergh et al., 2017) and technical (Liang, 2017) challenges for energy systems because these sources are intermittent, variable, and non-dispatchable. Integrating solar is especially challenging because its generation capacity decreases at the end of the day when energy demands increase (Badakhshan et al., 2019). As a compounding factor, large-scale energy storage is presently infeasible, so energy typically must be consumed shortly after its produced (Aneke and Wang, 2016).

One strategy that energy utilities can employ to mitigate these operational issues and increase the use of renewable energy is to incentivize customers to change the timing of their energy consumption using time-based pricing mechanisms (Palensky and Dietrich, 2011). Response to time-of-use (TOU) rates, demand response programs, or dynamic energy markets can smoothen the net load provided by dispatchable energy sources, alleviating stress on existing systems (Siano, 2014). Additionally, by shifting energy consumption into the times of day when renewable sources are abundant, renewable integration can increase, and greenhouse gas emissions can decrease (Paterakis et al., 2017).

As large energy users, the water sector has a significant opportunity to participate as energy demand resources through energy management. Water utilities' energy consumption accounts for approximately 4 % of global energy use (Kenway et al., 2019). Additionally, energy can account for 33 - 82 % of water and wastewater utilities' non-labor operating costs (Limaye and Jaywant, 2019). Participating in time-based pricing market mechanisms can both reduce costs for the facility and increase renewable integration for the energy sector. Water and wastewater systems also have several sources of operational flexibility that can be used to shift energy use, including water and

wastewater storage facilities (Lekov et al., 2009; Santhosh et al., 2014), excess system capacity (Oslen et al., 2012), and potential energy generation sources such as in-line turbines in water distribution systems (Williams, 1996) or cogeneration facilities at wastewater treatment plants (Schäfer et al., 2015).

At the same time, data is becoming increasingly available in both the water (Eggimann et al., 2017) and energy sectors (Zhou et al., 2016) with developments in sensors, increased instrumentation, and the expansion of the internet of things. Parallel advances in computing have increased opportunities to leverage this data to optimize water (Garrido-Baserba et al., 2020) and energy (Zhou et al., 2016) systems from both individual and interconnected water-energy nexus perspectives (Hamiche et al., 2016). Advances in visualization and interactive optimization can also aid in solving optimization problems for these types of complex multifactored systems (David et al., 2015).

Despite these opportunities and advancements, optimization and control of water and wastewater systems is challenging. These systems are often large and complex, with multifactored operational constraints and objectives. The primary purpose of water utilities is to provide safe and reliable drinking water, and wastewater treatment facilities must protect human health and the environment by ensuring their effluent meets the appropriate standards. Optimizing energy use cannot be at the expense of the principal goals of these systems.

In this dissertation, I present translational research focused on enabling water and wastewater utilities to perform energy demand management while still prioritizing key operation performance metrics such as water quality or system reliability. I also employ data analytics, machine learning, and visualization techniques to better characterize and improve optimization problem formulations and methods, specifically focused on optimizing water systems for energy demand management.

1.1 Dissertation Content and Structure

The remainder of this document is organized into the following chapters:

- **Chapter 2** investigates the use of several operational strategies to participate as a demand resource at a full-scale wastewater treatment plant through test demand response events, analyzing the impacts on water quality, and performing a cost-benefit analysis.
- **Chapter 3** presents a novel optimization problem formulation for the water distribution system pump optimization problem, secondary time-based controls, which allows utilities to maintain water reserves and operational system performance while responding to time-based energy incentives.
- **Chapter 4** expands on the work presented in Chapter 3 by developing a visual analytic framework designed to improve optimization problem formulation and search efficiency by characterizing the optimization fitness landscape, specifically focusing on the water distribution pump operation optimization problem. A corresponding optimization method that incorporates fitness landscape characterization techniques and machine learning into the search algorithm, the decision-tree guided genetic algorithm, is also presented.
- **Chapter 5** discusses the major contributions and conclusions of the research presented in this dissertation and provides suggestions for additional research to promote energy demand management in the water sector.

This introduction discusses the broader context which informs, motivates, and connects the individual research projects. Chapters 2-4 also contain their own abstract and introduction with additional context and background specific to each investigation.

2 Chapter 2: Load shifting at wastewater treatment plants: a case study for participating as an energy demand resource

2.1 Abstract

Energy load shifting can allow for increased renewable energy integration and reduced greenhouse gas intensity of the electricity grid. Recent research has demonstrated that wastewater treatment plants have considerable potential to shift energy loads and act as energy demand resources due to their energetic flexibility and energy production capacity. This paper investigates a wastewater treatment plant in Santa Rosa, California, participating as a demand resource on the wholesale energy market through the proxy demand resource program. Test demand response events showed that the facility was able to shift its energy load by modifying select operations without impacting wastewater effluent quality. A cost-benefit analysis based on projected program participation and the results from the test events, estimates that the Santa Rosa wastewater treatment plant could achieve up to 4.8% energy cost savings through the proxy demand resource program. Two main issues were identified from the test events: (1) the difficulty of correctly timing demand reduction periods and (2) the inaccuracy of using standard baseline methods to measure the energy load reduction. As a supplement to the case study, this paper also presents a roadmap outlining the technology necessary for wastewater treatment plants to participate in demand resource programs through energy load shifting. The roadmap identifies key instrumentation and automation infrastructure, and assets that can be utilized to provide energetic flexibility; it also recommends additional infrastructure that can stabilize energy loads and enhance controlled energy load shifting.

2.2 Introduction

Rising concerns over global warming and climate change have resulted in aggressive commitments to renewable energy integration in order to reduce the energy sector's greenhouse gas intensity. California has committed to 60 % renewable energy integration by 2030 and 100 % clean energy by 2045 (De León, 2018). Europe has similarly committed to 45 % renewable integration by 2030 (European Commission, 2014) and carbon neutrality by 2050 (European Commission, 2011).

However, renewable energy integration has caused new operational challenges for energy systems, particularly when integrating wind and solar due to their intermittent, variable, and non-dispatchable nature (Liang, 2017). These challenges are compounded by the fact that large scale energy storage is currently limited; most energy produced must be consumed immediately (Aneke and Wang, 2016). Renewable energy generation leaves increasingly substantial temporal gaps; in order to meet customer demand, energy providers must fill these gaps with dispatchable energy sources such as hydropower, geothermal, natural gas, and coal (Verzijlbergh et al., 2017).

Energy utilities help mitigate these operational issues by incentivizing customers to manage their energy use at particular time periods (Palensky and Dietrich, 2011). Customers can act as demand resources by reducing energy during the times of day when demand for dispatchable resources is the greatest or when the grid is most unbalanced (Siano, 2014). Furthermore, if energy users can shift energy loads into the times of day when renewable generation is readily available, they can increase renewable integration and reduce greenhouse gas emissions (Paterakis et al., 2017). Customers who are energetically flexible (i.e., can more significantly shift their energy load) are well suited to act as demand resources (Lund et al., 2015).

Energy utilities use several different time-based pricing mechanisms to promote demand-side management (Zhang and Li, 2012). Static time-of-use (TOU) rate structures based on historical

energy load patterns encourage customers to avoid using energy during historical peak usage hours. Demand response programs motivate customers by sending day-ahead requests to reduce load during either times of peak energy use or times that will improve the reliability of the grid. Some energy users can also access dynamic energy markets, either through energy utility programs or directly, which charge customers real-time energy prices based on the locational needs of the grid.

As large energy users, wastewater treatment plants (WWTPs) have a significant opportunity to participate as energy demand resources through energy management. In China, the United States, and Germany, WWTPs account for 0.3 %, 0.6 %, and 0.7 % of total energy consumption, respectively (Wang et al., 2016). In California, WWTPs make up nearly 0.8 % (GEI Consultants/Navigant Consulting Inc., 2010). Additionally, WWTPs have the potential to increase their energetic flexibility by cogenerating energy as a byproduct of their treatment processes (Schäfer et al., 2017). Depending on the size of the WWTP, energy accounts for approximately 25–40 % of operating costs (Panepinto et al., 2016). Participating in time-based pricing market mechanisms may help reduce these costs for the facility and increase renewable integration for the energy sector.

Wastewater treatment plants are complex systems that utilize physical, chemical, and biological treatment processes. Effluent water quality standards are met through primary, secondary, and optional advanced treatment. The energy requirements for primary treatment, which are largely driven by pumping, are low compared to secondary treatment, which must maintain microbial health often using energy-intensive aeration (Plappally and Lienhard, 2012). Advanced treatment processes, used for sensitive receiving waters, generally have a higher energy intensity than primary and secondary treatment processes (Wakeel et al., 2016). Energy consumption at wastewater plants is positively correlated pollutant loading (Gu et al., 2017). Both energy and loading vary seasonally and diurnally, with typically higher loads in the summer and daily peaks in

the late morning and early evening (Thompson et al., 2008). Previous research has explored modeling wastewater systems to predict effluent quality (Gernaey et al., 2004) and energy consumption (Mannina et al., 2019) based on influent parameters.

Energy research at WWTPs has typically focused on increasing energy efficiency or optimizing energy production (Maktabifard et al., 2018). However, recent investigations have begun evaluating operations based on time-based energy tariffs (Aymerich et al., 2015), as well as analyzing WWTPs' potential as energy demand resources (Seier and Schebek, 2017) and predicting the impacts of load shifting on treatment performance (Giberti et al., 2019). Several studies have identified the primary wastewater treatment processes that can provide energetic flexibility; Table 1 summarizes these processes and their related requirements and concerns.

Anaerobic digester operations are of particular interest for load shifting because of their potential for energy generation (Kirchem et al., 2018). The digester gas, supplemented with natural gas, can be used to generate electricity using combined heat and power engines (CHPs). WWTPs can shift CHP operations to generate energy during different periods, offsetting energy usage from the grid (Schäfer et al., 2015). Operators can also adjust the schedule of the heat pumps that maintain digester temperature. When heat pumps are off for less than 24 hours, insulated digester tanks lose no more than 1° C, with minimal impacts on final effluent quality (Schäfer et al., 2017). For sites where landfill leachate is being co-treated, there is the potential for additional energy generation using anaerobic digestion (Zairi et al., 2014) or microbial fuel cells (Greenman et al., 2009).

Aeration blower operations have also been targeted for load shifting due to their high energy use; however, operational shifts may degrade final water quality (Thompson et al., 2010). Facilities may be able to mitigate the effects of intermittent aeration on water quality with strict control parameters and monitoring oxygenation levels (Schäfer, 2019) or by over oxygenating wastewater

(Thompson et al., 2008). Sludge dewatering also consumes a large amount of energy. During dewatering centrifuge operation can be modified to perform load shifting; however, this may be restricted by site-specific requirements to operate in non-interruptible operation mode (Schäfer et al., 2017).

Table 1: Previous Research on Wastewater Treatment Process Energetic Flexibility

Treatment Process	**Studies**	**Requirements**	**Concerns**
Combined heat and power generation (CHP)	(Schäfer, 2019; Schäfer et al., 2017, 2015; Seier and Schebek, 2017)	Digester (anaerobic preferred to aerobic), gas storage, and CHP units.	-
Digester heat pump operations	(Schäfer, 2019; Schäfer et al., 2017)	Digester (anaerobic preferred to aerobic).	-
Transport Pump operations	(Aghajanzadeh et al., 2015; Oslen et al., 2012; Schäfer, 2019; Thompson et al., 2010)	Storage facilities to hold wastewater while pumps are not operating.	Further research is needed on the effects on effluent quality.
Aeration blower operations	(Aghajanzadeh et al., 2015; Giberti et al., 2019; Oslen et al., 2012; Schäfer, 2019; Thompson et al., 2010)	Blower equipment must allow for intermittent operations. Not all blowers operate non-constantly.	Intermittent blower operations can affect effluent quality. Researched and what controls or techniques might mitigate these effects.
Excess capacity	(Oslen et al., 2012)	Requires oversized facilities.	Further research is needed on the effects on effluent quality.
Centrifuge – sludge dewatering	(Oslen et al., 2012; Schäfer, 2019; Schäfer et al., 2017; Thompson et al., 2010)	Restricted by site-specific conditions.	-
Landfill leachate co-treatment	(Greenman et al., 2009; Zairi et al., 2014)	Digester, gas storage and CHP units or microbial fuel cell infrastructure	Research into generating energy using landfill leachate for microbial fuel cells is in its early stage.

If WWTPs have sufficient storage to retain wastewater temporarily, they can also use wastewater transport pumps for energy load shifting (Lekov et al., 2009). However, additional research is needed to understand the effects of intermittent pump operations on treatment efficacy (Kirchem et al., 2018). Excess facility capacity, in general, has been identified as an area that can

increase energetic flexibility. If additional capacity exists in either the facility or the sewer system before the wastewater enters the facility, operators can delay wastewater treatment operations and shift energy loads (Oslen et al., 2012). However, depending on the level of additional capacity and projected facility requirements, operators may hesitate to use excess capacity for this purpose (Aghajanzadeh et al., 2015).

A WWTP's level of automation can also significantly impact how effectively the facility can load shift operations (Lekov et al., 2009). Load shifting at manually operated facilities may require process control adjustments staff cannot adequately or easily provide.

Despite WWTPs' potential as demand resources, there are concerns that shifting operations could affect effluent quality (Burton and Stern, 1993). The primary goal of WWTPs is to treat wastewater to safe levels based on effluent use or discharge location; optimizing energy use must not interfere with the efficacy of treatment processes. Many of the biological treatment processes work best under steady conditions (LaGrega and Keenan, 1974). Therefore, definitive research is required for WWTPs to be willing to implement load shifting strategies that involve treatment processes. Research on full-scale WWTPs has been limited, with the majority of investigations analyzing flexibility by exploring existing operational data, focusing on the intermittent operation of pumps, blowers, and CHP units. Although several studies mention the potential to use excess capacity for load shifting, to the author's knowledge, no study has tested its implementation or impacts on wastewater quality. Furthermore, previous research has focused on the WWTP's ability to reduce energy and has not directly investigated participation in energy demand response programs. Additional research in this area can identify general technology and infrastructure requirements and determine any systematic barriers that current demand response programs may present to WWTP participation.

This study seeks to fill these gaps by testing three load shifting strategies with limited impact to wastewater treatment operations on a full-scale wastewater treatment system: (1) shifting the timing of CHPs for energy generation, (2) diverting flow equalization basins to reduce system-wide energy use, and (3) discharging an onsite battery. The primary objectives of this study are to use trial demand response events to examine how effectively the WWTP can participate in the demand response program, analyze the effects of flow diversion on wastewater operations and effluent quality, and use the trial event results to build a participation model and perform a cost-benefit analysis according to an existing demand response program. A supplementary goal of the study is to provide a roadmap for other WWTPs interested in participating as an energy demand resource.

2.3 Methods and Approach

This case study examines the ability of the Laguna Wastewater Treatment Plant, located in Santa Rosa, California, to shift energy loads based on price incentives from a demand response program.

2.3.1 The California Energy Landscape

The California Independent System Operator (CAISO) manages the statewide energy grid using a wholesale energy market, which promotes the purchase and sale of energy through locational marginal prices (LMPs); energy is bought and sold at real-time energy prices based on the locational value of energy. The three largest investor-owned utilities (IOUs) that participate in the CAISO market are Pacific Gas and Electric (PG&E), San Diego Gas and Electric, and Southern California Edison. Combined, these IOUs provide energy to nearly three-quarters of California customers. These IOUs, as well as the CAISO itself, have developed various time-based energy programs to promote energy demand response (Table 2). Each of these programs is designed to leverage specific customers and mitigate different operational issues of the grid. The programs

reduce energy through two methods: either the energy provider uses devices to automatically turn off energy or reduce its use or the users change their energy use to meet promised reductions.

Table 2: Demand Response Programs offered by the California Independent System Operator and the associated Investor Owned Utilities

Energy Provider	**Demand Response Programs**	**Control Type**
California Independent System Operator	Proxy Demand Resource	User-controlled
	Proxy Demand Resource - Load Shifting	User-controlled
	Reliability Demand Response Resource	User-controlled
Pacific Gas & Electric	Peak day Pricing	User-controlled
	Base Interruptible Program	User-controlled
	Capacity Bidding Program	User-controlled
	Optional Binding Mandatory Curtailment[1]	User-controlled
	Scheduled Load Reduction Program[1]	User-controlled
Southern California Edison	Agricultural and Pumping Interruptible	Automatic
	Automated Demand Response	Automatic
	Capacity Bidding Program	User-controlled
	Critical Peak Pricing	User-controlled
	Optional Binding Mandatory Curtailment	User-controlled
	Agricultural and Pumping Real-time Pricing	User-controlled
	Real-time Pricing	User-controlled
	Scheduled Load Reduction Program	User-controlled
	Summer Discount Plan for Businesses	Automatic
	Time-of-Use Base Interruptible Program[2]	User-controlled
San Diego Gas & Electricity	AC Saver (Summer Saver)	Automatic
	Base interruptible Program	User-Controlled
	Capacity Bidding Program	User-Controlled
	Smart Thermostats	Automatic

[1] The program is full and currently unavailable for enrollment
[2] The program only allows annual enrollment in April.

This study explores the Proxy Demand Resource (PDR) program managed by CAISO, a program that allows participants to bid into California's day-ahead and real-time wholesale energy markets, providing energy "supply" by curtailing their load by a promised amount. The PDR program requires a minimum bid of 0.1 megawatts (MW) to participate; smaller loads can be aggregated. Once all bids are placed, CAISO determines which bids to accept and notifies the participants. For each participation day, a moving average is calculated from the most recent non-participation days

to use as a baseline load. Once the PDR event has concluded, the user's metered data is compared to the calculated baseline to determine the amount of energy the user curtailed during the event. The final settlement amount is determined by multiplying the curtailed load by the LMPs for each 5-minute interval of CAISO's real-time wholesale energy market during the PDR event. Many other demand response programs evaluate energy reduction on a 15-minute or hourly interval. Note LMPs are used to determine revenue only when PDR participants supply energy; participants continue to pay their selected energy rate for consuming energy.

The Laguna Wastewater treatment plant previously participated in another demand response program: PG&E's Base Interruptible Program (BIP). This program requires participants to reduce their load below a pre-defined level when CAISO issues a day-of curtailment notice. PG&E provides a minimum of 30 minutes advance notice to participant and there are limits on the length and number of events that can be issued.

2.3.2 Case Study Site Description

The Laguna WWTP serves nearly 230,000 customers and treats an average daily flow of 66.2 megaliters (ML) or 17.5 million gallons (MG) for recycled water reuse. Its wastewater undergoes primary, secondary, and tertiary treatment (see Figure 1). These treatment processes all consume energy and are within the scope of this study. Primary treatment includes screens, grit removal, and primary clarification. Secondary treatment includes aeration and secondary clarification. The sludge collected in the primary and secondary clarifiers is then sent to anaerobic digesters. Methane is captured from the anaerobic digesters and used for energy cogeneration by the four CHP engines. After anaerobic digestion, the sludge is dewatered and composted for agricultural use. Tertiary treatment includes coal filtering and ultraviolet disinfection. The treated water, now classified as recycled, is either used to irrigate agricultural lands and urban landscaping or recharges the steamfield feeding the Geysers geothermal energy production facility. Water not

reused for these purposes is discharged to surface waters during the allowable discharge period: October 1 through May 14. All discharges are regulated by the National Pollutant Discharge Elimination System (NDPES) permit through the State Water Resources Control Board.

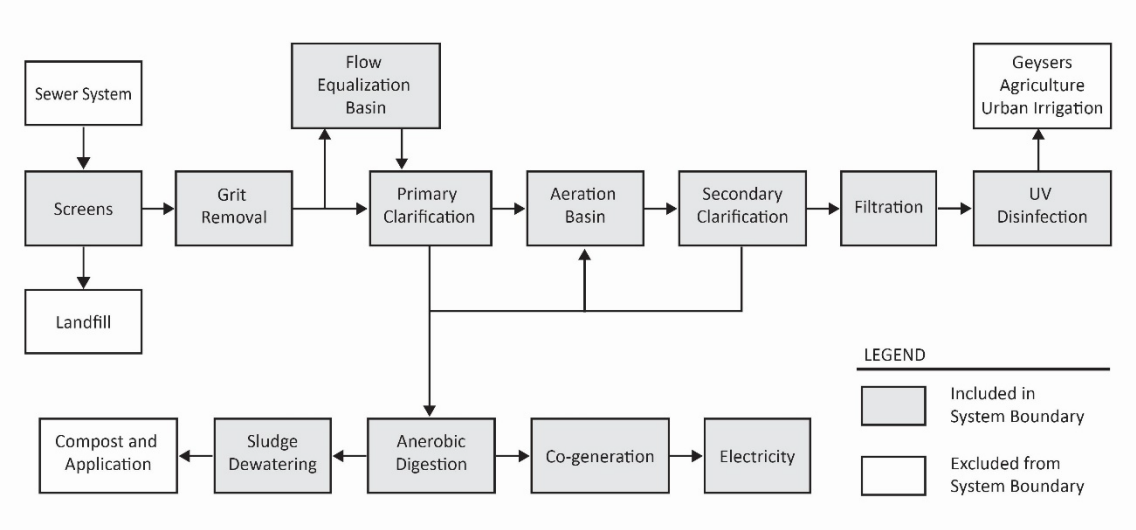

Figure 1. Laguna wastewater treatment plant treatment process diagram.

When the Laguna WWTP is not shifting energy loads, its energy consumption typically oscillates between 3 - 5 megawatt hours (MWh) throughout the day, based on the energy needs of the system. Each CHP engine cogenerates approximately 1 MWh of energy, reducing the WWTP's total energy requirements. PG&E provides the remaining energy to the Laguna WWTP from the California statewide grid.

The Laguna WWTP used both existing infrastructure and upgraded facilities to increase its load shifting capability and perform demand response (see Table 3). To reduce system-wide energy, the facility can divert wastewater flow, within certain operational limits, into two pre-existing 48.45 ML (12.8 MG) flow equalization basins (FEBs). Additionally, in 2018, the Laguna WWTP retrofitted two pre-existing CHPs with selective catalytic reduction (SCR) emissions control units, allowing up to four units to operate simultaneously. At previous emissions levels, state air regulations only allowed one unit to run at a time, mixing digester gas with natural gas. Now the two CHPs with SCRs operate using natural gas, and the remaining two CHPs are fueled by pure digester gas.

Table 3: Assets at Laguna Wastewater Treatment Plant That Facilitate Energy Load Shifting

Asset	Quantity	Make and Model	Size/Capacity
Additional Infrastructure			
Selective Catalytic Reduction Emissions Control Unit	2	Miratech[1]	NA
Battery	1	Tesla	2 MW
Outdoor Power Conversion System (Inverter)	1	Parker 890GTB-2206	2200 kVA (480 AC); 60 Hz
Existing Infrastructure			
Flow Equalization Basin	2	NA	48.45 ML (12.8 MG)
Combined Heat and Power Engines	2(of 4)[2]	Cummins QSK60G	Maximum power output: 1100 kW; Power Factor: 0.89

[1] Custom-built Selective Catalytic Reduction Unit by Miratech.
[2] Only the two engines fitted with SCR emissions control units are utilized for load shifting.

To further smoothen and control its energy load, the Laguna WWTP also installed several energy infrastructure assets onsite. It installed a 2-MW battery at its energy grid substation, allowing the WWTP to stabilize the energy load required by the electricity grid, and it installed an inverter to connect the battery and the grid, controlling the timing and rate at which the battery stores or releases energy. These assets facilitate rapid energy load reduction, easing the demand on the grid while maintaining the minimum energy required to sustain operations and adequate water quality.

2.3.3 Demand Response Test Events

From March to June 2019, the Laguna WWTP performed eighteen test events to determine how well it could participate as a demand resource in the PDR market. The facility conducted six trial runs for the following three operational modifications:

- **Scenario 1:** run one of the CHPs fitted with an SCR emissions control to generate energy from natural gas, reducing energy by an estimated 1 MWh.
- **Scenario 2:** divert flow to FEB to reduce system-wide energy use by an estimated 500 kWh.
- **Scenario 3:** discharge a 2-MW battery to offset energy use by approximately 200 kWh.

For scenario 2, additional energy is required to pump the diverted flow in and out of the FEB; however, this pumping energy use is minimal compared to the system-wide reduction provided by flow diversion. Wastewater treatment operators determined the extent and timing of flow diversion that would minimally impact effluent quality based on previous experience with the FEBs. Since CHP operation has no impact on wastewater treatment processes, there were no constraints. These two assets were the only operations that Laguna WWTP was comfortable altering for energy load shifting.

Supervisory control and data acquisition (SCADA) sensors and energy meters throughout the Laguna WWTP were used to collect time-series energy and flow data for the study duration. Water quality parameters were recorded, and water quality tests were performed to measure the constituents of concern (COCs) regulated by the NDPES Permit. Measurements were then compared to maximum contaminant levels (MCLs) to ensure that operational modifications were not impacting water quality.

Additionally, Laguna WWTP operators and program implementers were interviewed to identify opportunities and barriers to participation as a demand resource in the PDR program. A series of semi-structured telephonic and personal interviews were conducted during the planning and implementation phases and prior to the test events. The lessons learned from interviews and analysis of this case study were used to create a roadmap for other WWTPs to participate in similar demand resource programs.

2.3.3.1 Baseline Energy Calculations

To measure the energy load reduction for the demand response events, the energy load is compared to a calculated energy baseline. The PDR program uses the 10-in-10 baseline calculation methodology (CAISO, 2009). In this method, the calculated baseline, CBL, is determined for each

hour h, of the participation day d from the average energy consumption of the last n non-participation days in a 45-day window:

$$CBL_{d,h} = \frac{1}{n}\sum_{1}^{n} Energy\ Consumption_{d-n,h} \qquad (1)$$

If there are fewer than ten non-participation days, a minimum of five days may be used; if there are fewer than five days, the participation days with the highest load during a demand response event may be used. Separate baselines are calculated for weekdays and weekends/holidays

Additionally, on the day of participation, the raw CBL can be adjusted ±20 % using a same day adjustment multiplier:

$$Adjustment\ Multiplier = \frac{\sum Energy\ Consumption_{d,h}}{\sum CBL_{d,h}} \qquad (2)$$

The adjustment multiplier is calculated on the event day over the adjustment window: the four hours prior to the event, excluding the hour directly before demand response (CAISO, 2009).

Figure 2 provides an example of the adjusted baseline calculation for a single day. The figure includes the raw baseline, actual energy load, final adjusted baseline, adjusted baseline limitations, and time period that the adjusted baseline multiplier was calculated within.

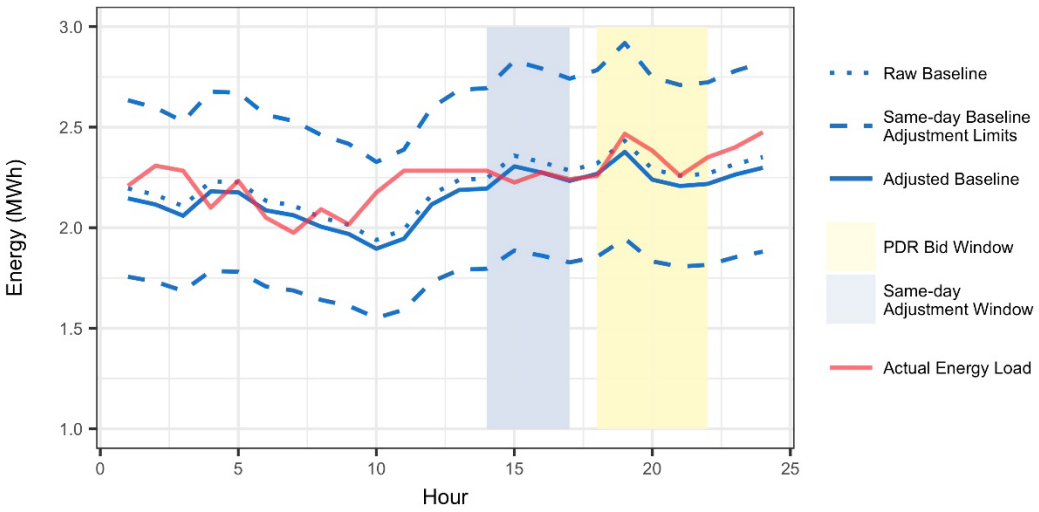

Figure 2. Example adjusted baseline calculation for a single day using the 10-in-10 baseline methodology.

2.3.4 Projected Cost-Benefit Analysis

A supplementary cost-benefit analysis was performed to estimate the potential revenue from participating in the PDR program. This analysis used historical five-minute interval energy load data from Laguna WWTP from May 2018 to October 2018 and corresponding CAISO real-time Locational Marginal Prices (CAISO, 2018a).

Wastewater treatment operators projected that they would be most likely to participate in the PDR program during the dry season from May to October. Based on staffing restrictions, they anticipated bidding into the market once a day, Tuesday through Friday. They did not plan on participating during the wet season because of system constraints: the WWTP experiences higher and more variable flows that limit plant energetic flexibility and require additional staff oversight.

For the analysis it was assumed the battery could fill any gap between the projected energy loads and actual energy produced during the PDR event. It was also assumed that PDR bid placement would be chosen to optimize revenue, and all bids placed would be accepted.

To determine the bid placement for each bid block size B, a rolling average of the day-ahead LMPs were taken for the bid period, and the bid placement with the maximum average price was selected as the bid start time. The cost was determined for 2-hour, 4-hour, and 6-hour bidding blocks:

$$Bid\ Start = argmax\left(\frac{1}{B}\sum_{j=0}^{B-1} LMP_{d,h+j}\right) \qquad (3)$$

Once the bid placement was determined, projected PDR revenue was calculated by multiplying the curtailed load by real-time LMPs during the event bid period for each 5-minute interval I of the bid:

$$Revenue = \sum_{i-1}^{12B} LMP_{d,i} \times Energy\ Reduction_{d,i} \qquad (4)$$

Finally, differences between the adjusted baseline and the actual energy load were calculated by transforming hourly adjusted baseline data into 5-minute interval data and comparing it to measured 5-minute interval data during the PDR event. For the Laguna system, this difference reflects the amount of supplemental energy the battery must provide:

$$Difference_{d,i} = Energy\ Consumption_{\ d,i} - CBL_{d,i} \qquad (5)$$

Additionally, because participants in the PDR program still purchase electricity from their supplier at static TOU rates, changes in electricity costs due to participation in the PDR program were also included in the cost-benefit analysis.

2.4 Results and Discussion

2.4.1 Demand Response Test Events

The daily time-series measurements for Scenarios 1, 2, and 3 are plotted in Figures 3, 4, and 5, respectively. For all three scenarios, energy and asset measurements show issues with correctly

timing the operational changes in order to participate in the PDR event. Additionally, each asset takes time to ramp up to the full energy offset capacity of the scenario. This delay affects the WWTP's ability to effectively meet its target energy load reductions. Issues with delays and ramping could be mitigated by iteratively analyzing SCADA data or employing WWTP models to improve the timing of operational modifications during future demand response events.

There is also an observable variation between the actual energy load and the baseline energy load. The energy load profile is noisy, with large fluctuations occurring over short intervals. In contrast, the baseline methodology averages energy, significantly reducing noise in the energy baseline and making it difficult to measure how effectively the operational changes shift energy.

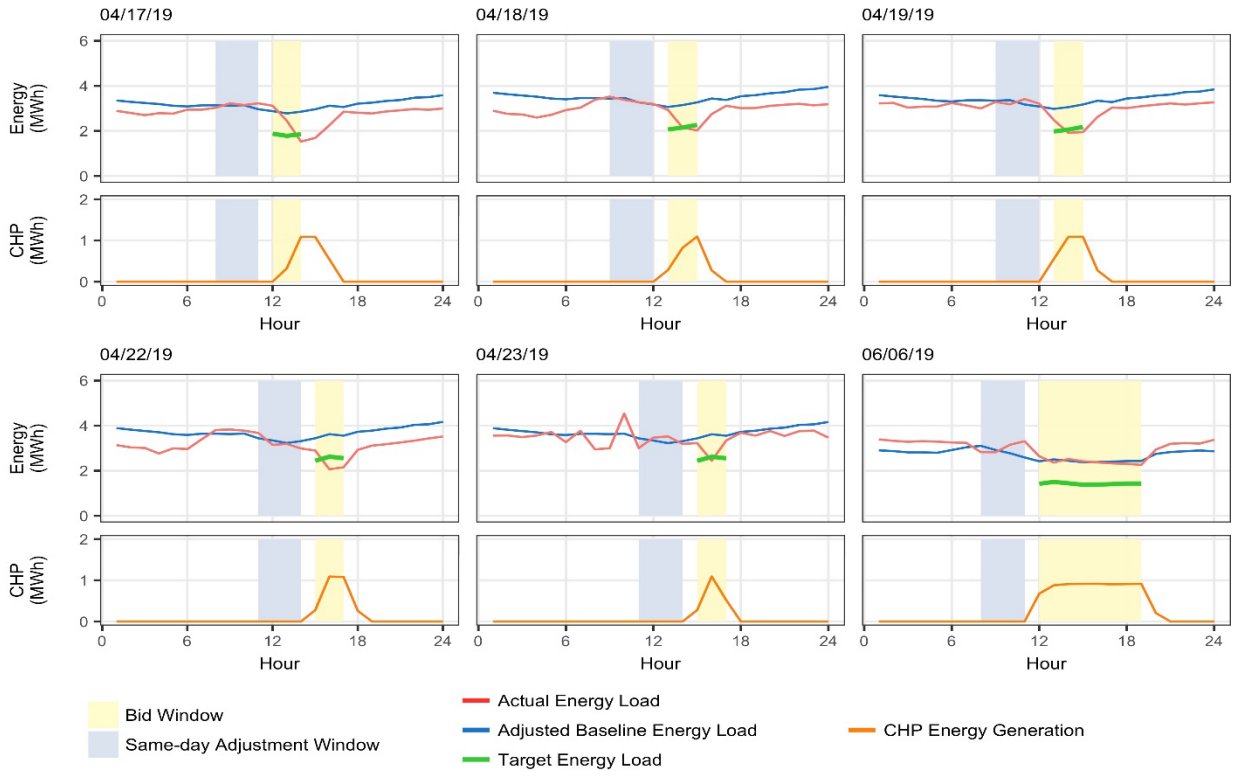

Figure 3. Daily time-series energy consumption and generation profiles for Scenario 1 PDR test events.

Despite the noise within the energy load, Scenario 1 test events demonstrate that turning on the CHP during the event period impacts the energy load profile (Figure 3).

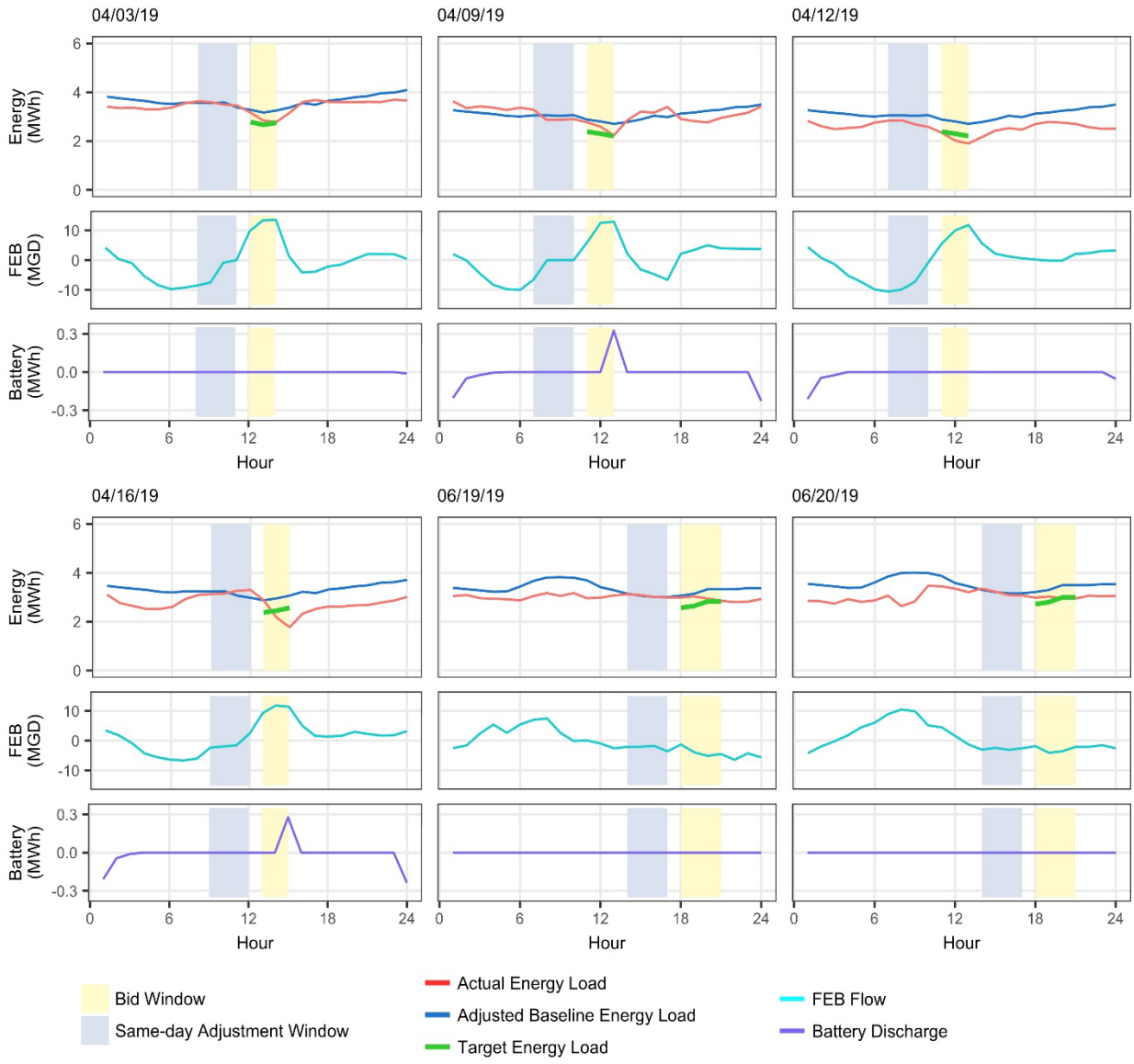

Figure 4. Daily time-series energy consumption, flow, and power profiles for Scenario 2 PDR test events.

In Figure 4, positive flows indicate the FEB is filling, and negative flows indicate it is draining. The first four trials show the operator ramping up the flow into the FEB slightly before the bid window and ramping down directly after the bid window. The timing of the flow diversion appears to be slightly offset from the timing of the energy load reduction. This is expected as energy consumption is reduced in treatment processes downstream from the FEB. Delays and ramping may be compounded by the energy requirements of pumping the diverted flow. For the final two

trials of Scenario 2, the flow diversion started significantly before the bid event, reducing energy before the bid window and therefore failing to meet the target energy reduction.

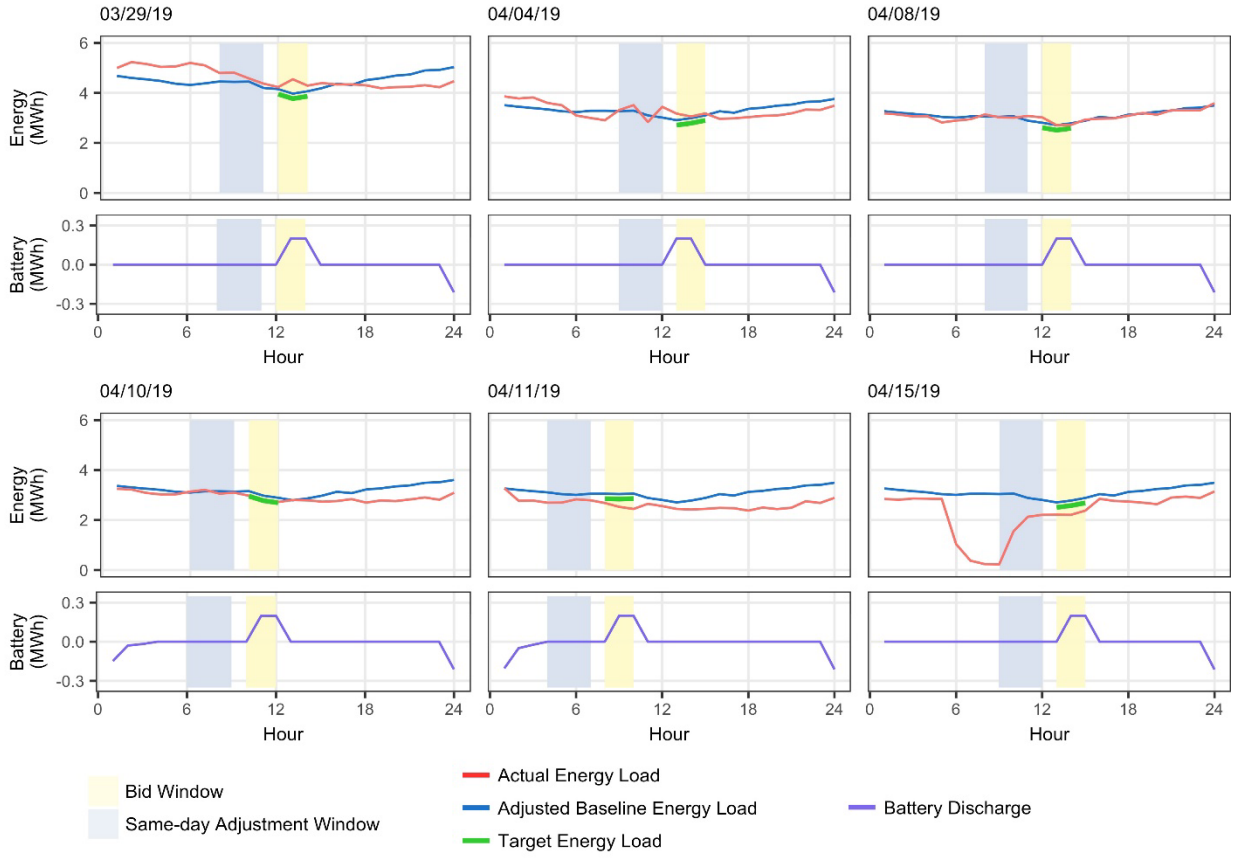

Figure 5. Daily time-series energy consumption and power profiles for Scenario 3 PDR test events.

Because of the fluctuations in the energy load profile, discharging the battery did not reliably allow the WWTP to meet the target energy load (Figure 5). The WWTP should, therefore, view battery discharge not as a stand-alone option but as a supplemental method for reducing the impacts of ramps and load fluctuations while implementing the other two load shifting scenarios.

Figure 6 depicts the hourly difference between the actual energy load and the target energy load for each trial. The bar plots to the right show how many of the total trials were successful or partially successful for each scenario. Scenario 3 had the most fully successful days, however, because the noise in the energy load profile has a similar magnitude to the energy reduction

supplied by the battery, the success rate in this scenario is likely influenced by random noise. Scenario 1 had the most partially successful days, with only the final trial showing no success over the entire bid window. Scenario 2 showed a lower success rate, although it had higher success in late April when the flow diversion was more closely aligned to the bid window. Additionally, when hourly targets were not met in Scenario 2, the difference between the target and actual energy load was generally lower than in the other scenarios.

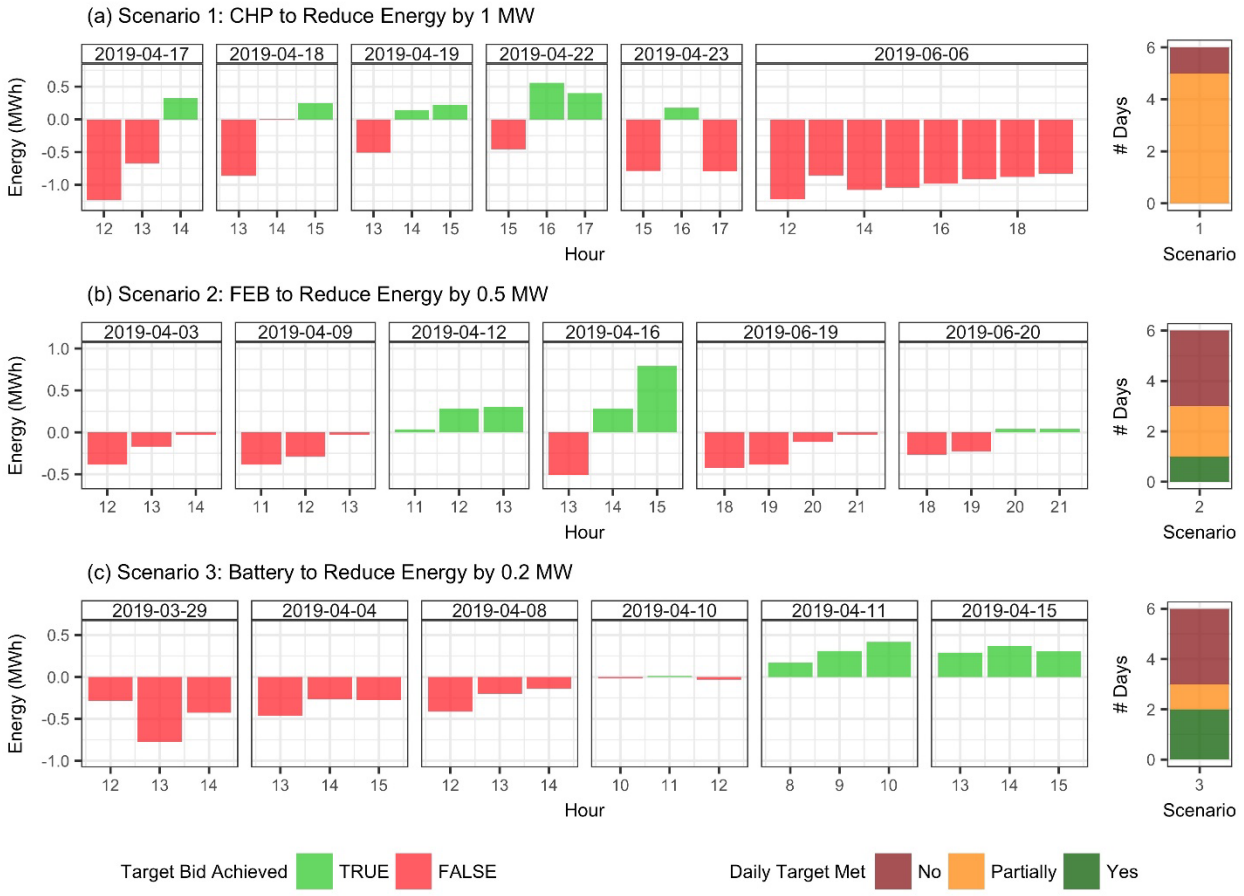

Figure 6. Difference between target energy load and actual measured energy load for all scenarios.

2.4.1.1 Water Quality Results

Table 4 compares the maximum measured concentrations for COCs regulated at Laguna WWTP during the demand response testing period to the MCLs outlined in Laguna WWTP's discharge permit (North Coast Regional Water Quality Control Board, 2013). The COCs that are regulated at

Laguna WWTP are biological oxygen demand (BOD), suspended solids, total coliform organisms and turbidity. Each COC's concentrations are regulated over a given combination of three possible time frames: monthly averages, weekly averages or daily maximum levels. As shown in Table 4, all of the water quality parameters measured during the study were below the required MCLs for each regulated time interval. This result was anticipated because the study's operational changes are within the WWTP's standard operating procedures: CHP generation and battery operation are not linked to water quality, and while flow diversion can influence the water treatment process, it was implemented within limits that reduce its impacts.

Table 4: Water Quality Results for PDR Facility during the PDR Test Event Study Period

Constituent	Unit	Monthly Average[b]		Weekly Average[c]		Daily Maximum	
		MCL[a]	Max Measured	MCL[a]	Max Measured	MCL[a]	Max Measured
BOD (20C, 5-day)	mg/L	10	2	15	2	--	--
Suspended Solids	mg/L	10	1.4	15	1.6	--	--
Total Coliform Organisms	MPN[e] /100 mL	--	--	2.2[f]	2	23[g]	7
Turbidity	NTU	2	0.84	--	--	5[h]	1.48

[a] Taken from NDPES Permit No. CA0022764 (State Water Resource Control Board, 2000)
[b] The mean of all effluent samples collected in a calendar month.
[c] The mean of all effluent samples collected in a calendar week, Sunday to Saturday.
[d] The daily discharge (kg/day) is obtained from the following calculation of any calendar day: $\frac{2.2}{N}\sum_i^N Q_i C_i$
Where N is the number of samples analyzed in any calendar day. Q_i and C_i are the flow rate (ML/day) and the constituent concentration (mg/L), respectively, which are associated with each of the N grab samples which may be taken in any calendar day. If a composite sample is taken, C_i is the concentration measured in the composite sample; and Q_i is the average flow rate occurring during the period over which samples are composited.
[e] Most Probable Number
[f] 7-day Median. The Median of all effluent samples collected in a 7-day period.
[g] May not occur in more than one daily result, where the result is the geometric mean of the online channels.
[h] Five NTU maximum not to be exceeded more than 5 percent of the time during any 24-hour period.

2.4.1.2 Operator Interview Results

Energy load shifting is not a typical part of wastewater operations. Laguna WWTP operators identified both barriers to implementation and opportunities for smoother integration of load shifting.

One of the main difficulties this facility encountered was navigating the design, permitting, and construction processes under the time constraints of grant funding. In particular, there were delays permitting the SCR units through the California Air Resources Board and processing the interconnection agreement with the energy utility. Delays for this project were compounded by the 2017 wildfires in this region, which put unanticipated strains on regulatory agencies reacting to emergency conditions. Future implementations might proceed more smoothly within the framework of municipal capital infrastructure development programs. These programs provide time for both the WWTP's administrative procedures and its established design and construction processes.

Prior to implementation, WWTP operators voiced concerns over their fragmented understanding of the PDR program. Specifically, they were not clear on how baselines would be calculated and settlements would be determined. This uncertainty may have been heightened by the fact that the PDR program was recently updated, and procedures and requirements had changed (CAISO, 2018b). Operators wanted to either develop in-house expertise or contract private consultants to interpret the energy data and market information they need to make informed operational decisions. Additional IOU or ISO outreach to educate facilities about demand resource programs might increase WWTP participation.

During the test events, operators identified several other difficulties. Although flow patterns are generally predictable, they are not predictable or controllable on an hourly basis. This variability made it difficult to forecast and adjust energy loads. The problem was compounded by issues with the same-day adjustment of the baseline. Flows in the morning are generally lower and, depending on the time of the bid, can affect the same-day baseline adjustment, limiting effective participation to later in the afternoon. Similarly, energy was reduced during the ramping periods for each scenario, and as a result of the same-day adjustment, the baseline was also reduced. In addition, operators found it difficult to achieve the smaller reduction targets (0.2-0.5 MW). Based on their

experiences with these test events, they believe it will be necessary to be conservative in their future energy load reduction bids.

2.4.2 Cost-Benefit Analysis Based on PDR Participation Model

Based on the results from the test events, it is estimated that Laguna WWTP could successfully participate in the PDR program if it were able to both improve the timing of the response to the bid and reduce the effects of ramping by discharging the battery at the appropriate time. The facility could reduce bid amounts to limit the risk of not meeting the promised reduction. The original bid amounts were used in this analysis to represent ideal participation levels.

Table 5 summarizes the results of the cost-benefit analysis. Based on projected participation, it is estimated that the facility could save between $8,015 to $45,563 annually through the PDR program, depending on the length of the facility's bid and which assets are used to load shift. It should be noted that these two strategies are independent of each other and can be combined for total savings up to $68,344. These savings include revenue from PDR participation and projected savings from shifting out of both peak and partial peak TOU pricing time periods. Given that the WWTP's total cost of electricity for 2017 was approximately $1,400,000, these savings represent a reduction in energy costs ranging up to 4.8 %. If the WWTP could increase participation year-round, it has the potential to achieve even larger cost reductions.

Table 5: Yearly Cost-Benefit Analysis Based on Projected PDR Participation

Event Length	Operational Changes for Load Shifting Event	Energy Load Shifted	PDR Revenue	TOU Savings[1]	Total Revenue
2 hours	Flow Diversion with FEB[2]	0.5 MWh	$5,175.574	$2,839.6	$8,015.17
	Run One CHP[3] with SCR[4]	1. 0 MWh	$10,351.15	$5,679.2	$16,030.35
4 hours	Flow Diversion with FEB[2]	0.5 MWh	$11,270.31	$3,405.28	$14,675.59
	Run One CHP[3] with SCR[4]	1. 0 MWh	$22,540.61	$6,810.56	$29,351.17
6 hours	Flow Diversion with FEB[2]	0.5 MWh	$16,393.95	$6,387.68	$22,781.63
	Run One CHP[3] with SCR[4]	1. 0 MWh	$32,787.9	$12,775.36	$45,563.26

[1]Energy cost taken from PG&E Tariff E-20T
[2]Flow Equalization Basin
[3]Combined Heat and Power Engine
[4]Selective Catalytic Reduction

Figure 7 presents the distribution of differences between the calculated adjusted baseline and the actual energy load at the 5-minute interval level that are used to determine financial settlements in the PDR program. If there is no difference, the adjusted baseline matches the actual energy load, meaning that the baseline correctly predicts the actual load and can be used to accurately determine curtailment. For most of the intervals, the adjusted baseline predicts the actual load reasonably well. However, there are several outlying data points that were as high as 1.25 MWh above or as low as 1.09 MWh below the calculated adjusted baseline, revealing that the adjusted baseline can sometimes greatly mispredict the energy load of this WWTP. It is not clear to what extent the battery will be able to provide sufficient energy load or load curtailment during these outlying events. Personnel should continue testing the battery to determine its operational limits. If the battery cannot provide the necessary energy, the WWTP can reduce energy bids to ensure it can meet the promised load curtailment; however, this reduction would decrease the profitability of participating in the PDR program.

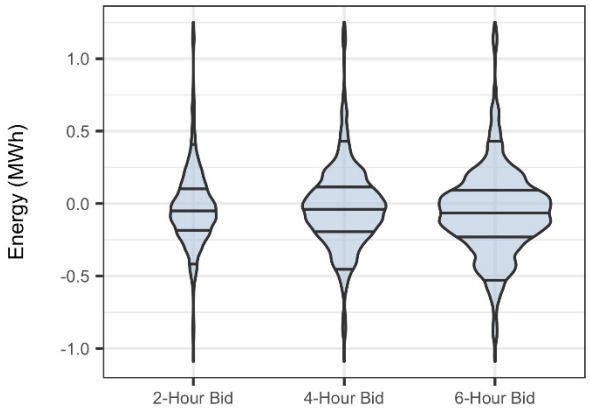

Figure 7. The difference between the adjusted baseline and actual energy load at the 5-minute settlement level. These violin plots show the full distribution points. The widths of each plot are relative based on the number of points. The horizontal lines represent the 0.05, 0.25, 0 0.75 and 0.95 quantiles.

Prior to the Laguna WWTP's participation in the PDR program, the facility participated in PG&E's BIP demand response program. From 2016 to 2018, the Laguna WWTP received an annual

average of $69,280 for its participation in BIP, or nearly 5 % of its annual electricity cost. Outside of regular audit events, the facility was only required to respond twice in those three years. The PDR program requires much more staff time and planned operational adjustments to achieve similar cost savings, suggesting the PDR program may not provide the financial incentives for complex industrial facilities such as Laguna WWTP to participate compared to the incentives of more traditional demand response programs.

2.4.3 Roadmap for Wastewater Treatment Plants to Participate as Demand Resource

As a part of this study, a roadmap was developed to guide WWTPs interested in participating as demand resources through energy load shifting (see Table 6). Three main requirements were identified for WWTPs to load shift to the standards needed to participate in typical demand resource programs: (1) instrumentation and automation infrastructure, (2) assets providing energetic flexibility, and (3) energy load stabilizing infrastructure.

Table 6: Roadmap for WWTPs Performing Energy Load Shifting

Requirements	**Recommended Assets**
Instrumentation/automation infrastructure	Time series energy metering[1] SCADA control system SCADA measurement instrumentation Water quality measurement instrumentation
Assets that can provide energetic flexibility	Combined heat and power engines Digester heat pumps Flow Diversion using equalization basins Flow Pumps Aeration Blowers[2]
Energy load stabilizing infrastructure	Battery Power conditioning system Microgrid controller (interfaced with SCADA systems)

[1]The PDR program requires 5-minute interval energy metering.
[2]Additional research is needed regarding the intermittent operation of aeration blowers to determine its effects on effluent quality.

At a minimum, demand resource programs require time-series energy metering to calculate baseline energy use and determine the amount of energy actually load shifted (CAISO, 2018b). Additional instrumentation, particularly automated control systems, can reduce user effort by

allowing remote optimization of operations. Without a certain level of automation, load shifting tasks may require excessive staff time (Schäfer et al., 2015). Additionally, water quality measurement instrumentation can confirm that load shifting operations are not affecting effluent quality (Aghajanzadeh et al., 2015).

Similarly, to perform energy load shifting assets must be selected whose operations can be modified without deteriorating effluent quality. Most demand response programs require a minimum energy load shift for a minimum time period. Based on their site-specific requirements, WWTPs may choose to shift energy loads using one or a combination of several assets to meet the minimum requirements of a particular demand response program.

Adding infrastructure that can stabilize energy load also allows WWTPs to more accurately project bids, reducing the possibility of failure to meet their energy reduction commitments. WWTPs have variable energy load profiles that primarily depend on inflow and BOD loading. Building in the ability to use residually stored energy to more precisely alter their loads is an important component for WWTPs to participate in energy load shifting.

Table 7 summarizes the requirements for participation in CAISO's PDR program. This table serves only as an introduction to the program and should not be considered a comprehensive set of all processes. Full and current details are available directly from CAISO.

Table 7: Requirements for Participation in the Proxy Demand Resource Program

Category	Requirement
Infrastructure requirements	Energy metering – 5-minute interval data
	Telemetry with a rated capacity greater than or equal to 10 MW
Minimum bidding requirements	100 kW minimum curtailment sustained for the duration of the bid (smaller loads may be aggregated to achieve the minimum)
CAISO PDR process requirements[1]	Execute a Demand Response Agreement
	Submit a new Demand Response Provider ID request
	Receive a Demand Response Provider ID
	Assign a certified Scheduling Coordinator [2]
	Receive Demand Response Registration System access
	Submit contacts to the Demand Response Registration System
	Obtain a user guide for Demand Response Registration System training
	Submit performance evaluation forms

[1] Additional details provided in the CAISO Demand Response User Guide (CAISO, 2018b).
[2] A list of certified Scheduling Coordinators is maintained on the CAISO website (CAISO, 2018b).

2.4.4 Study Limitations

The results of the cost-benefit analysis are subject to several limitations. Assumptions were made regarding when and how Laguna WWTP could participate in the PDR program; if altered, these assumptions could either increase or decrease the fiscal impact of participation. It was also assumed that all bids placed into the PDR program would be accepted, but because PDR is a competitive market, some bids might not be accepted, which would reduce earnings. Additionally, the study does not consider increases in energy required to divert flow into FEBs and increases in natural gas required to run additional CHP units. The study also did not consider changes in energy use based on changes in various system parameters and impact energy use including flow or pollutant loading through the treatment plant. An analysis of prolonged participation in the PDR program would provide further insight into both the economic potential and how load shifting might increase energy and natural gas use.

This case study used the PDR program's predominant baseline calculation methodology to analyze the test events and create the cost-benefit model. It was observed in both analyses that, in some cases, the adjusted baseline and actual energy load differed significantly. Such discrepancies

lead to inaccurate calculations of the amount of load curtailed. For Laguna WWTP the impacts of these discrepancies may be mitigated by the onsite battery installed to stabilize energy load.

2.4.5 Future Research

In general, additional research should be performed to determine a baseline calculation method that accurately predicts energy load and improves curtailment evaluation. A more accurate baseline will allow WWTPs to more confidently determine the financial impacts of energy load shifting. The PDR program has recently added two new methodologies for calculating baselines for non-residential customers: the weather matching baseline and the day matching baseline (CAISO, 2018b). Additionally, research on alternative baseline calculation methods utilizing more advanced models such as exponential weighting (Wi et al., 2009), regression (Mohajeryami et al., 2017) or supervised (Chen et al., 2017) or unsupervised (Park et al., 2015) machine learning methods have shown improvements in predicting residential and commercial energy loads. These methods should be explored to determine whether they can accurately predict energy load at WWTPs or whether the calculation requires an entirely new method for WWTPs specifically.

Future research should also address concerns regarding the capability of the WWTP to load shift itself. Researchers should continue exploring the impacts of load shifting on wastewater effluent quality, particularly for more sensitive energy-intensive components in biological treatment, such as aerators. Additionally, researchers should continue to explore how modeling and dynamic data can aid WWTPs in correctly timing and predicting the effects of the operational changes required for reducing energy.

2.5 Conclusion

WWTPs are energetically flexible facilities that energy grids can leverage as demand resources. This case study further illustrates that WWTPs are physically capable of shifting energy load with minimal or no impact on effluent quality and therefore, can reduce energy costs through

participation in demand resource programs. In this analysis, the BIP, a more traditional demand response program, was more cost-effective than PDR. However, the operational modifications that were explored can be used for any time-based incentive program, and the roadmap for WWTPs shifting energy loads and the summary of PDR requirements presented in this study are valuable resources for WWTPs interested in energy demand programs with time-based incentives.

This study also identified several barriers that make demand resource programs less attractive to WWTPs. First, it is difficult for WWTPs to correctly time and account for the ramping associated with operational modifications to meet the target energy load. WWTPs can improve the timing of operational modifications by analyzing SCADA data or utilizing WWTP models. Second, the current baseline calculation method averages energy, making it difficult to measure the impacts of operational modifications on the energy load accurately. Research should explore applying alternative baseline methodologies to wastewater treatment facilities. Finally, the financial incentives may not be strong enough to warrant the staff time and interruptions to the treatment processes to make participating in certain programs worthwhile. If the energy sector would like to incentivize WWTP participation in energy load shifting, they should design programs that are more easily accessible for WWTPs and that provide appropriate financial incentives.

Although further research will better enable WWTPs to shift energy loads and act as demand resources, there are already several assets operators can use with little or no concern for the negative effects on wastewater treatment processes. As renewable energy integration increases, the value of energetic flexibility in the energy market will likewise increase. As such, WWTPs should take advantage of opportunities they currently have to participate as demand resources and seek to expand those opportunities.

2.6 Acknowledgements

We would like to thank Laguna Wastewater Treatment Plant for participating in this study, with special acknowledgment to Joe Schwall - Deputy Director, Brian Bokkin - Wastewater Treatment Superintendent, and Tasha Wright - Energy and Sustainability Coordinator for all of their assistance. We also acknowledge Michael Day - Utility/SmartGrid Team Lead and Richard Swank - Microgrid Program Manager from Trane for their contributions to this project. Finally, we acknowledge Robert Good - Engineering Manager, from the Center for Water-Energy Efficiency for his help in managing this project.

This research project was generously funded by the California Energy Commission, grant number EPC 14-059. This article was prepared as a result of work sponsored by the California Energy Commission. It does not necessarily represent the views of the Energy Commission, its employees, or the State of California. The Energy Commission, the State of California, its employees, contractors, and subcontractors make no warranty, express or implied, and assume no legal liability for the information in this document; nor does any party represent that the use of this information will not infringe upon privately owned rights. This report has not been approved or disapproved by the Energy Commission nor has the Energy Commission passed upon the accuracy of the information in this report.

3 Chapter 3: Optimizing secondary time-based control structures to support renewable energy integration for water distribution systems

Abstract:

Shifting energy consumption to times when renewable energy is prevalent can allow for increased renewable energy integration and reduce greenhouse gas emissions from the electricity grid. As large energy users, water distribution utilities can participate in this type of energy demand management. Additionally, shifting rather than reducing pumping can help water utilities to decrease both costs and greenhouse gas emissions from energy consumption without sacrificing water reserves. Our study presents a new problem formulation for water distribution pump optimization: secondary time-based controls. This hierarchal pump control policy prioritizes maintaining water reserves but allows for energy load shifting to reduce costs under time-based energy incentives. We show this new control policy on three case studies of varying complexity and compare it to two benchmark pump control decision variable representations. In addition, a sensitivity analysis of the optimal policies using Monte Carlo simulation tests the robustness of the controls with uncertain water demands. Results show the proposed formulation reduces energy costs at similar or better levels than the two benchmark decision variable representations without reducing average water storage over the simulation period or violating system performance constraints. Additionally, the secondary time-based control policy better maintains water reserves and avoids violating system constraints with uncertain water demands. By shifting energy usage without depleting water reserves, water utilities can more comfortably manage energy demand and support the integration of renewable energy sources into the grid.

3.1 Introduction

Renewable energy generation is projected to increase significantly with the expansion of solar and wind power (International Energy Agency, 2019). However, these generation sources vary with weather, season, and time of day, causing operational and physical challenges for integration (Verzijlbergh et al., 2017). Integrating solar is especially difficult because its capacity falls at the end of the day when energy demand peaks (Badakhshan et al., 2019). Energy suppliers can increase solar integration by promoting energy demand management, incentivizing customers to change their timing of energy consumption through time-based pricing such as time-of-use (TOU) rates, demand response programs, or dynamic energy markets (Palensky and Dietrich, 2011). As energy utilities strive to expand renewable integration, incentives have moved from focusing on energy efficiency to energy demand management.

Water distribution systems (WDSs) have opportunities and incentives to optimize pumping operations to reduce energy costs. Water utilities consume about 4 % of global energy, with regional variations from 0.6 – 6.2 % (Kenway et al., 2019). Additionally, energy can account for 33 - 82% of water utilities' non-labor operating costs (Limaye and Jaywant, 2019), with pumping operations contributing to 85 % of energy costs (Sousa et al., 2016). Although energy drives costs at WDSs, there is generally a trade-off between energy reduction and two key reliability indicators: water storage levels and system pressure. WDSs must maintain sufficient water storage and minimum pressures to consistently meet customer and fire flow needs; reduced pumping generally decreases storage and system pressure. By shifting the timing of pump usage rather than reducing it, WDSs may more readily optimize costs to align with renewable energy generation with fewer impacts on system storage and pressure.

However, optimizing pump operations is challenging. WDS optimization is a mixed-integer nonlinear problem, and its complexity is NP-hard (D'Ambrosio et al., 2015). Real-world systems can be large and intricate, and they must balance optimization goals against consistently meeting

customer demands. Extensive research exists on optimizing WDS pump operations (Mala-Jetmarova et al., 2017). Several of these studies examine optimizing operations for different energy incentives that promote energy demand management, including TOU tariffs (Van Staden et al., 2011), demand response programs (Liu et al., 2020; Menke et al., 2016), and dynamic energy prices (Kernan et al., 2017). Others have looked at optimizing to minimize GHG emissions directly (Menke et al., 2017; Stokes et al., 2015). While these studies explored how WDSs could respond to different pricing structures or GHG emission factors, additional opportunities exist to restructure optimization strategies so storage levels and pressures can be maintained while managing energy demands.

When optimizing water systems, the problem formulation is key to method effectiveness and applicability (Maier et al., 2014). Comparing different control structures or objective formulations can show the benefits and drawbacks of different problem framings (Quinn et al., 2017). Several different pump operating control strategies exist for WDS operation optimization and many variations on how these control strategies can be represented as decision variables. A summary of previous decision variable representations for pump operation optimization appears in Table 8. The formulations presented focus on controlling fixed pumps, but many can be amended to work on variable speed pumps.

WDS operation optimization is typically formulated as a pump scheduling problem, where pumps are controlled explicitly by the time of day (Mala-Jetmarova et al., 2017). The most common decision variable representation is the status of each pump (on/off) for each time interval of the simulation (Mackle et al., 1995). Decision variables can be reformulated as pump start and end times (Bagirov et al., 2013), pump switching times (López-Ibáñez et al., 2008), or minimal equivalent pumps (Cimorelli et al., 2020) to reduce the size of the optimization search space. While explicit pump schedules can respond to time-based pricing structures directly, these control schemes depend highly on assumptions such as water demands and initial tank levels and need

frequent updating to be used in practice (Marchi et al., 2016). As a result, implementation of such optimized pump schedule policies has been minimal.

In practice, water utilities base pump controls on system state measurements. A few studies have explored optimizing these system-based controls, including controlling pumps based on pressures, flows (Cembrano et al., 2000), or water storage levels (Paschke et al., 2001). Researchers also have explored optimizing several different pump operating schemes that adjust tank level controls as a function of time (Alvisi and Franchini, 2017; Marchi et al., 2016; Quintiliani and Creaco, 2019). These time-dependent tank trigger controls help maintain storage levels and pressure with uncertain water demands while allowing pumps to respond to time-based pricing structures. However, because these control structures decrease the distance between the tank on and off triggers during some time periods, they can significantly increase the number of pump switches, accelerating pump deterioration (Quintiliani and Creaco, 2019). Additionally, several of the more intricate control structures may require complex programmable logic controller (PLC) programs to be implemented.

If operational requirements are not integrated into the optimization using system-based controls, they are often implemented as constraints. Storage reliability metrics can even be framed as a constraint (Kurek and Ostfeld, 2012). Although these constraints filter out broad performance issues, reducing energy costs is often prioritized over maintaining water supply. Another alternative is to use multiobjective optimization and include a reliability metric as an objective; this can help appropriately assess the trade-off between reliability and energy cost savings (Odan et al., 2015). However, many WDS performance measures do not fit within a traditional minimization or maximization framework, including water storage and pressure (Ostfeld et al., 2002). For example, increasing water storage and pressures may cause water quality issues or increased leakage (Van Zyl and Clayton, 2007). Optimal controls should maintain similar water storage levels and pressures to the current controls designed by engineers or system operators.

Table 8: Water Distribution System Pump Operation Optimization Decision Variable Formulations for Fixed Pumps

Decision Variables	Formulation Type	Variable Type	Number of Decision Variables	Reference
Pump status (on/off) during each time period	Pump schedule	Binary	$n_P \cdot n_t$	(Mackle et al., 1995)
Pump runtime during each time period	Pump schedule	Continuous/ discrete	$n_P \cdot n_t$	(Goldman and Mays, 1999)
Pump on/off durations	Pump schedule	Continuous/ discrete	$n_P \cdot n_{SW}$	(López-Ibáñez et al., 2008)
Pump on/off times	Pump schedule	Continuous/ discrete	$n_P \cdot n_{SW}$	(Bagirov et al., 2013)
Minimal equivalent pump status during each time period	Pump schedule	Binary	$n_t\left(\log_2\left(n_{mep}-1\right)+1\right)$	(Cimorelli et al., 2020)
Tank level to trigger pump on/off	Tank trigger	Continuous/ discrete	$2 \cdot n_P$	(Paschke et al., 2001; Van Zyl et al., 2004)
TT_{on}^{LT}, TT_{off}^{HT}	Tank trigger	Continuous/ discrete	$2 \cdot n_P$	(Marchi et al., 2016)
VTT_{on} during the LT, VTT_{off} during the HT.	Tank trigger	Continuous/ discrete	$2 \cdot n_P$	(Alvisi and Franchini, 2017; Housh and Salomons, 2018)
TT_{on}^{LT}, TT_{off}^{HT}, distance from the HT start, distance from the HT end.	Tank trigger	Continuous/ discrete	$4 \cdot n_P$	(Quintiliani and Creaco, 2019)

Definitions:

n_t = number of time periods within the simulation (time period is typically 1 hour)

n_P = number of pumps to be optimized

n_{SW} = number of pump switches in a 24-hour cycle

n_{mep} = Number of minimal equivalent pumps

LT = Low tariff time period

HT = High tariff time period

TT_{on}^{LT} = Tank level to trigger pump on for low tariff period

TT_{off}^{HT} = Tank level to trigger pump off for high tariff period

VTT_{on} = The endpoint for the variable tank level that triggers a pump on

VTT_{off} = The endpoint for the variable tank level that triggers a pump off

This work proposes a novel optimization formulation that adds and optimizes secondary time-based controls to pre-established system-based controls. The hierarchal control structure can optimize energy costs while maintaining water levels and pressures and requires a minimal number of decision variables in a structure that can easily be implemented through PLC logic. This decision variable representation is supported by an amended set of network performance

constraints that allow WDSs to maintain system storage and pressures. The optimization formulation is tested on several WDS systems of varying size and complexity to analyze its performance in minimizing energy cost using a metaheuristic optimization method. We compared this formulation to two more common decision variable representations, versions of the pump schedule and tank trigger controls, testing all formulations with and without a novel tank constraint. As our final performance metric, we examined each formulation's sensitivity to varying water demands to assess how uncertainty in demand forecasting affects the performance of these control structures. Water utilities can use this optimization formulation to shift energy usage without depleting storage reserves, supporting the integration of renewable energy sources in the grid through energy demand management.

3.2 Methods

3.2.1 Problem Formulation

The proposed secondary time-based (STB) control structure formulation was developed to mimic the typical operating procedures of water distribution systems responding to TOU prices. This control structure retains the original fixed tank level controls formulated by water system operators to maintain proper system pressures and ensure water reserves. The control structure then adds secondary time-based controls (Figure 8) that can be optimized based on time-dependent energy tariffs. The decision variables in this formulation are the times of day that the pump is triggered on or off in the second priority time-based rule. For each pump, p, for the given set of pumps to be optimized, P_{opt}, the pump control decision variables are:

$$\{x_{p,\ on}, x_{p,\ off}\}, \quad x \in [0,23] \tag{1}$$

Where $x_{p,on}$ is the time of day to turn pump p on, and $x_{p,off}$ is the time of day to turn pump p off.

We compared the hierarchal STB control structure to two established optimization formulations: the status-time pump schedule (STPS) formulation, a reduced variable variation of the original pump schedule representation (Bagirov et al., 2013), and the tank trigger (TT)

representation (Van Zyl et al., 2004) (Figure 8). These decision variable formulations represent the two extremes of pump control structures: controls only dependent on time or dependent on system parameters. These formulations have fewer decision variables than many other control structure variations (see Table 8), with an optimization search space size comparable to the STB formulation.

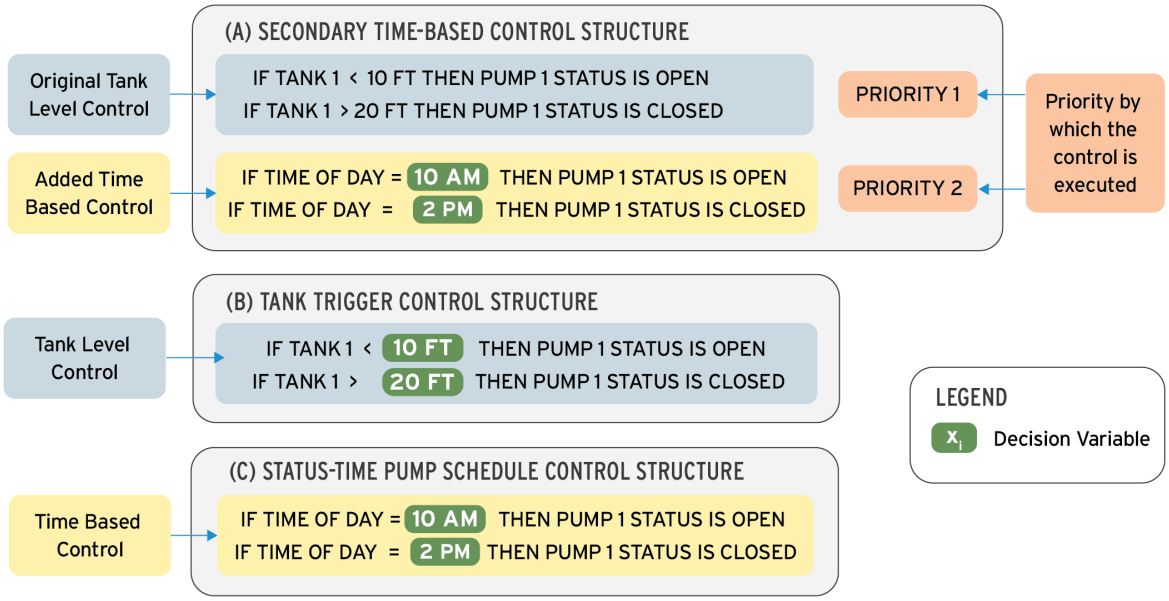

Figure 8. Pump Control Decision Variable Formulations. Illustrative depictions of (A) the proposed secondary time-based control structure, (B) the tank trigger control structure, and (C) the status-time pump schedule.

For this study, slight alterations were made to the STPS formulation proposed in Bagirov et al. (2013). The optimization formulation initially specified the same number of pump switches for all pumps. We amended the formulation so that number of pump switches can be unique for each pump, allowing tanks to maintain different cycling rates. Additionally, the initial operating status was not treated as a decision variable, utilizing the initial pump status from the original EPANET model.

3.2.1.1 Objective Value

By responding to time varying energy prices designed to promote energy demand management, utilities can reduce costs and support renewable integration. For time varying energy prices, energy

costs are determined from the energy tariff and the energy consumption for each time interval over a specified duration. The amount of energy used by a pump depends on the pump head, flow, and efficiency. Time series hydraulic measurements, including pump flow and head, are determined for each pump schedule using the EPANET hydraulic simulator (Rossman et al., 2020), assuming specified water demand patterns and initial conditions. The Python package WNTR (Katherine A Klise et al., 2017; Katherine A. Klise et al., 2017) was used to integrate EPANET simulations into a simulation-optimization framework. For this study, we omitted water source or pump maintenance costs from the objective function. Such additional features can be added to the objective function if applicable. The objective function is the total pumping energy cost over the simulation period:

$$\text{minimize } C_E = \sum_{p=1}^{n_p} R_t \sum_{t=1}^{n_t} \mathcal{P}_{p,t} \Delta t \;=\; \sum_{p=1}^{n_p} R_t \sum_{t=1}^{n_t} \frac{\gamma Q_{p,t} H_{p,t}}{\eta_p} \Delta t \tag{2}$$

where C_E is the total cost of energy for pumping, p is the pump index, n_p is the total number of pumps, t is the time interval index, n_t is the total number of time intervals, R_t is the energy rate at time t, $\mathcal{P}_{p,t}$ is the power for pump p at time interval t, γ is the specific weight of water, $Q_{p,t}$ is the flow in pump p at time interval t, $H_{p,t}$ is the head in pump p at time interval t, η_p is the efficiency of pump p, and Δt is the time interval duration (typically 1 hour). When the pump is off, no flow is supplied by the pump and the energy consumption and energy cost for that time frame are zero.

The STB control structure is formulated to respond to any time-based energy incentive. It is particularly suitable for areas with high variable renewable generation (IEA, 2020). For this study, we used historical prices from the California Independent System Operator's day-ahead wholesale electricity market as the energy tariff to calculate energy costs. Specifically, the tariff used was the average hourly prices from 2020 for all nodes, presented in Figure 9 (CAISO, 2020). This energy market has a high infiltration of solar energy and so typically incentivizes energy consumption in the middle of the day and disincentivizes energy use in the evening.

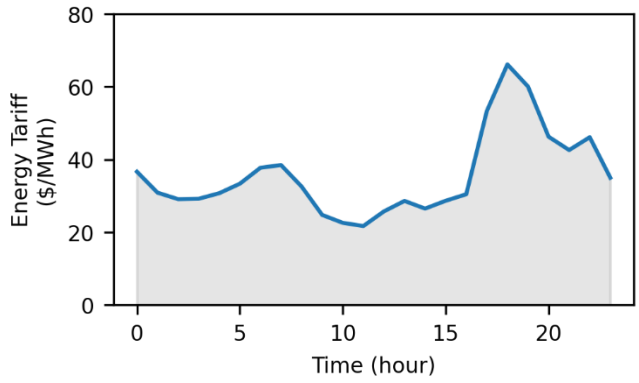

Figure 9. Average Hourly Energy Prices over a 24-Hour Period. Average hourly prices from the California Independent System Operator day-ahead electricity market (CAISO 2020).

3.2.1.2 Constraints

This optimization is subject to several constraints which ensure that pump operating policies are feasible and maintain hydraulic reliability and network performance. These include hydraulic constraints enforced by the simulation model, such as conservation of mass and energy, and operational constraints handled as part of the optimization formulation by penalizing the objective function. To maintain similar reliability to the original system controls, we propose to base the operational constraints on the simulation results of the original system. These constraints require that the proposed controls provide at least the same minimum pressure, maximum velocity, and minimum/maximum tank levels as the initial control set within a given tolerance.

Pressure and velocity are commonly constrained as adequate pressure is needed to deliver demands, and high velocities may erode pipes. The pressure and velocity constraints are defined as:

$$min(P_{n,t}) \geq min({P^0}_{n,t}) - \varepsilon_P \quad \forall\, n,t \tag{3}$$

$$max(V_{l,t}) \leq max({V^0}_{n,t}) + \varepsilon_V \quad \forall\, l,t \tag{4}$$

where $P_{n,t}$ is the pressure at junction n for time t, ${P^0}_{n,t}$ is the pressure at junction n for time t for the original control set, ε_P is the allowable pressure difference, $V_{l,t}$ is the velocity at pipe l for time

t, $V^0{}_{l,t}$ the velocity at pipe l for time t for the original control set, and ε_V is the allowable velocity difference.

As part of the optimization, it is also important to constrain minimum and maximum tank levels. The minimum and maximum levels specified in the hydraulic models will serve as the tank level constraints within a given tolerance. However, if the simulated tank levels for the original controls are above the maximum or below the minimum, the simulated tank level serves as the constraint within a given tolerance. The minimum and maximum tank level constraints are formulated as:

$$min_s(L_{s,\,t}) < min\{min_s(L^0{}_{s,\,t} - \ \varepsilon_{SM}),\ S_{\min level} + \ \varepsilon_{SM}\} \qquad \forall\, s \ \in S \tag{5}$$

$$max_s(L_{s,\,t}) < max\{max_s(L^0{}_{s,\,t} - \ \varepsilon_{SM}),\ S_{\max level} - \ \varepsilon_{SM}\} \qquad \forall\ s \ \in S \tag{6}$$

where $L_{s,\,t}$ is the level at tank s for time t, $L^0{}_{s,\,t}$ is the level at tank s for time t for the original control set, $S_{\min level}$ is the minimum water storage level for tank s, $S_{\max level}$ is the maximum water storage level for tank s, and ε_{SM} is the allowable maximum and minimum tank storage level difference.

Finally, it is critical to maintain storage levels to ensure the system is resilient to uncertain demands and operational issues. This is typically achieved by adding a constraint that refills each storage facility to a similar level by the end of the simulation. However, using this differencing storage constraint, storage levels can remain low much of the time. Adding a constraint that controls the average tank level over the simulation period better maintains system storage resilience, a key performance parameter for emergency and fire supplies. To test the impact of adding this average storage constraint, we optimized each case study twice, once using a variation of the common differencing storage constraint (equation 7) and once with the proposed average storage constraint (equation 8):

$$\frac{\left|L_{s,\ 1} - L_{s,\ n_t}\right|}{S_{\max level}} > \frac{\left|L^0{}_{s,\ 1} - L^0{}_{s,\ n_t}\right|}{S_{\max level}} + \varepsilon_{\mathrm{SD}} \tag{7}$$

$$\frac{\sum_{t=1}^{n_t}\left(L_{s,\ t}\right)}{T \cdot S_{\max level}} > \frac{\sum_{t=1}^{n_t}\left(L^0{}_{s,\ t}\right)}{T \cdot S_{\max level}} - \varepsilon_{\mathrm{SA}} \tag{8}$$

where T is the total simulation period, $\varepsilon_{\mathrm{SD}}$ is the allowable percent difference for the differencing storage constraint, and $\varepsilon_{\mathrm{SA}}$ is the allowable percent difference for the average storage constraint.

3.2.2 Optimization Method

To optimize the formulation presented above, we used particle swarm optimization (PSO), a search-based metaheuristic approach (Kennedy and Eberhart, 1995). The PSO algorithm was implemented using DEAP, a computational framework that allows users to build evolutionary algorithms in Python (Fortin et al., 2012). Since heuristic or metaheuristic methods are search-based, global optimization is not guaranteed (Blum and Roli, 2003). However, metaheuristic methods can be applied to the pump optimization problem without simplifying the hydraulic relationships, allowing a more confident interpretation of the impacts of operational changes on the system (López-Ibáñez et al., 2008). Along with genetic algorithms, PSO has frequently been applied to the WDS design problem (Mala-Jetmarova et al., 2018), as well as the pump operations problem (Wegley et al., 2004). While several variations of this algorithm increase its efficiency for different applications, we employ the original PSO strategy outlined in Poli et al. (2007).

In the PSO algorithm, a population (swarm) of NP candidate solutions (particles) evolve over g generations. The particles are moved around the search space based on each particle's best-known position and the best-known position overall. The population of candidate solutions is initialized with random positions and velocities within a given range. This population is evaluated using the objective value and constraints specified in the previous section. Each particle's position and velocity are updated and reevaluated until the population converges within some tolerance or until the maximum number of generations, G_{max}.

The cognitive coefficient, φ_p, affects how much the optimization searches locally, while the social coefficient, φ_g, determines the influence of the global best solution. The performance of the PSO algorithm can vary with the parameter values chosen for NP, φ_p , and φ_g. The parameter values used in this study appear in Table 9. Since this investigation uses a heuristic optimization method, each optimization was replicated with a different random seed and reported as the average of three replications. To assist the optimization in finding solutions within operational constraints, the initial population was seeded with a single feasible solution (Savic et al., 1997). Additional details on the optimization formulation and method and full optimization results are included in the supplemental information.

Table 9: Particle Swarm Optimization Parameter Values

Optimization parameters	**Variable**	**Value**
Cognitive coefficient	φ_p	2
Social coefficient	φ_g	2
Population size	NP	100^a /125^b
Maximum number of generations	G_{max}	80^a /100^b

[a] *For the Net 3 case study.*

[b] *For the Richmond Skeleton and the Moulton Niguel Water District case studies.*

3.2.3 Case Studies

Three hydraulic simulation models of varying sizes and complexity were analyzed to assess the performance of different pump control and decision variable formulations. The first two models, the Net 3 and skeletonized Richmond (RMSK) systems, are frequently used for benchmarking optimization methods (Rossman et al., 2020; Van Zyl et al., 2004). These networks are fairly simple, and the models are publicly available. The final model is of a larger and more complex network, the Moulton Niguel Water District (MNWD) reclaimed water system, presented for the first time in this paper. This system serves over 1,300 irrigation customers in Southern California. These WDSs are summarized in Table 10.

Some minor adjustments were made to the models for this study. One of the original controls in the Net 3 model was modified so that all initial controls for these three systems follow the TT control scheme, where pumps are turned on and off based on specified tank levels. The Net 3 and RMSK models have demand patterns for a single day. However, we simulated operations over seven days, with the demand pattern repeating every 24 hours, allowing us to examine how well the controls maintain storage levels over a longer time. Additionally, the number of optimized pumps in the MNWD system was limited to reduce the number of decision variables and computation effort. The nine pumps whose operations depend on tank levels with the largest energy use were selected for optimization; the remaining pumps operate using the original TT controls.

Table 10: Case Study Properties

	Net 3[a]	**RMSK[b]**	**MNWD[c]**
Total number of pumps	2	7	31
Number of pumps to be optimized	2	6[d]	8
Number of tanks	3	6	13
Total storage volume (m^3)	25,863	2,598	51,284
Number of pipes / junctions	117 / 91	44 / 41	4343 / 4277
Number of decision variables	4	12[d]/16[e]	16
Number of constraints	12	21	42

[a] *For the Net 3 case study.*

[b] *For the Richmond Skeleton case study.*

[c] *Moulton Niguel Water District case study.*

[d] *Number of decision variables for the secondary time-based control structure and tank trigger cases*

[e] *Number of decision variables for the explicit pump schedule case*

Six cases were examined for each pump control formulation: two optimizations for each system, either using the average or differencing storage constraint. Several metrics were analyzed to understand the performance of each control formulation and the impacts of the different tank storage constraints: the objective value of the best-found control policy, average percent storage, and the number of pump switches. The objective value indicates how well the optimization control structure allows the system to reduce energy costs without violating operating constraints. The

average percent storage indicates how well storage reserves were maintained, and the number of pump switches signifies potentially greater pump maintenance costs.

3.2.4 Demand Sensitivity Analysis

Using Monte Carlo (MC) simulation, a sensitivity analysis was performed to examine how the objective value, system storage, and pump switches vary with uncertain water demands for the different optimized pump control schemes. We also explored how water demand variation can affect constraint violations for each case study. The sensitivity analysis was applied only to the optimized control structures identified using the average storage constraint. The objective value also was calculated using the average storage constraint; the differencing value constraint was not applied.

All three models simulate average demands over a week. For this study, demands vary from the average following a truncated normal distribution, allowing only positive water demands, with a standard deviation of ten percent of the mean (Babayan et al. 2005). The demands were randomly sampled each timestep, assuming randomness is temporally independent. For each optimized control set for all replicates of every optimization run, we ran 2,000 simulations for the Net 3 cases, 7,000 simulations for the skeletonized Richmond cases, and 10,000 for the larger MNWD cases. To verify that this sample size was sufficient, an additional 100 random samples were drawn to confirm that there were minor impacts on overall distributions.

The objective value, average storage, and the number of pump switches were found for each simulation. The distribution of these values was then analyzed and compared across the three control structures. If the optimized control structure is stable, the objective value will have a lower unpenalized value (< 1e7). Lower variation in these three metrics indicates that the control structure is robust for uncertain water demands.

3.3 Results

3.3.1 Optimization Results

For each of the six optimized cases, we evaluated the objective value (energy cost over the simulation period plus any constraint violation penalties) of the best-found control policy, along with operational performance metrics of average percent storage and the number of pump switches (shown in Figure 10).

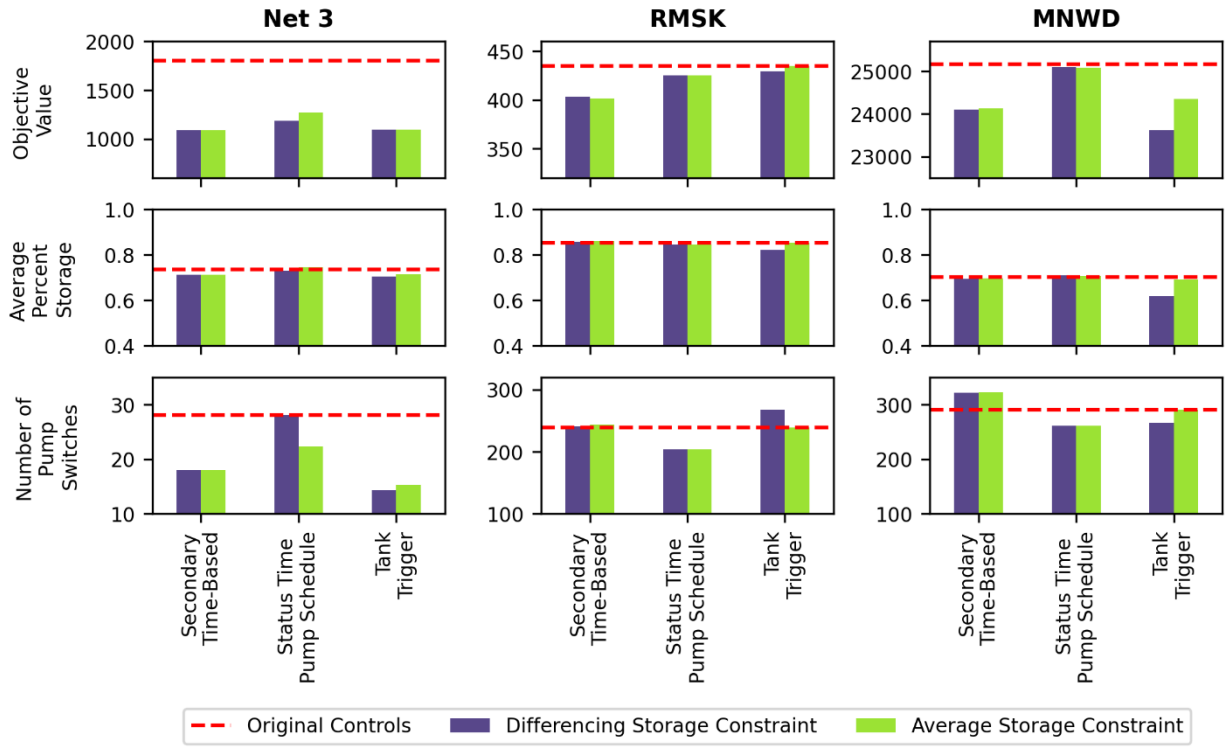

Figure 10. Best-Found Policy Results from Optimization. Simulated results from the three optimized control structures for the Net3, skeletonized Richmond (RMSK), and Moulton Niguel Water District (MNWD) case studies using either the average storage or the differencing storage constraint.

Results of the Net 3 optimizations show energy cost was reduced significantly for all three control structures. Energy cost was most reduced by the STB formulation, followed closely by the TT structure for both storage constraints. The STPS optimization found a lower energy cost solution

using the differencing constraint, but the TT and STB formulations still allowed greater reductions for the given computational budget. All three optimized policies perform similarly well in maintaining storage over the simulation period. The STPS runs had more pump switches, followed by STB, with TT performing the best, indicating a slight reduction in pump maintenance.

For the RMSK system, the STB control structure more clearly outperforms the other two control formulations. The STB optimization results were similar regardless of which storage constraint was applied. The different effects of these storage constraints are observable in the TT control optimization. The TT optimization found a slightly lower objective value when using the differencing storage constraint; however, the storage reserves also were reduced during optimization, decreasing system resiliency. Water utilities typically prioritize maintaining appropriate water storage levels. Therefore, including an average constraint for these formulations should lead to solutions more aligned with water utility priorities despite potentially higher energy and maintenance costs.

For the MNWD system, the optimizations of the STB and the TT control structures led to similar energy cost reductions. The STB performed slightly better with the average constraint, and the TT formulation performed better with the differencing constraint. As observed in the RMSK case, the TT formulation could further reduce costs with the differencing constraint at the expense of storage reserves. Again, the STB formulation converged on similar solutions regardless of whether the averaging or differencing storage constraint was applied. There are slight increases in the number of pump switches for the STB formulation, but this is less prioritized than maintaining water reserves.

Across all three case studies, the STB control structure most successfully reduced energy costs while maintaining storage levels without significantly increasing the number of pump switches from baseline operations. In the one case where the TT formulation outperformed the STB

structure, the energy cost reductions were at the expense of lower average water storage levels. For the larger systems, the STPS formulation could only marginally reduce energy costs from the baseline operations. Since the STPS formulation most directly responds to energy cost, this limited reduction is likely from the difficulty of this formulation in meeting the operating constraints. Applying the average tank constraint when using the TT or STPS formulations ensures that the optimized control policies maintain storage reserves. However, the STB formulation did not need the average tank constraint to maintain system storage. These results indicate that the hierarchal STB control structure can enable water utilities to take advantage of time-based energy tariffs, allowing systems to reduce costs without sacrificing water storage redundancy or significantly increasing pump maintenance.

Comparing the time series simulation results for the baseline and optimized scenarios for one of the case studies shows the differences between these control formulations. Figure 11 shows average tank levels and energy consumption for the RMSK system over the first 48 hours. The CAISO energy tariff is plotted on the secondary axis to illustrate how the optimized formulations respond to the price incentive.

For the first control structure formulation, STPS, optimized operations vary minimally from baseline conditions. Increasing the number of pump switches per day may allow this formulation more flexibility to respond to the price tariff while maintaining operational constraints. However, increases in pump switches exponentially increase the search space, making optimization intractable for larger systems. Operations under the optimized TT control and averaging storage constraint are similar to baseline operations; this is likely because only a few tank trigger settings maintain the same level of water reserves. Under the differencing constraint, we see more flexibility in operations, but there are corresponding reductions in water storage levels over the simulation. While there is more observable energy load management, the pumping increases do not target the lowest tariff timeframes, and energy reductions are not within the highest tariff periods. The STB

formulation shows observable changes in the timing of energy consumption without reductions in water reserves. The plot also shows that the STB formulation has the flexibility to reduce energy during the highest tariff periods and increase energy in the lowest cost timeframes. These results again indicate the suitability of the STB formulation to facilitate energy demand management.

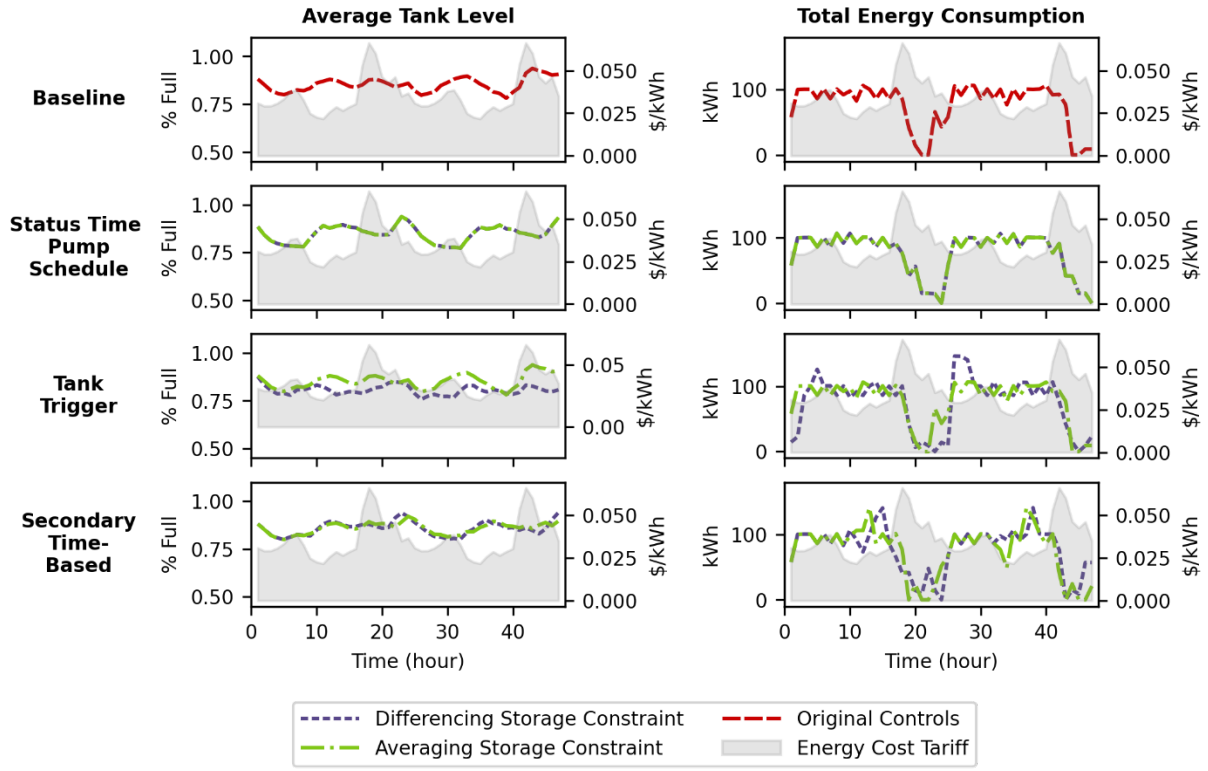

Figure 11. Energy Consumption and Average Tank Level for Richmond Skeletonized Case. The simulated results from baseline operations and the three optimized control structures for the first 48 hours are plotted on the primary axis; the energy tariff is plotted on the secondary axis.

3.3.2 Control Structure Impacts on Optimization Search Space

The results in Figure 10 were simulated using the best policies from the search-based optimization. Many other control policies were sampled during the optimization, providing insights into how the objective values change with decision variable settings. The "black box" nature of these types of heuristic optimization routines can mask underlying issues with the optimization formulation. By analyzing all policies evaluated during the optimization routine, we can better

understand the stability of the different control structures. If a large proportion of the tested policies violate the given constraints, the control structure is very sensitive to the setting of the decision variable. A high constraint violation rate indicates that the stability of the solution may be low and that it may be difficult to optimize the control structure because of the high rate of constraint penalties over the optimization search space. This may also suggest that small changes in the system may affect the validity of optimized controls, which is more explicitly explored in the sensitivity analysis. In this section, we analyzed the solutions tested over the six optimizations to see what percent of the operations were feasible (violating no constraints) and which constraints were violated for the infeasible solutions (see Figure 12).

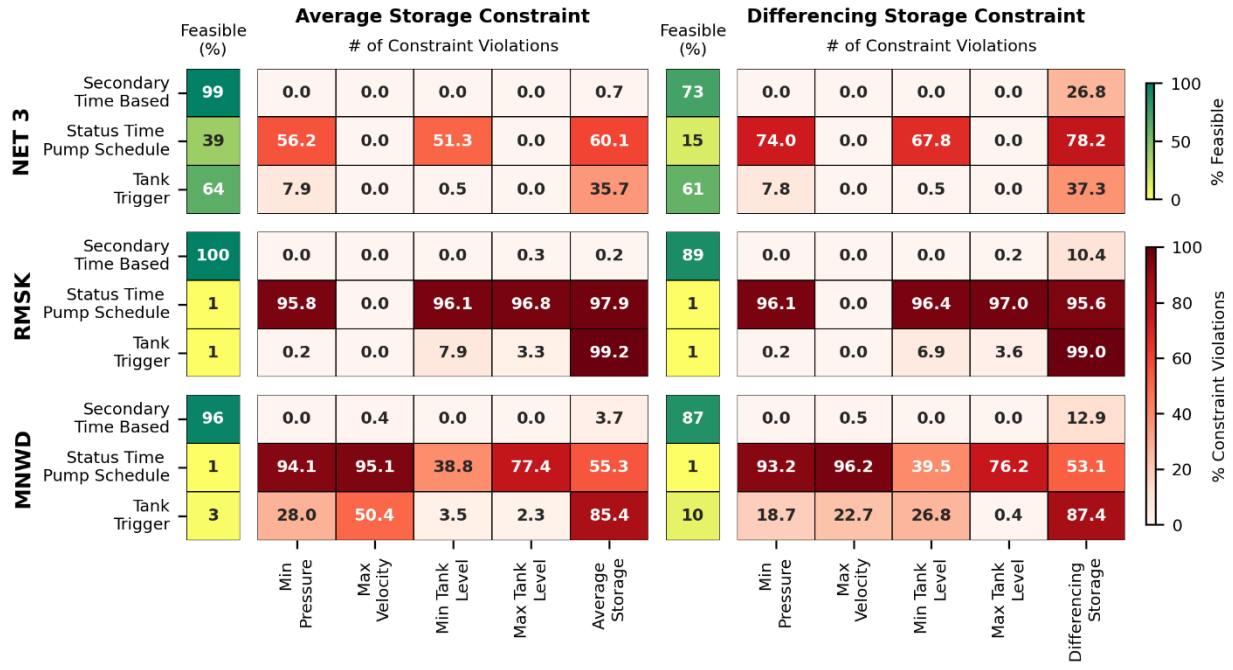

Figure 12. Feasible Solutions and Constraint Violations during Energy Cost Optimization. The percentage of the feasible policies tested and violated constraints for the optimizations of each of the three tested control structures. These results are presented for the Net3, skeletonized Richmond (RMSK), and the Moulton Niguel Water District (MNWD) case studies testing either the average storage or differencing storage constraint.

Figure 12 shows that the optimizations using the STB control structure generally have a high percentage of feasible solutions over the search space, particularly under the averaging storage constraint (96 - 99 %).

For the Net 3 optimizations, the highest number of constraint violations for the STB control policy was for the average or differencing storage constraints. When optimizing the TT formulation, the tank filling constraints again have the highest number of violations, but a much larger proportion of the evaluated policies violate this constraint. For the Net 3 STPS optimization, a much larger proportion of solutions tested were infeasible; 61% under the averaging storage constraint and 85 % for the differencing storage constraint. Here we see a much higher rate of minimum pressure and minimum tank level constraint violations and a large percent of storage filling constraint violations.

For the larger, more complex RMSK and MNWD systems, the TT and STPS formulations show even more constraint violations across additional constraint categories, including minimum pressure, minimum and maximum tank level, and the average storage constraints. Only 1 % of solutions were feasible for the RMSK cases using the STPS and TT formulations for either tank filling constraint over the given number of generations. The TT formulation performed slightly better with the MNWD system under the differencing constraint but still found only 3 % of tested solutions were feasible under the average storage constraint. Especially considering that these optimizations were all seeded with initial feasible solutions, this indicates that these control structures are difficult to optimize in larger systems, and the feasibility of solutions is very sensitive to decision variable settings.

By investigating what proportion of the sampled search space was infeasible, we can gain significant insights into how stable the solution is and how difficult a formulation may be to optimize. Understanding the impacts of decision variable formulations is particularly important for

critical applications such as optimizing water distribution operations since constraint violations may indicate significant operational performance issues.

3.3.3 Sensitivity Analysis Results

An MC sensitivity analysis assessed the impacts of water demand variation on energy cost, average storage levels, and the number of pump switches for each control structure. Three case studies were tested, the Net3, RMSK, and MNWD systems using the control structures optimized under the average storage constraint (see Figure 13).

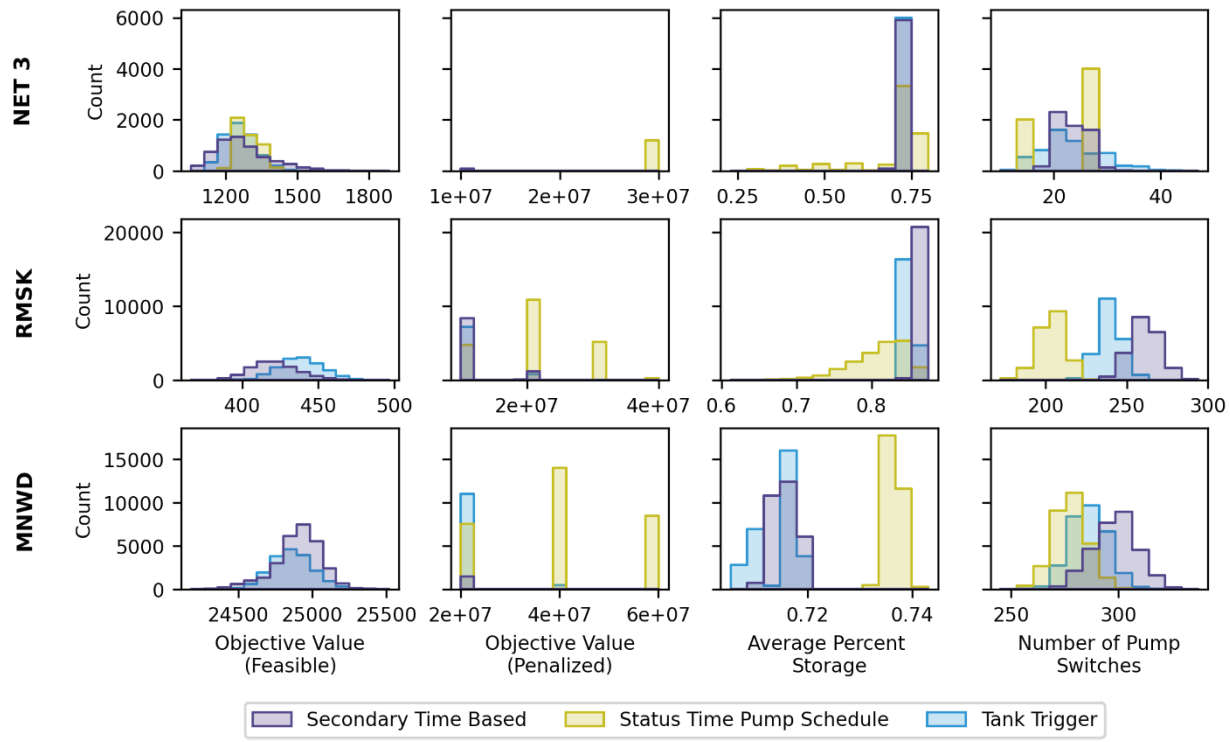

Figure 13. Water Demand Sensitivity Analysis Results. Simulated results from performing a Monte Carlo sensitivity analysis using the optimized control structures for the Net3, skeletonized Richmond (RMSK), and Moulton Niguel Water District (MNWD) case studies using the average storage constraint.

For the Net 3 system, energy cost varies for both optimized STB and TT control policies, but very few solutions become infeasible due to the variation. However, the STPS structure has many more constraint violations under the same variation; this is expected as the STPS control structure

does not respond to system measurements. Similar trends occur for the RMSK and the MNWD systems. The MNWD system appears more sensitive to changes in TT settings because many more constraint violations occur with the optimized control than in the other systems. This result reaffirms conclusions from the optimization search space analysis. The optimized TT controls may be unstable if the system is sensitive to changes in tank levels because adjusting the tank level control can impact key system parameters, such as pressure.

Across all three systems, variation in the average percent storage for both the STB and the TT optimized control policies was minimal; this is intuitive as the controls are activated as a function of storage level and dynamically respond to changes in water demand. The STPS optimized control policies generally have higher variation in storage level as demands change. Conversely, there are larger variations in the number of pump switches for the optimized tank-based control structures compared to the STPS control since variation in demand can change the number of times the tank-based control triggers a pump.

In addition to analyzing the variation of key system parameters as a function of changes in water demand, we also tracked if and which constraints were violated for each MC simulation (see Figure 14).

Figure 14 shows that there were fewer constraint violations when using the STB control and less variation in what constraints are violated compared to TT and STPS policies. The strength of the STB control structure is because tank levels can directly affect pressure levels and velocity in pipes. System operators developed the original tank controls to maintain reliable system operations. As a result, if the original tank controls are well suited for the expected variation of the system, the optimized STB control will perform similarly well. This gives the STB formulation an advantage even over the TT formulation, which could potentially find tank trigger levels that may work under current conditions but not be resilient to changes in demands.

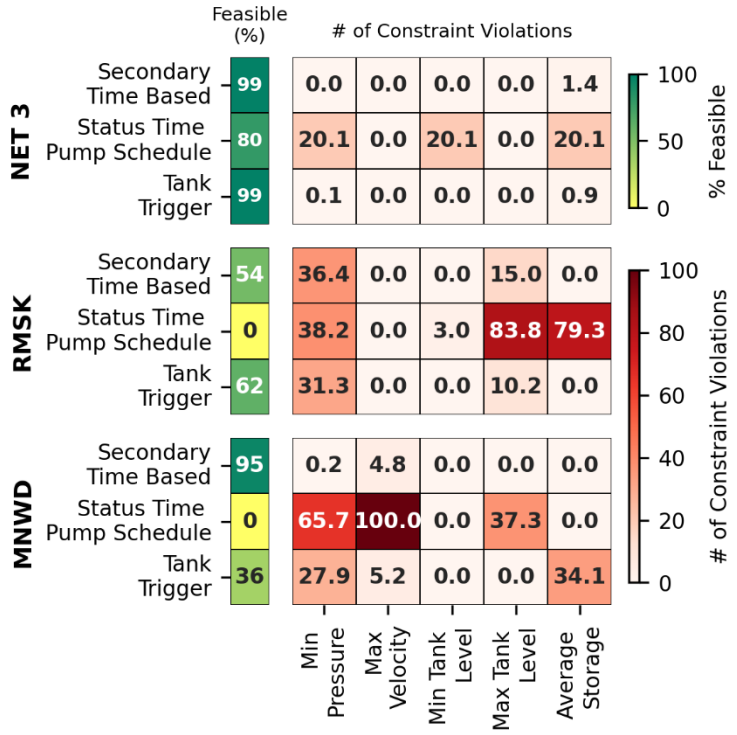

Figure 14. Feasible Solutions and Constraint Violations during the Water Demand Sensitivity Analysis. The percentage of the feasible policies tested and the number of violated constraints for the Monte Carlo simulations using the optimized control structures for the Net3, skeletonized Richmond (RMSK), and Moulton Niguel Water District (MNWD) case studies using the average storage constraint.

3.4 Conclusions

A new optimization formulation for distribution system operations considering water reliability and energy use, the STB control structure, was presented and compared to two standard formulations, the TT and STPS control structures. This hierarchal control structure allows water utilities to maintain water reserves while allowing pumps to respond to time-based energy incentives to support renewable energy integration. The formulations were tested on three case studies of various sizes, investigating their performance with a novel averaging tank constraint or a common differencing tank constraint. By using the STB formulation, water utilities can reduce

energy costs without reducing water storage reserves, with little or no increases in the number of pump switches. Compared to traditional operation structures, the STB formulation better maintained water reserves even without the average tank level constraint, with similar or better reductions in energy cost.

In addition to the optimization results, we analyzed simulation data collected during the optimization routine to better understand the effects of model formulation on the fitness function. Throughout the optimization routine for the STB control structure, a high percentage of the tested policies were feasible (had no constraint violations) compared to the STPS and TT formulations. This indicates that the STB control structure is more stable since different control settings can stay within the constrained performance bounds. It also may be easier to optimize since there are fewer areas penalized by constraint violations, leading to a noisy objective function. Finally, a Monte Carlo sensitivity analysis showed how changes in water demand affect optimized control policy performance. The STB performed consistently well across all three case studies; very few tested simulations resulted in constraint violations, and water storage was maintained. As this control structure aligns closely with water utility priorities, it may allow water utilities to more easily accept and implement optimized solutions, particularly as they move to increase energy demand management to navigate the growth of renewable energy sources.

3.5 Data Availability

The MNWD hydraulic model used during the study was provided by a third party. Direct request for these materials may be made to the provider as indicated in the Acknowledgments. Additional data, models, or code generated or used during the study are available in a repository or online in accordance with funder data retention policies (https://github.com/emusabandesu/wdpo.git).

3.6 Acknowledgments

Special thanks to the California Energy Commission for financing this study with funds from the Electric Program Investment Charge, the ratepayer surcharge authorized by the California Public Utilities Commission, through agreement number EPC-16-062. Special thanks to the MNWD and its staff for their excellence in supporting this and other research into the water-energy nexus and energy efficiency in the water sector.

4 Chapter 4: A visual framework for analyzing fitness functions and a correspondent decision tree guided genetic algorithm for water distribution system optimization

4.1 Abstract:

The water distribution system operation optimization problem has been extensively researched. Pump operations can be optimized for a range of objectives to improve the efficiency and sustainability. However, optimizing water distribution operations is difficult as a mixed-integer nonlinear problem with many decision variables and constraints. Heuristic or metaheuristic methods can allow the water distribution operation optimization without simplifying the hydraulic relationships. Nonetheless, it can be challenging to assess issues with problem formulation or optimization search strategies using these "black box" methods. We present a visual analytic framework that allows users to characterize the optimization fitness landscape, which relates optimization inputs and the objective function, to aid analysis of the optimization problem formulation and search efficiency. Visualizations, decision tree machine learning models, and fitness landscape characterization metrics are used to extract actionable information. We also present a new optimization method, the decision tree guided genetic algorithm, to help guide the optimization using similar machine learning methods and analytics in the characterization approach. We observed differences in the fitness landscapes due to problem formulation by examining two decision variable representations on a simple case study network. We then tested the performance of several existing optimization methods on these two problem formulations, using either standard optimization or optimization with guidance from the visual analytic framework. We compared this to the proposed optimization method. Both the existing methods informed using the visual analytic framework and the new optimization methods showed improvement over existing standard methods.

4.2 Introduction

Extensive research has been conducted on optimizing pump operations for water distribution systems (WDS) (D'Ambrosio et al., 2015; Mala-Jetmarova et al., 2017). Pump operations may be optimized for a range of objectives to improve system efficiency or sustainability. These objectives may include energy use, energy cost (Bagirov et al., 2013), greenhouse gases (Blinco et al., 2016; Stokes et al., 2015), water quality (Arai et al., 2013; Kurek and Ostfeld, 2013), water losses (Giustolisi et al., 2013) and system reliability (Odan et al., 2015). However, optimizing WDS operations is computationally demanding; WDS hydraulics are nonlinear, and optimization is a mixed-integer nonlinear problem and NP-hard (D'Ambrosio et al., 2015). The complexity of the problem is compounded by the many decision variables in real-world networks, limiting the effectiveness of many optimization techniques (Rao and Salomons, 2007).

Early research focused on classical optimization approaches such as linear control theory (Fallside and Perry, 1975), linear programming (Jowitt and Germanopoulos, 1992), nonlinear programming (Yu et al., 1994), and dynamic programming (Sterling and Coulbeck, 1975). However, these classical methods can only be applied to simple networks or by simplifying the mathematical representation of WDS hydraulics (D'Ambrosio et al., 2015). In contrast, heuristic or meta-heuristic methods can be applied to the pump optimization problem without simplifying the hydraulic relationships (López-Ibáñez et al., 2008), although global optimization is not guaranteed (Blum and Roli, 2003). Some prominent heuristic methods applied to this problem are genetic algorithms (Mackle et al., 1995), hybrid genetic algorithms (Van Zyl et al., 2004), ant colony optimization (López-Ibáñez et al., 2008), particle swarm optimization (Wegley et al., 2004), and simulated annealing (Goldman and Mays, 1999).

Although there are benefits to heuristic approaches to complex optimization problems such as WDS pump operation, there are several common challenges. It can be difficult to assess issues with

optimization problem formulation or the search method because of the black-box nature of heuristic simulation-optimization methods (Maier et al., 2014). As a result of this lack of transparency and the uncertain effectiveness of the methods, users often mistrust these types of automated systems (Lee and See, 2004). One approach to improving optimization results and user acceptance is to more formally integrate humans into the optimization process through human-in-the-loop or human-as-the-loop optimization. Although this is a developing field, research on human-integrated optimization systems dates back to as early as the 1970s (Benayoun et al., 1971; Wallenius, 1975). Interactive optimization can be used to adjust or enrich the formulation of the optimization problem itself (Meignan, 2015), identify preferences for multiobjective solutions (Miettinen et al., 2010), or be used to help guide the search method in finding an appropriate solution (Klau et al., 2002). Most research in this field focuses on developing interactive systems for problems with multiple objectives (Xin et al., 2018). An in-depth review and taxonomy of human integrated optimization methods for operations research is provided in David et al. (David et al., 2015).

Visualization of optimization results and search spaces can aid in understanding optimization performance and facilitate interactive optimization. Research in this area often focuses on visualizing multiobjective optimization results. Parallel coordinate systems (Bagajewicz and Cabrera, 2003), heatmaps (Pryke et al., 2007), self-organizing maps (Zhang et al., 2018), Sammon mapping (Valdés and Barton, 2007), and radial visualization (Hoffman et al., 1997) have been used to assess tradeoffs between different optimization goals. Visualizing the fitness landscape, which is the relationship between optimization inputs and outputs, has also been explored. Because of the high dimensional nature of the optimization inputs in the fitness function, visualizing the relationship between optimization inputs and outputs typically requires some sort of dimension reduction or mapping. Methods used to visualize the optimization fitness landscape or the optimization algorithm's trajectory through the fitness landscape include principal component

analysis (Collins and O'Neill, 1999; Mccandlish, 2011), T-SNE (Duan et al., 2017), or Sammon mapping (Kim and Moon, 2003).

Characterizing the fitness landscape analytically can give additional insights into the complexity of the optimization problem and what methods might be effective (Pitzer and Affenzeller, 2012). Several metrics exist to measure different characteristics of the fitness function. Key areas of measurement include landscape ruggedness (Hordijk and Nm, 1996), the correlation between fitness and location (Jones and Forrest, 1995), and how separable feasible areas are from areas with constraint violations (Malan et al., 2015). These landscape fitness characterization metrics can also be directly incorporated into the optimization method itself (Li et al., 2022; W. Li et al., 2021).

A selection of interactive optimization, optimization visualization, and characterization techniques have also been applied to WDS operation optimization problems or WDS optimization design, which strongly parallels WDS operation optimization. Researchers have explored how the visualization of multiobjective results (Fu et al., 2012) and the spatial-temporal visualization of decision variable changes in coordination with optimization results (Keedwell et al., 2015) can aid in optimizing WDS design. Li et al. (2021) investigated visualization of the WDS operation fitness function, presenting an interactive visual analytics system that incorporated decision trees as interpretable machine learning to analyze input-output relationships. Gibbs et al. (2004) also applied fitness landscape measures to the WDS optimization problem.

This study presents a framework for the visual and analytical characterization of WDS operation optimization fitness landscapes to aid in analysis of both optimization problem formulation and optimization search efficiency. This approach expands the visual and machine learning components from the visual analytic tool presented in Li et al. (2021) by incorporating fitness landscape analysis. We apply this framework to characterize the fitness landscape and to show how users can employ this type of characterization to analyze both the problem formulation

and the behavior of optimization methods through a set of case studies. In addition, we present a new optimization method to help guide the optimization using similar interpretable machine learning methods and analytics to the characterization approach. The optimization method uses iteratively updated decision trees to guide a genetic algorithm to predicted optimal regions. Although both the optimization characterization approach and the decision-tree genetic algorithm (DTGA) were designed for the WDS operation optimization problem, these methods can be applied to a large range of complex optimization problems.

4.3 Methods

An overview of the proposed visual analytic framework is provided in Figure 15.

First we collect a sample of the fitness landscape to analyze the fitness function. To perform the initial characterization, we use Latin hypercube sampling to ensure good coverage of the search space. For this study, we analyze two sample sets (10,000 samples and 2,000 samples) to see the impact of sample size on our analysis. From the sample, we characterize the fitness landscape using the proposed set of visualizations and analytic metrics. Details on this characterization are provided below.

The user can then use the results of this characterization to extract and apply information for either the optimization problem itself or to the search method. Problem-oriented interactions focus on assessing the impacts of different problem formulations, revealing underlying challenges or benefits of optimization model design. Problem attributes such as decision variable construction, objectives, or constraints can be modified based on iterative formulation comparisons. Search-orientated interactions focus more on determining how the attributes of the fitness landscape can impact the efficacy of the search methodology; this information can be used to find more optimal solutions. Users can apply a selection of the proposed visual tools to observe how different search methods traverse or cover the search space. They can also use the visual analytic framework to

select parameter settings or algorithm search methods that may improve search behavior based on the characteristics of the fitness landscape. The decision tree feasibility models also can be used to find regions of more optimal solutions, and users can guide the optimization to these regions.

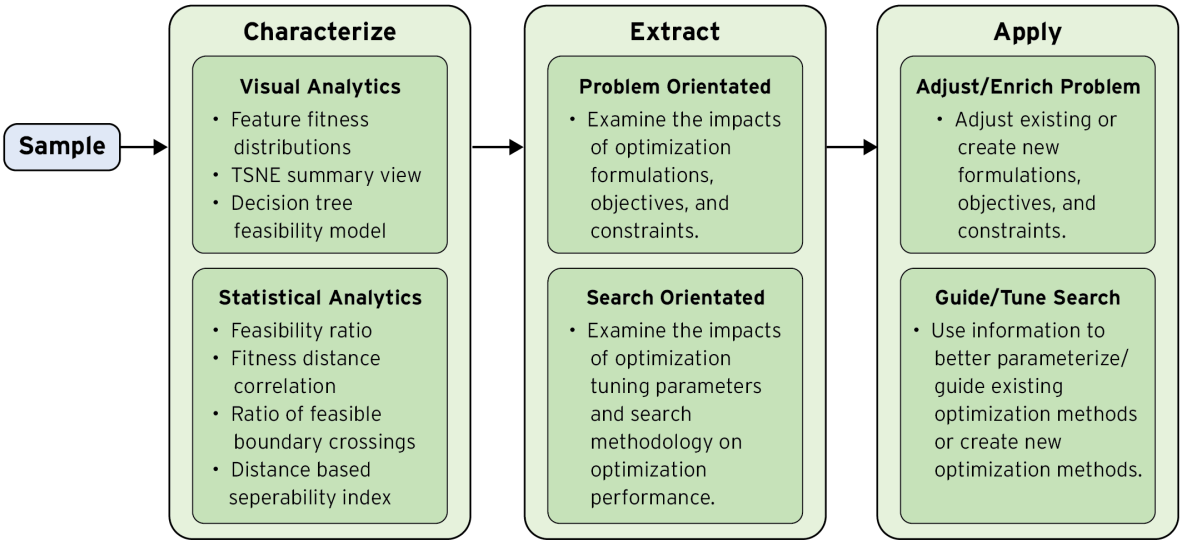

Figure 15. Proposed visual analytic framework for characterizing water distribution operation optimization fitness landscapes.

4.3.1 Visual Characterization

We incorporate three visualizations from the visual analytic system presented in Li et al. (2021b) to characterize and analyze the fitness landscape. For reference, we show the full visual analytic system in Figure 16.

First, we use the feature distribution view. This visualization displays the distribution of input settings for each feature using a stacked histogram. Each feature is a different decision variable in the optimization problem. Colors encode the optimization results, showing a single color for infeasible solutions (with constraint violations) and the range of optimal values based on a colormap. This color encoding is repeated across all subsequent visualizations.

Next, we include the simulation overview. This visualization uses T-SNE (van der Maaten and Hinton, 2008) to map the high-dimensional data presented in the features distribution view into a

2-dimensional plot. The distance between points represents feature values similarity. This view allows users to see overall patterns within the data and identify clusters of infeasible or highly optimal solutions.

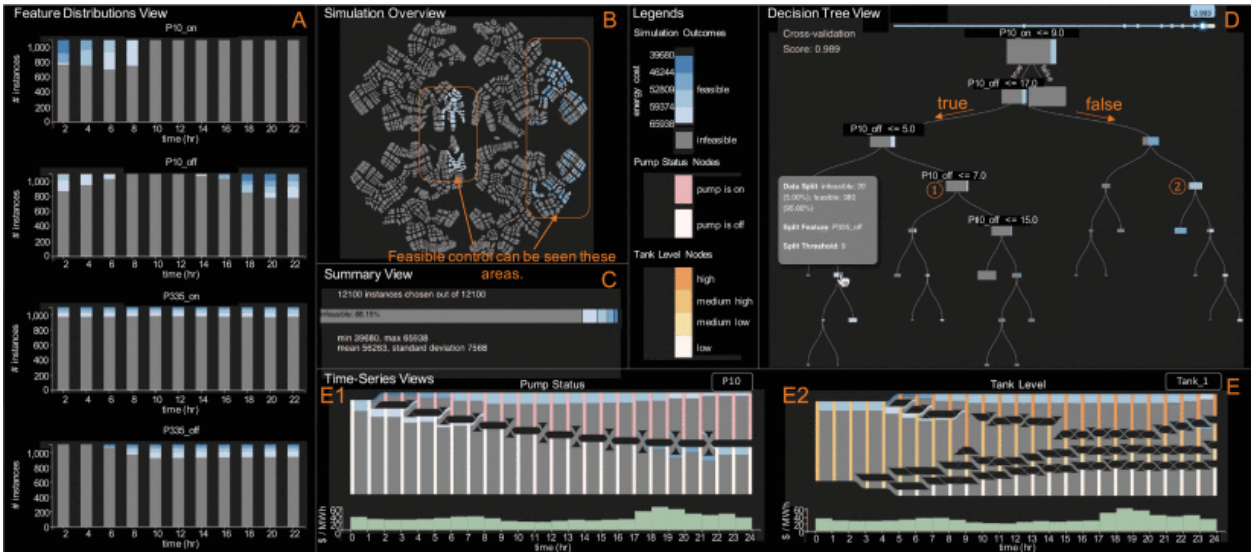

Figure 16. The interface of the visual analytics system for WDS optimization. The system is composed of (A) the feature distributions view, (B) simulation overview, (C) summary view, (D) decision tree view, and (E) time-series views.

Finally, we apply the decision tree view. This view uses a decision tree classification model (Breiman et al., 2017) trained on the sampling data to identify combinations of input values likely to yield feasible solutions and guide further optimization. To ensure the rules are interpretable, we apply cost-complexity pruning to reduce the tree size while optimizing prediction accuracy (Breiman et al., 2017). Cost complexity pruning, is post-training pruning algorithm that generates a series of trees from the original decision tree model, iteratively removing the weakest-link branch, with the final tree being the root. For our analysis, we use the smallest tree that provides a prediction accuracy greater than the feasibility ratio threshold FR_{thr} , and the colors of each decision tree nodes represent the percent of samples at that node that are feasible.

In this study, we do not include the summary or time-series views. Although these visualizations can aid deeper understanding of the data in the interactive visual analytics system

presented in Li et al. (2021b), we focused our analysis on visualizations that provide the most actionable information.

4.3.2 Analytic Characterization

For the analytic characterization of the search space, we calculate several fitness and data analysis metrics on the sample sets of the optimization fitness function.

The feasibility ratio (FR) estimates the feasible percent of the search space (samples with no constraint violations) (Malan et al., 2015). Given a sample of size n, the FR is defined as:

$$FR = \frac{n_f}{n} \tag{1}$$

where n_f is the number of samples that are feasible.

It is important in characterizing the fitness landscape to also assess how intermixed feasible and infeasible regions are. To estimate this intermixing, we use the ratio of feasible boundary crossings (RBFC) (Malan et al., 2015)

The RBFC calculates the fraction of points within a random walk of the search space that cross from a feasible to an infeasible fitness value:

$$RBFC = \frac{\sum_{i=1}^{n-1} cross(i)}{n-1} \tag{2}$$

$$cross(i) = \begin{cases} 0 \; if \; class(x_i) = class(x_{i+1}) \\ 1 \; if \; class(x_i) \neq class(x_{i+1}) \end{cases} \tag{3}$$

where n is the total number of samples, $class(\,)$ is the classification of the sample as feasible or infeasible, and x_i is the i^{th}sample.

Finally, the fitness distance correlation (FDC) provides a global measurement of the correlation between fitness values and distance to the optimum (Jones and Forrest, 1995):

$$FDC = \frac{\frac{1}{n}\sum_{i=1}^{n}(f_i - \bar{f})(d_i - \bar{d})}{\sigma_F \sigma_D} \quad (5)$$

where f is the fitness of the point x, d is the distance between point x and the best solution, σ_F is the variance in the fitness values and σ_D is the variance of the distance values. We use the best-found solution in the sample set as the optimum for this application.

FDC gives a general indication of problem difficulty, with values near to one signifying more correlation between location and fitness, allowing for easier optimization and values near to zero indicating little correlation and increased difficulty.

4.3.3 Optimization Test Problem Formulation

For our optimization problem test cases, we compared two decision variable formulations: the secondary time-based (STB) control (Musabandesu et al., 2022) and the status-time pump schedule control (STPS) (Bagirov et al., 2013). By contrasting these formulations, we examine how the visual and analytical characterization of the fitness landscape can aid in problem-oriented analysis.

The STB control structure preserves the original fixed tank level controls and then adds secondary time-based controls that can be optimized based on time-dependent energy tariffs. This allows water systems to maintain storage reserves while responding to price incentives. The STPS removes the original controls and replaces them with exclusive time-based controls. The decision variables in both formulations are the times of day that the pump is triggered on or off in either the exclusive or second priority time-based rule. For each pump, p, for the given set of pumps to be optimized, P_{opt}, the pump control decision variables are:

$$\{x_{p,\ on}, x_{p,\ off}\}, \quad x \in [0,23] \quad (6)$$

where $x_{p,on}$ is the time of day to turn pump p on, and $x_{p,off}$ is the time of day to turn pump p off.

4.3.3.1 Objective Value

For this study, we optimized energy costs under a time-variant energy rate. In particular, we used the average hourly prices from the California day-ahead wholesale electricity market in 2020 (CAISO 2020).

We calculated energy consumption from time-series hydraulic measurements using WNTR (Katherine A Klise et al., 2017; Katherine A. Klise et al., 2017), a python package that simulates water distribution operations utilizing the EPANET hydraulic simulator (Rossman et al., 2020). The objective function, or fitness function, is defined as the total pumping energy cost over the simulation period:

$$\text{minimize } C_E = \sum_{p=1}^{n_p} R_t \sum_{t=1}^{n_t} \mathcal{P}_{p,t} \Delta t \;=\; \sum_{p=1}^{n_p} R_t \sum_{t=1}^{n_t} \frac{\gamma Q_{p,\,t} H_{p,\,t}}{\eta_p} \Delta t \tag{7}$$

where C_E is the total cost of energy for pumping, p is the pump index, n_p is the total number of pumps, t is the time interval index, n_t is the total number of time intervals, R_t is the energy rate at time t, $\mathcal{P}_{p,t}$ is the power for pump p at time interval t, γ is the specific weight of water, $Q_{p,t}$ is the flow in pump p at time interval t, $H_{p,t}$ is the head in pump p at time interval t, η_p is the efficiency of pump p, and Δt is the time interval duration (typically 1 hour). Energy cost is related to the decision variables because when the pump is off, no flow is supplied by the pump during that time period.

4.3.3.2 Constraints

To maintain the hydraulic reliability of the WDS, the optimization model is subject to several constraints. Hydraulic constraints, such as conservation of mass and energy, are enforced by the simulation model. Operational constraints are incorporated into the optimization formulation by penalizing the objective function. For this study, we require that the proposed controls provide similar reliability to the initial controls; they must maintain minimum pressure, maximum velocity, and minimum/maximum tank levels within a given tolerance:

$$min(P_{n,t}) \geq min(P^0{}_{n,t}) - \varepsilon_P \quad \forall n, t \tag{8}$$

$$max(V_{l,t}) \leq max(V^0{}_{n,t}) + \varepsilon_V \quad \forall l, t \tag{9}$$

$$min_s(L_{s,\,t}) < min\{min_s(L^0{}_{s,\,t} - \varepsilon_{\mathrm{SM}}),\, S_{\min level} + \varepsilon_{\mathrm{SM}}\} \qquad \forall s \in S \tag{10}$$

$$max_s(L_{s,\,t}) < max\{max_s(L^0{}_{s,\,t} - \varepsilon_{\mathrm{SM}}),\, S_{\max level} - \varepsilon_{\mathrm{SM}}\} \qquad \forall s \in S \tag{11}$$

where $P_{n,t}$ is the pressure at junction n for time t, $P^0{}_{n,t}$ is the pressure at junction n for time t for the original control set, ε_{P} is the allowable pressure difference, $V_{l,t}$ is the velocity at pipe l for time t, $V^0{}_{l,t}$ the velocity at pipe l for time t for the original control set, and ε_{V} is the allowable velocity difference. $L_{s,\,t}$ is the level at tank s for time t, $L^0{}_{s,\,t}$ is the level at tank s for time t for the original control set, $S_{\min level}$ is the minimum water storage level for tank s, $S_{\max level}$ is the maximum water storage level for tank s, and $\varepsilon_{\mathrm{SM}}$ is the allowable maximum and minimum tank storage level difference.

In addition, we include a constraint that controls the average tank level over the simulation period so the system maintains water storage for emergency and fire supplies:

$$\frac{\sum_{t=1}^{n_t}(L_{s,\,t})}{T \cdot S_{\max level}} > \frac{\sum_{t=1}^{n_t}(L^0{}_{s,\,t})}{T \cdot S_{\max level}} - \varepsilon_{\mathrm{SA}} \tag{12}$$

where T is the total simulation period, $\varepsilon_{\mathrm{SD}}$ is the allowable percent difference for the differencing storage constraint, and $\varepsilon_{\mathrm{SA}}$ is the allowable percent difference for the average storage constraint.

4.3.3.3 Water Distribution System Case Study

We tested our visual analytic framework on the Net 3 water system, a case study often used for benchmarking optimization methods (Rossman et al., 2020). This system is fairly simple, with two pumps (P10 and P335), three tanks, 117 junctions, and 91 pipes. For use with the STB optimization,

the Net 3 model was modified so all initial controls for all three systems follow a tank trigger control scheme, where pumps are turned on and off based on specified tank levels.

4.3.4 Existing Optimization Methods

We also used our system to compare the performance of three heuristic search-based optimization methods on the two problem formulations: genetic algorithm (GA), particle swarm optimization (PSO), and the covariance matrix adaptation evolution strategy (CMA-ES). GA evolves a population of candidate solutions using the biologically inspired mechanisms, selecting individuals with good performance and using these to create new candidates (Bäck et al., 2018). The PSO algorithm is inspired by the movement of a flock of birds or a school of fish; the position and velocity of the candidate solutions, or particles, is updated based on each particle's best-known location and the global best-known location (Clerc and Kennedy, 2002). CMA-ES draws candidates from a sequence of multivariate gaussian distributions, where the centroid is the initial mean and sigma is the initial standard deviation (Hansen and Ostermeier, 2001). Additional hyperparameters control iterative changes to the distribution. Because these algorithms have a range of search mechanisms, we can use the visual analytic framework to analyze the effectiveness of different search characteristics on the various case studies' fitness functions.

All three algorithms were implemented using DEAP, a computational framework for constructing evolutionary algorithms in Python (Fortin et al. 2012). For GA, we used the simple form (Bäck et al., 2018), which includes selection, crossover, and mutation, specifying three-way tournament selection, one-point crossover, and uniform random mutation. We used the constricted coefficient strategy outlined in Clerc and Kennedy (2002) to implement PSO, using the Type 1" constriction particle swarm. For the CMA-ES algorithm, we used the weighted recombination form (Hansen and Ostermeier, 2001). Weights help regulate the algorithm's convergence and, in the DEAP implementation, follow either an equal, linear, or superlinear decreasing scheme.

We compared the performance of each optimization method using typical optimization parameter value settings against the method's performance with guidance and tuning parameters informed by the visual and analytic framework. The initial parameter settings appear in Table 11. For the CMA-ES parameter values not itemized in Table 11, we used default values specified in DEAP documentation (Fortin et al., 2012). The combinations of population size and maximum number of generations were selected so each algorithm performed a similar number of function evaluations.

Table 11: Initial Optimization Parameter Values for Existing Methods

Algorithm	**Parameter**	**Variable**	**Value**
Genetic Algorithm	Crossover rate (ratio)	α_{cx}	0.8
	Mutation rate (ratio)	α_{mut}	0.3
	Independent probability for each attribute to be mutated	α_{indmut}	0.2
	Population size	NP	155
	Maximum number of generations	G_{max}	100
Particle Swarm Optimization	Cognitive coefficient	φ_p	2.05
	Social coefficient	φ_g	2.05
	Population size	NP	80
	Maximum number of generations	G_{max}	125
Covariance Matrix Adaptation Evolution Strategy	Centroid	μ	A random number drawn from the uniform distribution
	Sigma	σ	3.0
	Weight strategy	w	Superlinear
	Population size	NP	100
	Maximum number of generations	G_{max}	100

For the informed optimization run, we used the best n_{ws} policies to seed the optimization method with warm solutions. Seeding optimizations with warm solutions improves optimization performance (Savic et al., 1997). We also drew n_{DT} candidates of the initial population from a region of the decision search space found by the decision tree to have a high probability of feasible and optimal solutions. We trained a decision tree on the smaller characterization sample set (2,000 samples) to find this target region and then pruned the tree using cost-complexity pruning

(Breiman et al., 2017). On the pruned tree, we located $node_{target}$, the node with the minimum objective value and a feasibility ratio greater than the feasibility threshold, FR_{thr}. The rules from this node were extracted to create the bounds for the target region.

4.3.5 Proposed Decision Tree Genetic Algorithm

In addition to testing the application of extracted information from the visual analytic framework on a set of existing optimization methods, we propose a new optimization method that automatically integrates some of the optimization guidance methods into a heuristic search. The algorithm employs a decision tree classifier to identify regions with a high probability of feasible and low objective values and use rules extracted from this decision tree to guide the genetic algorithm. Similar methods that incorporate machine learning models into heuristic searches have been developed, including IBEA-SVM (Li et al., 2019) and the classification-based surrogate-assisted evolutionary algorithm (Pan et al., 2019).

The proposed method, named the decision tree guided genetic algorithm (DTG-GA), is presented in Algorithm 1. First, a characterization population of NCP candidate solutions is initialized using Latin hypercube sampling. The population is classified as infeasible and feasible. A decision tree is trained on the classified data and then pruned using cost-complexity pruning.

On the pruned tree, we locate $node_{target}$, the node with the minimum objective value and a feasibility ratio greater than the feasibility threshold, FR_{thr}. The rules are extracted for this node and used to create a new set of lower bounds (l_{DT}) and upper bounds (u_{DT}) for the mutation operator. At this point, we start the genetic algorithm. We initialize a new population of NP candidates and seed this population with n_{ws} best candidates from the characterization population.

Algorithm 1: DTG-GA

Require: NCP: the size of characterization population; NP: the size of optimization population; α_{DT}: the probability of reliance on the decision tree model; α_{mut}: mutation rate; α_{cx}: crossover rate; FR_{thr}: Feasibility ratio threshold;
n_{ws}: Number of warm solutions; Termination condition

Initialize the characterization population P_c using Latin hypercube sampling
Classify the characterization population as feasible (F) or infeasible (I)
Train a DT classification model using $\{F\}$ and $\{I\}$
Find $node_{target}$ the decision tree node with minimum objective value and a feasibility ratio greater than FR_{thr}
Use rules for $node_{target}$ to adjust lower (l_{DT}) and upper (u_{DT}) bounds of mutation to target mutation
Initialize optimization population P
Seed population with n_{ws} best solutions from characterization population
while the Termination condition is not satisfied, **do**
Find Fitness values for P if not calculated already
Generate a new offspring population P_o using tournament selection
Apply crossover function $P_o' = crossover(P_c, \alpha_{cx})$
Apply mutation function $P_o'' = mutate(P_c', \alpha_{mut}, \alpha_{DT}, l_{DT}, u_{DT})$
Classify P_o' as feasible (F) or infeasible (I)
Train a DT classification model using the full set of $\{F\}$ and $\{I\}$ from all sampled points
Find $node_{target}$ the decision tree node with the lowest average objective value for feasible solutions with a feasibility ratio greater than FR_{thr}
Use rules for $node_{target}$ to adjust lower (l_{DT}) and upper (u_{DT}) bounds of mutation
P = P_c''
end while

In the main loop, a new offspring population is generated using tournament selection, and then the crossover operator is applied to the offspring population at a crossover rate α_{cx}. Next, the mutation operator is applied at the mutation rate α_{mut}. At this point, information from the decision tree is incorporated to guide the evolution of the population. Within the mutation operator at a probability of α_{DT} the bounds l_{DT} and u_{DT} are used in place of the original bounds. The mutation occurs within the full decision variable limits for all other instances. A new decision tree is trained on the full set of data sampled up to this point, and new l_{DT} and u_{DT} are generated. Finally, the offspring population becomes the population. This loop is continued until the termination condition is met.

Optimization parameter settings for the DTG-GA are provided in Table 12. For this study, we used the maximum number of generations as the termination condition.

Table 12: Optimization Parameter Values for Decision Tree Guided Genetic Algorithm

Parameter	Variable	Value
Crossover rate (ratio)	α_{cx}	0.8
Mutation rate (ratio)	α_{mut}	0.3
Probability of reliance on DT model	α_{DT}	0.25
Independent probability for each attribute to be mutated	α_{indmut}	0.2
Feasibility ratio threshold	FR_{thr}	0.5[a]/0.9[b]
Number of warm solutions	n_{ws}	5
Population size	NP	155
Characterization population size	NCP	$3 \cdot NP$
Maximum number of generations	G_{max}	100

[a] For status-time pump schedule formulation; [b] For secondary time-based formulation

4.4 Results

4.4.1 Latin Hypercube Sampling Results

Figure 17 presents the visual characterization of the large sampling set for both the STB and STPS formulations. From the simulation overview, the STPS has a much larger rate of infeasible solutions than the STB formulation, with small clusters of feasible and optimal solutions. The STPS feature distributions view shows a large proportion of the optimal solutions require that P10 turn on in the earliest hours. We can also see from the decision tree view that the model predicts a high percentage of infeasible solutions when P10 is turned on at or after 15.4 hours or 3:24 PM. The STB formulation is much more stable with a lower percentage of infeasible regions. From the feature distributions view, it appears that P10 off has the largest influence on when samples are feasible, with a large clustering of infeasible solutions between 2 - 6 AM. The decision tree also shows there is a high prediction of feasibility when P10 is turned off at or after 7.9 hours or 7:54 AM.

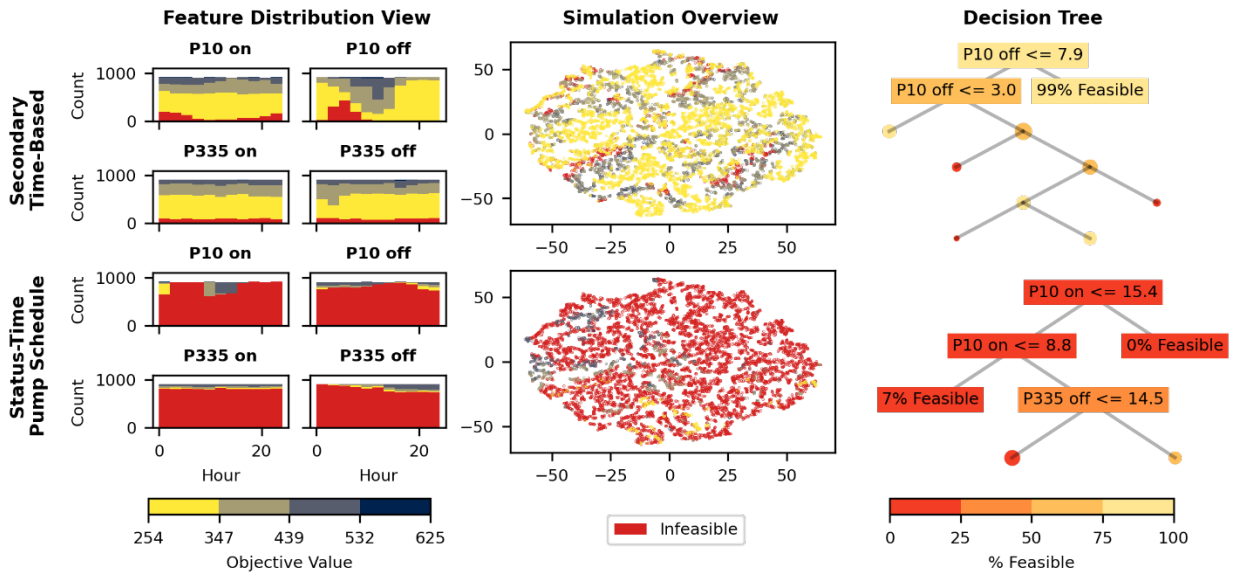

Figure 17. Latin Hypercube Sampling of the Fitness Landscape for Large Sample Set. Visual characterization of secondary time-based control problem formulations for the Net 3 water system using 10,000 samples.

Figure 18 shows the fitness landscapes for the smaller sampling set (2,000 samples). In this figure, we see similar patterns to the large sampling set. Again, the STPS formulation has a much higher rate of infeasibility, and the feasible solutions are mostly clustered with P10 turning on before 2 AM. There are slight differences between the splitting thresholds on the decision tree, but generally, the models are similar. The STB formulation shows equal similarities, with the same patterns as the large sampling set emerging for P10 within the feature distribution view. The decision tree again has a similar structure to the one trained on the larger sampling set with slight changes to the splitting thresholds. These consistent findings indicate that patterns begin to emerge even with a more limited characterization sample; users can use these patterns to inform subsequent optimizations.

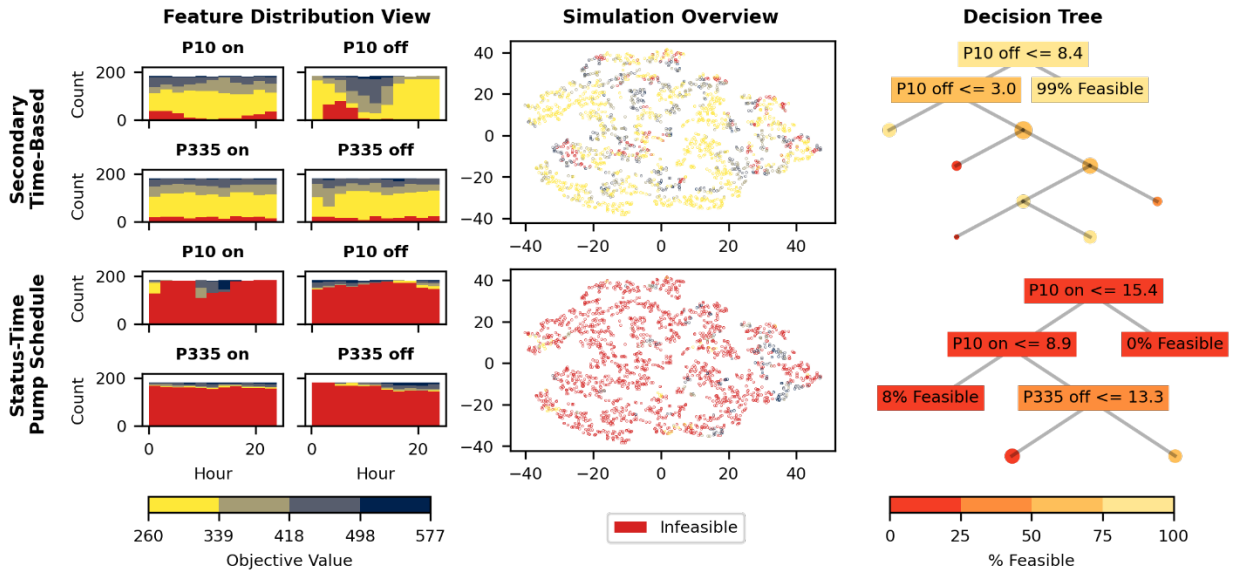

Figure 18. Latin Hypercube Sampling of the Fitness Landscape for Limited Sample Set. Visual characterization of secondary time-based control problem formulations for the Net 3 water system using 2,000 samples.

Table 13 presents the analytic characterization of the STB and STPS fitness landscapes using both larger and smaller sampling sets. The feasibility ratio is much higher for the STB formulation than the STPS formulation, as seen in Figures 17 and 18. This indicates that the STB control structure is more stable across many settings and may provide more consistent system control.

The RFBC has similar values for both control structures. This metric indicates the level of separability or intermixing between feasible and infeasible solutions. Both fitness landscapes have a fairly large ratio of crossings, with nearly 20% of random steps crossing from one feasibility state to another. This intermixing is also observable within the visualizations of the characterization sampling sets, particularly in the simulation overview plots.

Finally, the FDC is slightly smaller for the STB formulation than the STPS, signifying its fitness is more correlated with position. We can see in the simulation overview that the STPS landscape has many fewer optimal solutions, and they tend to cluster together. Conversely, the STB fitness landscape contains many low objective value solutions that range over many different decision variable values. From this, we may infer that because we do not have a large number of decision

variables for the test WDS, it may be easier to find the best solution for the STPS. However, given the low rate of feasible options, the small clusters of optimal solutions may be difficult to locate for a system with more decision variables.

Table 13: Analytic Characterization of Fitness Landscapes

Metric	Case Study Control Type	Sample Size	Value
Feasibility Ratio	Secondary Time Based	10,000	0.90
		2,000	0.90
	Status Time Pump Schedule	10,000	0.11
		2,000	0.12
Ratio of Feasible Boundary Crossings	Secondary Time Based	10,000	0.18
		2,000	0.18
	Status Time Pump Schedule	10,000	0.20
		2,000	0.21
Fitness Distance Correlation	Secondary Time Based	10,000	0.11
		2,000	0.11
	Status Time Pump Schedule	10,000	0.18
		2,000	0.08

For all metrics except the FDC for the STPS case studies, there is minimal difference between the metric values calculated from the small and large sample sets. This consistency again indicates that we can gather information about an optimization problem's fitness landscape from even a limited dataset.

4.4.2 Informed Optimization Parameterization

Using information from the visual and analytic characterization of the search space sampling sets, we can infer what parameter settings might be more appropriate based on the behavior of each algorithm. Our decision to guide the optimizations by seeding the population with good initial solutions also influences our parameter settings because we know some portion of the population already has better objective values. Table 14 provides parameters values for the informed optimization runs selected based on details extracted from the fitness landscape characterization.

Table 14: Informed Optimization Parameter Values

Algorithm	Parameter	Variable	Value
Genetic Algorithm	Crossover rate (ratio)	α_{cx}	0.8
	Mutation rate (ratio)	α_{mut}	0.6
	Independent probability for each attribute to be mutated	α_{indmut}	0.2
	Population size	NP	155
	Number of warm seeds	n_{ws}	5
	Number of individuals drawn from the best region in the decision tree	n_{DT}	$0.2 \cdot NP$
	Maximum number of generations	G_{max}	100
Particle Swarm Optimization	Cognitive coefficient	φ_p	0.5
	Social coefficient	φ_g	3.6
	Population size	NP	80
	Number of warm seeds	n_{ws}	5
	Number of individuals drawn from the best region in the decision tree	n_{DT}	$0.2 \cdot NP$
	Maximum number of generations	G_{max}	125
Covariance Matrix Adaptation Evolution Strategy	Centroid	μ	The best performing individual from the characterization
	Sigma	σ	3.0
	Weight strategy	w	Superlinear
	Population size	NP	100
	Maximum number of generations	G_{max}	100

For the genetic algorithm, because the initial population has been seeded with warm solutions, a high probability of mutation and crossover allows these methods to explore how different settings might improve the initial population. The PSO algorithm uses the social and cognitive coefficients to pull particles towards global or local best solutions (Clerc and Kennedy, 2002). Because the particles are distributed throughout the search space, and changes between particle iterations are incremental, we set the social coefficient higher to give it more influence to pull solutions into the optimal region. To guide the CMA-ES, we set the centroid as the minimum value found in characterization. The CMA-ES algorithm can quickly converge if it finds a set of solutions that appear to be a minimum; to force it to explore more, we increase sigma (Hansen and Ostermeier, 2001).

4.4.3 Optimization Results

Table 15 summarizes the final optimization results for the existing optimization algorithms and the new optimization method DTG-GA.

For the existing algorithms, the genetic algorithm performed well across all cases, especially when informed using the warm solutions and samples drawn from the predicted optimal region. The crossover and mutation strategies allowed good solutions to evolve more efficiently. This consistent performance is the main reason we chose GA as the base for our proposed method DTG-GA. Although PSO covered a larger range of the search space, in general, the GA found better solutions. The guidance improved the algorithm's performance in the STPS search space, but it did not benefit the optimization of the STB controls.

The CMA-ES performed similarly to GA for the STPS formulation. Again, the informed version had no performance increase. Since many of the optimization runs for STPS converged on this solution, this may be the global minimum. Conversely, CMA-ES performance was the poorest of the three algorithms with the STB formulation. This poor performance is likely because CMA-ES tends to converge on a minimum quickly, getting stuck in local optima, and the STB fitness landscape is multimodal with many different clusters of low objective value solutions. Employing a restarting strategy can help the CMA-ES find solutions when there are many local minimums; but, this strategy requires a larger computational budget (Auger and Hansen, 2005).

Our proposed method, the DTG-GA, performed nearly as well as the informed GA for the STB formulation, with a much smaller computational budget for the initial characterization population. DTG-GA also performed as well as GA and CMA-ES on the STPS problem. From these results, we infer that incorporating an initial characterization and incorporating decision tree models to either inform the initial population or as an integral part of the optimization routine can improve optimization performance.

Table 15. Optimization Results

Algorithm	Control Formulation	Version	Objective Value
Genetic Algorithm	Secondary Time Based	Standard	260.94
		Informed	**224.72**
	Status Time Pump Schedule	Standard	260.19
		Informed	260.19
Particle Swarm Optimization	Secondary Time Based	Standard	260.93
		Informed	260.93
	Status Time Pump Schedule	Standard	285.07
		Informed	260.20
Covariance Matrix Adaption Evolution Strategy	Secondary Time Based	Standard	290.28
		Informed	263.40
	Status Time Pump Schedule	Standard	260.19
		Informed	260.19
Decision Tree Guided	Secondary Time Based	Standard	**225.94**
	Status Time Pump Schedule	Standard	260.19

4.4.4 Visual Analysis of Optimization Coverage

We can use a subset of the visualizations to analyze each algorithm's coverage of the optimization search space to understand how the behavioral characteristics of each algorithm interact with the fitness landscape for each problem formulation and influence optimization performance. In particular, we used the feature distribution view and the simulation overview to visualize the search range of each optimization method. This visual analysis provides insights into the challenges or benefits of different approaches and shows the multifaceted use of these visual analyses.

Figure 19 presents a visual analysis of the GA trajectory or coverage of each fitness landscape. Generally, GA spends most of simulations testing combinations within a narrow range of values. Although the vast majority of samples tested across both the STB and STPS formulations are feasible, there is clear intermixing between the feasible and infeasible solutions within the range of values sampled. The informed version spends more time testing a larger range of values, particularly for the STB formulation. When comparing method performance, the informed GA

performed best, which indicates that evolving good solutions using the mutation and crossover strategies to diversify the population was effective for the two fitness landscapes presented.

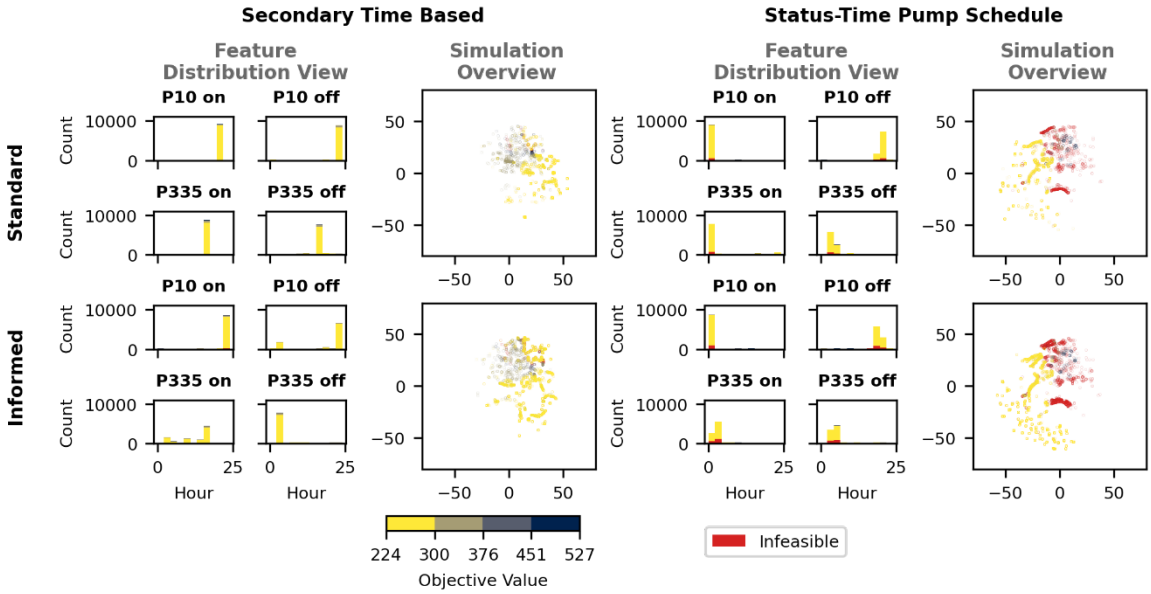

Figure 19. Visual Analysis of the Genetic Algorithm Search Coverage Comparing Standard Optimization and Informed Optimization. Applied to both the secondary time-based control and the status-time pump schedule case studies.

From the visual analysis, the PSO algorithm spends many more simulations across a larger range of values (Figure 20). For the STPS formulation, a very high proportion of the sampled instances are infeasible. The informed version for the STPS formulation more clearly targets a region of optimal samples than the standard version. This trend is also visible in the STB formulation; the feature distributions show higher proportions of the samples targeted within tighter ranges. Although this algorithm showed good coverage of the fitness landscape, this did not translate into an increase in optimization performance. Once good solutions were located, PSO did not test enough variations of these solutions to escape the local minimum. PSO might perform better for these landscapes if combined with a local search method.

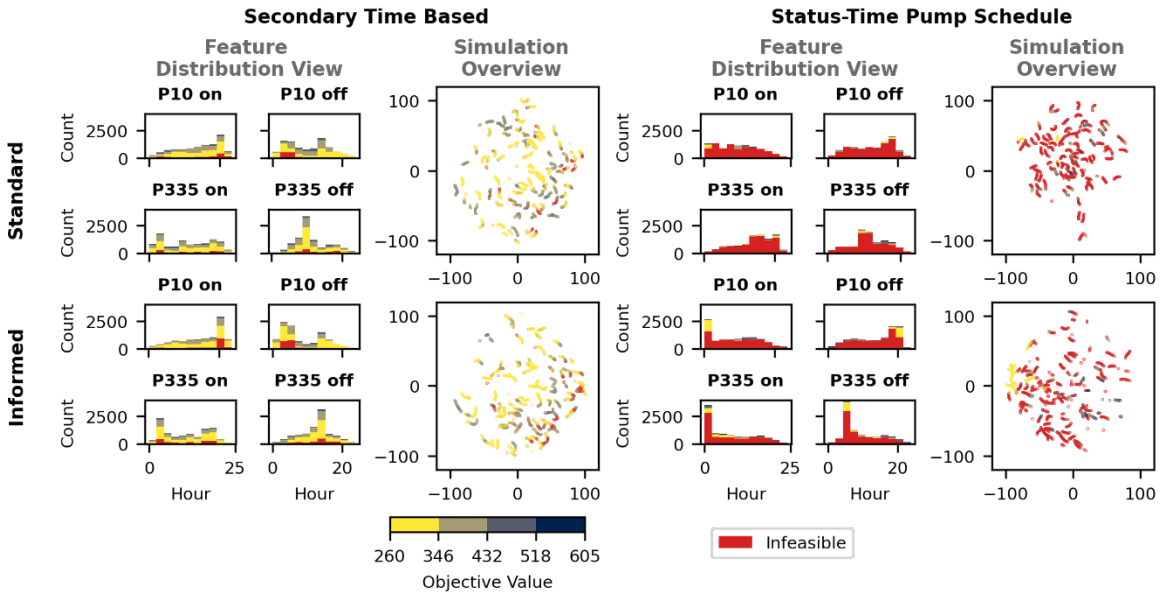

Figure 20. Visual Analysis of the Particle Swarm Optimization Search Space Comparing Standard Optimization and Informed Optimization. Applied to both the secondary time-based control and the status-time pump schedule case studies.

Like GA, the CMA-ES focuses its search on a small range of values (Figure 21). In the CMA-ES coverage, for both formulations, many simulations have low objective values, with a minimal number of infeasible and high objective solutions. For the informed optimization of the STB case, the algorithm focused on a different and narrower range of solutions than the standard algorithm. Although the information improved the performance of CMA-ES on the STB landscape, both versions settled on local minima and were not able to perform as well as GA. For the STPS formulation, there is more overlap in the coverage between the standard and informed versions. This similarity in coverage indicates that the guidance did not seem to influence the CMA-ES trajectory in the STPS fitness landscape. This again supports the conclusion that CMA-ES performs well when the function has a more distinct global minimum.

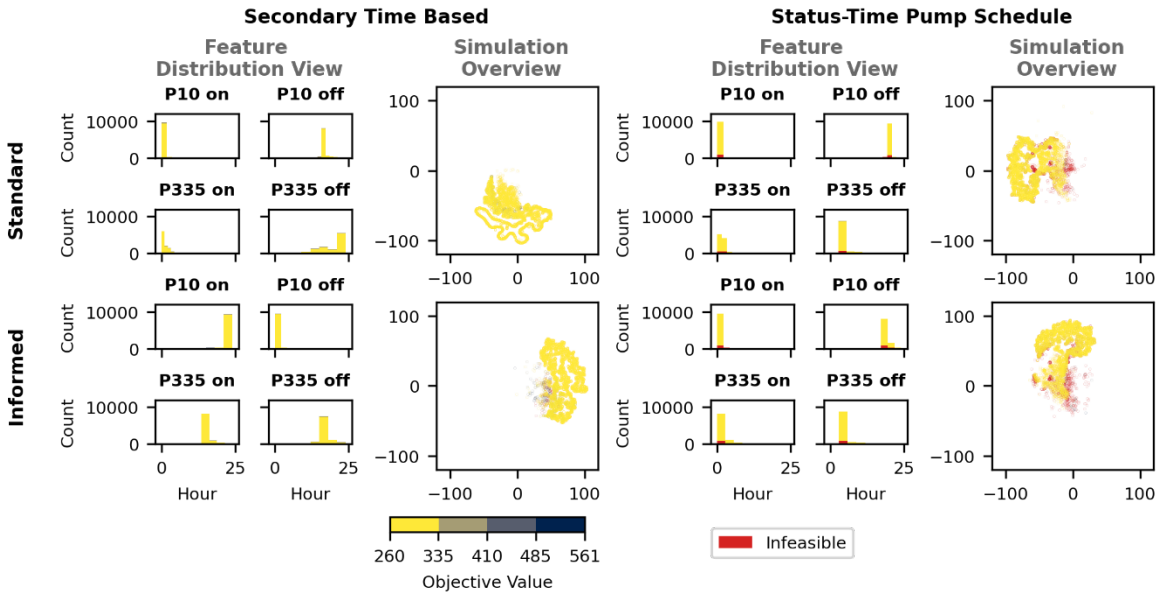

Figure 21. Visual Analysis of the Covariance Matrix Evolutionary Strategy Search Space Comparing Standard Optimization and Informed Optimization. Applied to both the secondary time-based control and the status-time pump schedule case studies.

Figure 22 visualizes the DTG-GA coverage of the STB and STPS fitness landscapes. DTG-GA leverages fitness function characterization to guide the optimization using similar search mechanisms to GA. For the STB formulation, we can observe that the DTG-GA optimization generally targeted a small range of values, as seen with the GA. If we compare the search range more closely, the algorithm was drawn into similar regions as the informed GA optimization, which allowed it to perform comparably well even with a smaller characteristic population. One difference is that this algorithm more clearly focused on solutions that turned P10 off earlier in the day. The similarities between the GA and DTG-GA search are also apparent in the STPS case. Both algorithms could quickly find the minimum value and focus on feasible options in a highly infeasible search space.

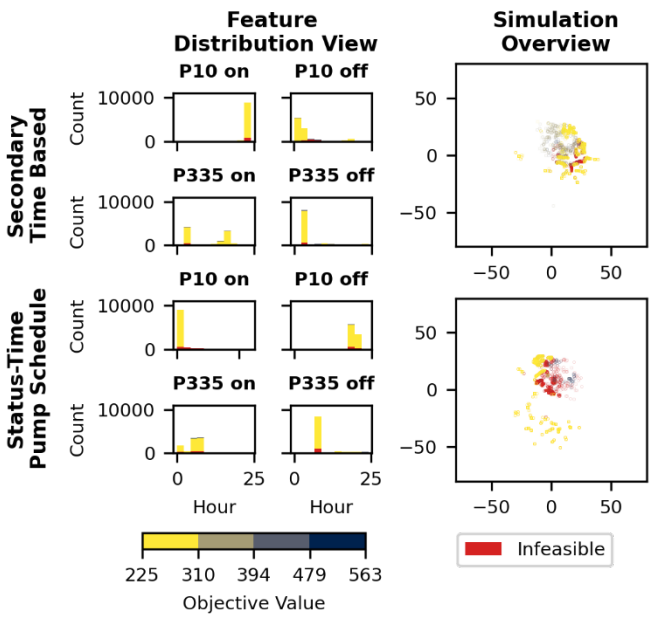

Figure 22. Visual Analysis of the Decision Tree Guided Genetic Algorithm Search Coverage. Applied to both the secondary time-based control and the status-time pump schedule case studies.

4.5 Conclusions

This study presents a visual analytic framework to characterize WDS operation optimization fitness landscapes. The framework allows the user to examine the impacts of problem formulation and search method on optimization effectiveness and extract information to improve the problem formulation and the optimization search.

We demonstrated this framework on a simple WDS network, examining the fitness landscapes of two different pump control structures. The framework was effective in analyzing and visualizing the differences between problem formulations. The STB formulation yielded a very different fitness landscape than the STPS formulation, with a much higher percentage of feasible and low objective value solutions. Because of this higher rate of feasibility, the STB control structure is more stable across many settings and may provide more consistent control of the system.

We also demonstrated the use of this framework to extract actionable information to select, guide, or create new optimization methods to improve optimization search performance. We tested

three existing methods, GA, PSO, and CMA-ES, examining their performance with and without guidance from the initial characterization.

To inform the existing methods, we seeded the initial optimization population with warm solutions extracted from the characterization and samples drawn from predicted feasible and low objective value regions identified by decision tree models. We also used landscape fitness characteristics to select tuning parameters for each algorithm. The efficiency of optimization methods generally improved when informed using these strategies, particularly in the STB fitness landscape, which has many local optima. Overall, the genetic algorithm performed the best across the two fitness landscapes.

Additionally, we presented a novel optimization method, DTG-GA, that incorporates fitness landscape characterization and decision tree machine learning to guide a genetic algorithm to regions with high feasibility and optimal solutions. This optimization method performed similarly well to the informed genetic algorithm with a much smaller characterization population. By integrating machine learning into the search algorithm, the decision tree model could iteratively guide the optimization into high-performing areas.

Although we designed this framework and the corresponding optimization method for use in the WDS operation optimization problem, many of the techniques presented can be extended to other complex optimization problems, particularly problems solved using heuristic or metaheuristic search techniques.

4.6 Data Availability

All data, models, or code generated or used during the study are available in a repository or online in accordance with funder data retention policies.

4.7 Acknowledgments

Special thanks to the California Energy Commission for financing this study with funds from the Electric Program Investment Charge, the ratepayer surcharge authorized by the California Public Utilities Commission, through agreement number EPC-16-062. Special thanks to the MNWD and its staff for their excellence in supporting this and other research into the water-energy nexus and energy efficiency in the water sector.

This document was prepared as a result of work sponsored by the California Energy Commission. It does not necessarily represent the views of the Energy Commission, its employees, or the State of California. The Energy Commission, the State of California, its employees, contractors, and subcontractors make no warranty, express or implied, and assume no legal liability for the information in this document; nor does any party represent that the use of this information will not infringe upon privately owned rights. This report has not been approved or disapproved by the Energy Commission, nor has the Energy Commission passed upon the accuracy of the information in this report.

5 Chapter 5: Conclusions

The research presented in this dissertation focused on enabling the water sector to respond to time-based energy incentives to support renewable integration. From these investigations, we can summarize several main contributions and overarching conclusions.

First, we identified opportunities within the water sector that allow utilities to shift energy use with little or no impact on water quality or water service performance goals. For wastewater facilities, we showed that flow equalization basins or combined heat and power engines could be used to shift energy with minimal impact on wastewater quality. We also provided a wastewater treatment plant roadmap outlining the instrumentation, automation, and infrastructure necessary to participate in demand resource programs. For water distribution utilities, we developed a new

optimization formulation that allows these systems to optimize pump operations based on time-dependent energy incentives while maintaining storage reserves and minimizing the violation of system constraints under uncertain demands.

Although the research we presented can support water utilities performing energy demand management, the design of the energy incentive itself can be a barrier to water sector participation. Incentives need to be accessible and provide adequate compensation to encourage response. In particular, we identified significant challenges for water utilities participating in demand response programs that use averaging baseline methodologies to measure demand reductions and award compensation. Averaging baseline methodologies do not accurately estimate energy patterns for wastewater facilities, making it difficult to measure the impacts of operational modifications on the energy load.

Through our research, we also developed procedures to analyze optimization problem formulations and the performance of optimization methods based on fitness function characteristics. In Chapter 3, we analyzed the simulations tested during the optimization search, determining what percent of the simulations resulted in feasible solutions and which constraints were violated for infeasible solutions. This analysis gave us insights into how model formulation can affect the fitness function characteristics, which constraints were more sensitive to changes in formulations, and how stable formulations are across various settings. In Chapter 4, we presented a more formal set of procedures as part of a visual analytic framework designed to characterize optimization fitness functions. This characterization can be used to examine the impacts of objective model design, modifying problem attributes, including objective values, constraints, or decision variable formulations. The visual analytic framework can also be used to improve the search efficiency of optimization methods, using characterization to guide, tune, or create new methods.

There are several areas of research that could expand the work presented in this dissertation. Future work should investigate additional load-shifting strategies for WWTPs that have minimal impacts on water quality, particularly focusing on developing techniques for load-shifting processes that are more sensitive to operational changes. Similar approaches could be explored for other types of water treatment plants. Researchers could also examine how modeling and dynamic data could be leveraged to correctly time and predict the effects of operational changes.

Our research into performing energy demand management at water distribution facilities could also be expanded. Researchers could explore the iterative optimization of hierarchal rules, first optimizing tank-based rules to maximize reliable system performance under uncertain demands and then optimizing secondary time-based controls to reduce energy costs. It would also be valuable to analyze how different system characteristics can impact a utility's ability to shift energy, including water storage redundancy, system elevation distributions, and pumping capacity. Studies implemented at full-scale facilities could identify additional barriers to energy demand management as well.

There are also many opportunities to improve optimization formulations and methods using insights from our visual analytic framework. This research is not limited to the optimization of water systems but can be applied to a broad range of complex optimization problems. This framework should serve as a starting point for the visual and analytical characterization of optimization search spaces. Researchers may continue to develop new visualizations and landscape fitness metrics to improve our understanding of optimization problems and optimization method efficiency.

Appendix A: Supplemental Information for Chapter 3

Additional information is provided in this document to supplement the manuscript, including adjustments to the original NET 3 case study EPANET file, optimization details, and the optimization results.

Case Study Amendments

A set of controls for one of the pumps in the Net 3 model (Rossman et al., 2020) was modified to follow a tank trigger control scheme rather than a time-of-day trigger. The new controls were constructed to result in similar operations to the original control. Both the original and modified controls are provided below:

Original Controls:		**Modified Controls:**
Link 10 OPEN AT TIME 1	→	Link 10 OPEN IF NODE 2 BELOW 21.5
Link 10 CLOSED AT TIME 15		Link 10 CLOSED IF NODE 2 ABOVE 29

Optimization Details

Further optimization details are provided in the following sections, including parameter values for the operational constraints and the control setting values used to seed the initial population of the particle swarm optimization (PSO) for each control structure type.

Operational Constraint Parameter Values

The full constraint formulations are provided in the manuscript. Table 16 provides the allowable difference parameter values that were used in this study for each constraint.

Table 16. Operational Constraint Parameter Values

Parameter	Description	Value
ε_P	Allowable pressure difference	5 m
ε_V	Allowable velocity difference	2 m/s
ε_{SM}	Allowable maximum/minimum tank storage level difference.	0.05
ε_{SD}	Allowable percent difference for the differencing storage constraint	0.1
ε_{SA}	Allowable percent difference for the average storage constraint	0.05 m

Initial Seeding Particles

To assist the optimization in finding solutions that do not violate operational constraints, the initial population for each optimization was seeded with a single feasible solution (Savic et al., 1997). All other particles in the initial population were selected randomly.

Each control structure has a different initial seed for each case study. The tank trigger settings from the original controls were used as seeds for the tank trigger control structures. Table 17 provides the settings of the initial seeds for the secondary time-based and explicit pump schedule controls.

Table 17. Initial Feasible Population Seed for Secondary Time-Based and Explicit Pump Schedule Controls

Network	Pump Name	Explicit Pump Schedule		Secondary Time-Based Control	
		On Time	Off Time	On Time	Off Time
Net 3 RU	10	2	16	11	16
	335	22.5	5	11	16
Richmond Skeletonized[a]	1A	0	0	11	16
	2A	7.1	3.7	11	16
	3A	8.3	1.25	11	16
	$5C_{SW1}$	15	4	11	16
	$5C_{SW2}$	18.2	6	--	--
	$6D_{SW1}$	8.7	3.33	11	16
	$6D_{SW2}$	5.8	23.5	--	--
	7F	19	21.1	11	16
MNWD	1	21.5	14	17	21
	2	0.4	13.5	17	21
	3	1	11	17	21
	4	21.2	8.7	17	21
	5	1	3	17	21
	6	0.6	5.4	17	21
	7	13.5	13.5	17	21
	8	21.5	11.2	17	21

[a]Pumps 5C and 6D require two sets of pump switches to cycle tanks at a similar rate to the original system. All other pumps only require a single set of pump switches.

Optimization Results

Three replications were performed for each optimization case study utilizing different random seeds. The manuscript reports the average values for each parameter of interest; Table 3 provides full results for each replication.

Table 18. Full Optimization Results for Optimization Run Replications

Network	Evaluation Type	Tank Constraint Type	Random Seed	Objective Value ($)	Number of Pump Switches	Average Storage (%)
Net 3	Secondary Time Based	Averaging	0	1086.27	18	71.1
			1	1086.27	18	71.1
			2	1086.29	18	71.1
		Differencing	0	1086.26	18	71.2
			1	1086.31	18	71.1
			2	1086.30	18	71.1
	Explicit Pump Schedule	Averaging	0	1312.98	26	73.8
			1	1242.81	14	73.1
			2	1247.02	27	76.8
		Differencing	0	1183.02	28	73.5
			1	1183.47	28	72.6
			2	1183.12	28	73.2
	Tank Trigger	Averaging	0	1093.10	14	71.9
			1	1098.11	15	71.6
			2	1088.12	17	71.0
		Differencing	0	1092.79	15	70.4
			1	1093.64	14	70.6
			2	1093.66	14	70.6
Richmond Skeletonized	Secondary Time Based	Averaging	0	397.49	251	86.2
			1	407.59	239	85.8
			2	399.24	240	86.2
		Differencing	0	400.10	234	86.1
			1	407.59	239	85.8
			2	401.86	249	85.8
	Explicit Pump Schedule	Averaging	0	425.38	204	84.4
			1	425.38	204	84.4
			2	425.38	204	84.4
		Differencing	0	425.38	204	84.4
			1	425.38	204	84.4
			2	425.38	204	84.4
	Tank Trigger	Averaging	0	434.83	239	85.2
			1	434.83	239	85.2
			2	434.83	239	85.2
		Differencing	0	430.99	244	79.8
			1	429.62	278	80.3
			2	427.15	281	86.2

Network	Evaluation Type	Tank Constraint Type	Random Seed	Objective Value ($)	Number of Pump Switches	Average Storage (%)
MNWD	Secondary Time Based	Averaging	0	24177.68	320	70.0
			1	24050.71	324	69.3
			2	24149.51	324	69.8
		Differencing	0	24148.92	304	69.4
			1	24017.36	336	69.4
			2	24149.51	324	69.8
	Explicit Pump Schedule	Averaging	0	25083.76	261	70.8
			1	25083.76	261	70.8
			2	25083.76	261	70.8
		Differencing	0	25101.00	261	70.8
			1	25101.00	261	70.8
			2	25101.00	261	70.8
	Tank Trigger	Averaging	0	24450.96	270	69.2
			1	24426.95	298	69.4
			2	24177.68	304	69.1
		Differencing	0	23800.96	272	61.3
			1	23558.97	270	62.8
			2	23481.26	258	61.2

Bibliography

Aghajanzadeh, A., Wray, C.P., McKane, A.T., 2015. Opportunities for Automated Demand Response in California Wastewater Treatment Facilities. California, USA.

Alvisi, S., Franchini, M., 2017. A robust approach based on time variable trigger levels for pump control, in: Journal of Hydroinformatics. IWA Publishing, pp. 811–822. https://doi.org/10.2166/hydro.2017.141

Aneke, M., Wang, M., 2016. Energy storage technologies and real life applications – A state of the art review. Appl. Energy 179, 350–377. https://doi.org/10.1016/j.apenergy.2016.06.097

Arai, Y., Koizumi, A., Inakazu, T., Masuko, A., Tamura, S., 2013. Optimized operation of water distribution system using multipurpose fuzzy LP model. Water Sci. Technol. Water Supply 13, 66–73. https://doi.org/10.2166/ws.2012.080

Auger, A., Hansen, N., 2005. A restart CMA evolution strategy with increasing population size. 2005 IEEE Congr. Evol. Comput. IEEE CEC 2005. Proc. 2, 1769–1776. https://doi.org/10.1109/CEC.2005.1554902

Aymerich, I., Rieger, L., Sobhani, R., Rosso, D., Corominas, L., 2015. The difference between energy consumption and energy cost: Modelling energy tariff structures for water resource recovery facilities. Water Res. 81, 113–123. https://doi.org/10.1016/j.watres.2015.04.033

Bäck, T., Fogel, D.B., Michalewicz, Z., 2018. Evolutionary Computation 1: Basic Algorithms and Operators. Bristol Philadelphia Inst. Phys. Publ.

Badakhshan, S., Hajibandeh, N., Shafie-khah, M., Catalão, J.P.S., 2019. Impact of solar energy on the integrated operation of electricity-gas grids. Energy 183, 844–853. https://doi.org/10.1016/j.energy.2019.06.107

Bagajewicz, M., Cabrera, E., 2003. Pareto optimal solutions visualization techniques for multiobjective design and upgrade of instrumentation networks. Ind. Eng. Chem. Res. 42, 5195–5203. https://doi.org/10.1021/IE020865G/ASSET/IMAGES/LARGE/IE020865GF00018.JPEG

Bagirov, A.M., Barton, A.F., Mala-Jetmarova, H., Al Nuaimat, A., Ahmed, S.T., Sultanova, N., Yearwood, J., 2013. An algorithm for minimization of pumping costs in water distribution systems using a novel approach to pump scheduling. Math. Comput. Model. 57, 873–886. https://doi.org/10.1016/j.mcm.2012.09.015

Benayoun, R., de Montgolfier, J., Tergny, J., Laritchev, O., 1971. Linear programming with multiple objective functions: Step method (stem). Math. Program. 1971 11 1, 366–375. https://doi.org/10.1007/BF01584098

Blinco, L.J., Simpson, A.R., Lambert, M.F., Marchi, A., 2016. Comparison of Pumping Regimes for Water Distribution Systems to Minimize Cost and Greenhouse Gases. J. Water Resour. Plan. Manag. 142, 04016010. https://doi.org/10.1061/(ASCE)WR.1943-5452.0000633

Blum, C., Roli, A., 2003. Metaheuristics in Combinatorial Optimization: Overview and Conceptual Comparison. ACM Comput. Surv. 35, 268–308. https://doi.org/10.1145/937503.937505

Breiman, L., Friedman, J.H., Olshen, R.A., Stone, C.J., 2017. Classification and regression trees. Classif. Regres. Trees 1–358. https://doi.org/10.1201/9781315139470/CLASSIFICATION-REGRESSION-TREES-LEO-BREIMAN-JEROME-FRIEDMAN-RICHARD-OLSHEN-CHARLES-STONE

Burton, F.L., Stern, F., 1993. Water and wastewater industries: Characteristics and DSM opportunities. California, United States.

CAISO, 2020. Open Access Open Access Same-time Information System - Interval Locational

Marginal Pricing from January 2020 to December 2020 [WWW Document]. URL oasis.caiso.com (accessed 1.10.21).

CAISO, 2018a. Open Access Same-time Information System - Interval Locational Marginal Pricing from May 2018 to October 2018 [WWW Document]. URL oasis.caiso.com (accessed 1.8.19).

CAISO, 2018b. Demand Response User Guide, Version 4.5. California, USA.

CAISO, 2009. Proxy Demand Resource (PDR) Project Implementation Plan (DRAFT). California, USA.

Cembrano, G., Wells, G., Quevedo, J., Pérez, R., Argelaguet, R., 2000. Optimal control of a water distribution network in a supervisory control system. Control Eng. Pract. 8, 1177–1188. https://doi.org/10.1016/S0967-0661(00)00058-7

Chen, Y., Xu, P., Chu, Y., Li, W., Wu, Y., Ni, L., Bao, Y., Wang, K., 2017. Short-term electrical load forecasting using the Support Vector Regression (SVR) model to calculate the demand response baseline for office buildings. Appl. Energy 195, 659–670. https://doi.org/10.1016/j.apenergy.2017.03.034

Cimorelli, L., D'Aniello, A., Cozzolino, L., 2020. Boosting Genetic Algorithm Performance in Pump Scheduling Problems with a Novel Decision-Variable Representation. J. Water Resour. Plan. Manag. 146, 04020023. https://doi.org/10.1061/(ASCE)WR.1943-5452.0001198

Clerc, M., Kennedy, J., 2002. The particle swarm-explosion, stability, and convergence in a multidimensional complex space. IEEE Trans. Evol. Comput. 6, 58–73. https://doi.org/10.1109/4235.985692

Collins, J.J., O'Neill, M., 1999. Visualization of evolutionary algorithms using principal component analysis, in: GECCO Workshop Program. pp. 99–100.

D'Ambrosio, C., Lodi, A., Wiese, S., Bragalli, C., 2015. Mathematical programming techniques in

water network optimization. Eur. J. Oper. Res. https://doi.org/10.1016/j.ejor.2014.12.039

David, M., Sigrid, K., Frayret, J.M., Gilles, P., Nicolas, G., 2015. A Review and Taxonomy of Interactive Optimization Methods in Operations Research. ACM Trans. Interact. Intell. Syst. 5, 17. https://doi.org/10.1145/2808234

De León, S.K., 2018. SB 100: California Renewables Portfolio Standard Program: emissions of greenhouse gases, California Senate.

Duan, Q., Shao, C., Li, X., Shi, Y., 2017. Visualizing the search dynamics in a high-dimensional space for a particle swarm optimizer. Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics) 10593 LNCS, 994–1002. https://doi.org/10.1007/978-3-319-68759-9_82/FIGURES/4

Eggimann, S., Mutzner, L., Wani, O., Schneider, M.Y., Spuhler, D., Moy De Vitry, M., Beutler, P., Maurer, M., 2017. The Potential of Knowing More: A Review of Data-Driven Urban Water Management. Environ. Sci. Technol. https://doi.org/10.1021/acs.est.6b04267

European Commission, 2014. Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions. A policy framework for climate and energy in the period from 2020 to 2030, COM 15 final. Brussels, Belgium.

European Commission, 2011. Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions. Energy Roadmap 2050, COM 885 final. Brussels, Belgium.

Fallside, F., Perry, P.F., 1975. HIERARCHICAL OPTIMISATION OF A WATER-SUPPLY NETWORK. Proc. Inst. Electr. Eng. 122, 202–208. https://doi.org/10.1049/piee.1975.0048

Fortin, F.-A., Marc-André Gardner, U., Parizeau, M., Gagné, C., 2012. DEAP: Evolutionary Algorithms

Made Easy. J. Mach. Learn. Res. 13, 2171–2175. https://doi.org/10.5555/2503308.2503311

Fu, G., Kapelan, Z., Kasprzyk, J.R., Reed, P., Asce, M., 2012. Optimal Design of Water Distribution Systems Using Many-Objective Visual Analytics. J. Water Resour. Plan. Manag. 139, 624–633. https://doi.org/10.1061/(ASCE)WR.1943-5452.0000311

Garrido-Baserba, M., Corominas, L., Cortés, U., Rosso, D., Poch, M., 2020. The Fourth-Revolution in the Water Sector Encounters the Digital Revolution. Environ. Sci. Technol. https://doi.org/10.1021/acs.est.9b04251

GEI Consultants/Navigant Consulting Inc., 2010. Embedded Energy in Water Studies, Study 1: Statewide and Regional Water-Energy Relationship - Draft Final Report. California Public Utilities Commission - Energy Division, California, USA.

Gernaey, K. V., Van Loosdrecht, M.C.M., Henze, M., Lind, M., Jørgensen, S.B., 2004. Activated sludge wastewater treatment plant modelling and simulation: State of the art. Environ. Model. Softw. 19, 763–783. https://doi.org/10.1016/j.envsoft.2003.03.005

Gibbs, M.S., Maier, H.R., Dandy, G.C., 2004. APPLYING FITNESS LANDSCAPE MEASURES TO WATER DISTRIBUTION OPTIMIZATION PROBLEMS. Hydroinformatics 795–802. https://doi.org/10.1142/9789812702838_0098

Giberti, M., Dereli, R.K., Flynn, D., Casey, E., 2019. Predicting wastewater treatment plant performance during aeration demand shifting with a dual-layer reaction settling model. Water Sci. Technol. 81, 1365–1374. https://doi.org/10.2166/wst.2019.262

Giustolisi, O., Laucelli, D., Berardi, L., 2013. Operational Optimization: Water Losses versus Energy Costs. J. Hydraul. Eng. 139, 410–423. https://doi.org/10.1061/(ASCE)HY.1943-7900.0000681

Goldman, F.E., Mays, L.W., 1999. The Application of Simulated Annealing to the Optimal Operation of Water Systems, in: WRPMD'99. American Society of Civil Engineers, Reston, VA, pp. 1–16.

https://doi.org/10.1061/40430(1999)56

Greenman, J., Gálvez, A., Giusti, L., Ieropoulos, I., 2009. Electricity from landfill leachate using microbial fuel cells: Comparison with a biological aerated filter. Enzyme Microb. Technol. 44, 112–119. https://doi.org/10.1016/j.enzmictec.2008.09.012

Gu, Y., Li, Y., Li, X., Luo, P., Wang, H., Robinson, Z.P., Wang, X., Wu, J., Li, F., 2017. The feasibility and challenges of energy self-sufficient wastewater treatment plants. Appl. Energy 204, 1463–1475. https://doi.org/10.1016/j.apenergy.2017.02.069

Hamiche, A.M., Stambouli, A.B., Flazi, S., 2016. A review of the water-energy nexus. Renew. Sustain. Energy Rev. https://doi.org/10.1016/j.rser.2016.07.020

Hansen, N., Ostermeier, A., 2001. Completely derandomized self-adaptation in evolution strategies. Evol. Comput. 9, 159–195. https://doi.org/10.1162/106365601750190398

Hoffman, P., Grinstein, G., Marx, K., Grosse, I., Stanley, E., 1997. DNA visual and analytic data mining. Proc. IEEE Vis. Conf. 437–441. https://doi.org/10.1109/VISUAL.1997.663916

Hordijk, W., Nm, S.F., 1996. A measure of landscapes. Evol. Comput. 4, 335–360. https://doi.org/10.1162/EVCO.1996.4.4.335

Housh, M., Salomons, E., 2018. Optimal Dynamic Pump Triggers for Cost Saving and Robust Water Distribution System Operations. J. Water Resour. Plan. Manag. 145, 04018095. https://doi.org/10.1061/(ASCE)WR.1943-5452.0001028

IEA, 2020. Renewables 2020: Analysis and forecasts to 2025, International Energy Agency.

International Energy Agency, 2019. World Energy Outlook 2019 – Analysis - IEA [WWW Document]. World Energy Outlook 2019.

Jones, T., Forrest, S., 1995. Fitness Distance Correlation as a Measure of Problem Difficulty for

Genetic Algorithms, in: Proceedings of the 6th International Conference on Genetic Algorithms. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, pp. 184–192.

Jowitt, P.W., Germanopoulos, G., 1992. Optimal Pump Scheduling in Water-Supply Networks. J. Water Resour. Plan. Manag. 118, 406–422. https://doi.org/10.1061/(ASCE)0733-9496(1992)118:4(406)

Keedwell, E., Johns, M., Savic, D., 2015. Spatial and Temporal Visualisation of Evolutionary Algorithm Decisions in Water Distribution Network Optimisation, in: Proceedings of the Companion Publication of the 2015 Annual Conference on Genetic and Evolutionary Computation. ACM, New York, NY, USA, pp. 941–948. https://doi.org/10.1145/2739482

Kennedy, J., Eberhart, R., 1995. Particle swarm optimization, in: Proceedings of ICNN'95 - International Conference on Neural Networks. IEEE, pp. 1942–1948. https://doi.org/10.1109/ICNN.1995.488968

Kenway, S.J., Lam, K.L., Stokes-Draut, J., Sanders, K.T., Binks, A.N., Bors, J., Head, B., Olsson, G., McMahon, J.E., 2019. Defining water-related energy for global comparison, clearer communication, and sharper policy. J. Clean. Prod. 236, 117502. https://doi.org/10.1016/j.jclepro.2019.06.333

Kernan, R., Liu, X., McLoone, S., Fox, B., 2017. Demand side management of an urban water supply using wholesale electricity price. Appl. Energy 189, 395–402. https://doi.org/10.1016/j.apenergy.2016.12.082

Kim, Y.H., Moon, B.R., 2003. New usage of Sammon's mapping for genetic visualization. Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics) 2723, 1136–1147. https://doi.org/10.1007/3-540-45105-6_122/COVER/

Kirchem, D., Lynch, M., Bertsch, V., Casey, E., 2018. Market Effects of Industrial Demand Response

and Flexibility Potential from Wastewater Treatment Facilities, in: 2018 15th International Conference on the European Energy Market (EEM). IEEE, pp. 1–6. https://doi.org/10.1109/EEM.2018.8469974

Klau, G.W., Lesh, N., Marks, J., Mitzenmacher, M., Schafer, G.T., 2002. The HuGS platform: A toolkit for interactive optimization. Proc. Work. Adv. Vis. Interfaces AVI 324–330. https://doi.org/10.1145/1556262.1556314

Klise, Katherine A., Bynum, M., Moriarty, D., Murray, R., 2017. A software framework for assessing the resilience of drinking water systems to disasters with an example earthquake case study. Environ. Model. Softw. 95, 420–431. https://doi.org/10.1016/j.envsoft.2017.06.022

Klise, Katherine A, Hart, D., Moriarty, D.M., Bynum, M.L., Murray, R., Burkhardt, J., Haxton, T., 2017. Water network tool for resilience (WNTR) user manual.

Kurek, W., Ostfeld, A., 2013. Multi-objective optimization of water quality, pumps operation, and storage sizing of water distribution systems. J. Environ. Manage. 115, 189–197. https://doi.org/10.1016/j.jenvman.2012.11.030

Kurek, W., Ostfeld, A., 2012. Multiobjective Water Distribution Systems Control of Pumping Cost, Water Quality, and Storage-Reliability Constraints. J. Water Resour. Plan. Manag. 140, 184–193. https://doi.org/10.1061/(ASCE)WR.1943-5452.0000309

LaGrega, M.D., Keenan, J.D., 1974. Effects of Equalizing Wastewater Flows. J. (Water Pollut. Control Fed. 46, 123–132.

Lee, J.D., See, K.A., 2004. Trust in automation: Designing for appropriate reliance. Hum. Factors 46, 50–80. https://doi.org/10.1518/hfes.46.1.50_30392

Lekov, A., Thompson, L., Mckane, A., Song, K., Piette, M.A., 2009. Opportunities for Energy Efficiency and Open Automated Demand Response in Wastewater Treatment Facilities in California –

Phase I Report. California, USA.

Li, H. ran, He, F. zhi, Yan, X. hu, 2019. IBEA-SVM: An Indicator-based Evolutionary Algorithm Based on Pre-selection with Classification Guided by SVM. Appl. Math. J. Chinese Univ. 2019 341 34, 1–26. https://doi.org/10.1007/S11766-019-3706-1

Li, W., Meng, X., Huang, Y., 2021. Fitness distance correlation and mixed search strategy for differential evolution. Neurocomputing 458, 514–525. https://doi.org/10.1016/J.NEUCOM.2019.12.141

Li, W., Sun, B., Huang, Y., Mahmoodi, S., 2022. Adaptive complex network topology with fitness distance correlation framework for particle swarm optimization. Int. J. Intell. Syst. 37, 5217–5247. https://doi.org/10.1002/INT.22790

Li, Y., Musabandesu, E., Fujiwara, T., Loge, F.J., Ma, K.L., 2021. A Visual Analytics System for Water Distribution System Optimization. Proc. - 2021 IEEE Vis. Conf. - Short Pap. VIS 2021 126–130. https://doi.org/10.1109/VIS49827.2021.9623272

Liang, X., 2017. Emerging Power Quality Challenges Due to Integration of Renewable Energy Sources. IEEE Trans. Ind. Appl. 53, 855–866. https://doi.org/10.1109/TIA.2016.2626253

Limaye, D.R., Jaywant, S., 2019. Development of Indicators from the World Bank IBNET Database.

Liu, Y., Barrows, C., Macknick, J., Mauter, M., 2020. Optimization Framework to Assess the Demand Response Capacity of a Water Distribution System. J. Water Resour. Plan. Manag. 146, 04020063. https://doi.org/10.1061/(ASCE)WR.1943-5452.0001258

López-Ibáñez, M., Prasad, T.D., Paechter, B., 2008. Ant Colony Optimization for Optimal Control of Pumps in Water Distribution Networks. J. Water Resour. Plan. Manag. 134, 337–346. https://doi.org/10.1061/(ASCE)0733-9496(2008)134:4(337)

Lund, P.D., Lindgren, J., Mikkola, J., Salpakari, J., 2015. Review of energy system flexibility measures to enable high levels of variable renewable electricity. Renew. Sustain. Energy Rev. 45, 785–807. https://doi.org/10.1016/j.rser.2015.01.057

Mackle, G., Savic, D.A., Walters, G.A., 1995. Application of genetic algorithms to pump scheduling for water supply, in: First International Conference on Genetic Algorithms in Engineering Systems: Innovations and Applications. Institution of Engineering and Technology, Sheffield, UK, pp. 400–405. https://doi.org/10.1049/cp:19951082

Maier, H.R., Kapelan, Z., Kasprzyk, J., Kollat, J., Matott, L.S., Cunha, M.C., Dandy, G.C., Gibbs, M.S., Keedwell, E., Marchi, A., Ostfeld, A., Savic, D., Solomatine, D.P., Vrugt, J.A., Zecchin, A.C., Minsker, B.S., Barbour, E.J., Kuczera, G., Pasha, F., Castelletti, A., Giuliani, M., Reed, P.M., 2014. Evolutionary algorithms and other metaheuristics in water resources: Current status, research challenges and future directions. Environ. Model. Softw. 62, 271–299. https://doi.org/10.1016/j.envsoft.2014.09.013

Maktabifard, M., Zaborowska, E., Makinia, J., 2018. Achieving energy neutrality in wastewater treatment plants through energy savings and enhancing renewable energy production. Rev. Environ. Sci. Biotechnol. 17, 655–689. https://doi.org/10.1007/s11157-018-9478-x

Mala-Jetmarova, H., Sultanova, N., Savic, D., 2018. Lost in Optimisation of Water Distribution Systems? A Literature Review of System Design. Water 2018, Vol. 10, Page 307 10, 307. https://doi.org/10.3390/W10030307

Mala-Jetmarova, H., Sultanova, N., Savic, D., 2017. Lost in optimisation of water distribution systems? Environ. Model. Softw. 93, 209–254. https://doi.org/10.1016/j.envsoft.2017.02.009

Malan, K.M., Oberholzer, J.F., Engelbrecht, A.P., 2015. Characterising constrained continuous optimisation problems. 2015 IEEE Congr. Evol. Comput. CEC 2015 - Proc. 1351–1358.

https://doi.org/10.1109/CEC.2015.7257045

Mannina, G., Rebouças, T.F., Cosenza, A., Chandran, K., 2019. A plant-wide wastewater treatment plant model for carbon and energy footprint: Model application and scenario analysis. J. Clean. Prod. 217, 244–256. https://doi.org/10.1016/j.jclepro.2019.01.255

Marchi, A., Simpson, A.R., Lambert, M.F., 2016. Optimization of Pump Operation Using Rule-Based Controls in EPANET2: New ETTAR Toolkit and Correction of Energy Computation. J. Water Resour. Plan. Manag. 142, 04016012. https://doi.org/10.1061/(asce)wr.1943-5452.0000637

Mccandlish, D.M., 2011. VISUALIZING FITNESS LANDSCAPES. Evolution (N. Y). 65, 1544–1558. https://doi.org/10.1111/J.1558-5646.2011.01236.X

Meignan, D., 2015. An experimental investigation of reoptimization for shift scheduling, in: Proceedings of the 11th Metaheuristics International Conference. pp. 1–10.

Menke, R., Abraham, E., Parpas, P., Stoianov, I., 2016. Demonstrating demand response from water distribution system through pump scheduling. Appl. Energy 170, 377–387. https://doi.org/10.1016/j.apenergy.2016.02.136

Menke, R., Kadehjian, K., Abraham, E., Stoianov, I., 2017. Investigating trade-offs between the operating cost and green house gas emissions from water distribution systems. Sustain. Energy Technol. Assessments 21, 13–22. https://doi.org/10.1016/J.SETA.2017.03.002

Miettinen, K., Eskelinen, P., Ruiz, F., Luque, M., 2010. NAUTILUS method: An interactive technique in multiobjective optimization based on the nadir point. Eur. J. Oper. Res. 206, 426–434. https://doi.org/10.1016/J.EJOR.2010.02.041

Mohajeryami, S., Doostan, M., Asadinejad, A., Schwarz, P., 2017. Error analysis of customer baseline load (CBL) calculation methods for residential customers. IEEE Trans. Ind. Appl. 53, 5–14. https://doi.org/10.1109/TIA.2016.2613985

Musabandesu, E., Good, R.T., Herman, J.D., Loge, F.J., 2022. Optimizing secondary time-based control structures to support renewable energy integration for water distribution systems. J. Water Resour. Plan. Manag. (In Rev.

North Coast Regional Water Quality Control Board, 2013. Regional Water Board Order No. R1-2013-0001 NPDES Permit No. CA0022764 WDID No. lB830990SON Waste Discharge Requirements and Master Reclamation Permit for the City of Santa Rosa Subregional Water Reclamation System Sonoma County [WWW Document]. URL www.waterboards.ca.gov/northcoast/board_decisions/adopted_orders/pdf/2013/131121_0001_SantaRosaNPDES.pdf (accessed 10.25.19).

Odan, F.K., Ribeiro Reis, L.F., Kapelan, Z., 2015. Real-Time Multiobjective Optimization of Operation of Water Supply Systems. J. Water Resour. Plan. Manag. 141, 04015011. https://doi.org/10.1061/(ASCE)WR.1943-5452.0000515

Oslen, D., Goli, S., Faulkner, D., Mckane, A., 2012. Opportunities for automated demand response in wastewater treatment facilities in California-Southeast Water Pollution Control Plant Case Study, Lawrence Berkeley National Laboratory. California, USA.

Ostfeld, A., Kogan, D., Shamir, U., 2002. Reliability simulation of water distribution systems – single and multiquality. Urban Water 4, 53–61. https://doi.org/10.1016/S1462-0758(01)00055-3

Palensky, P., Dietrich, D., 2011. Demand Side Management: Demand Response, Intelligent Energy Systems, and Smart Loads. IEEE Trans. Ind. Informatics 7, 381–388. https://doi.org/10.1109/TII.2011.2158841

Pan, L., He, C., Tian, Y., Wang, H., Zhang, X., Jin, Y., 2019. A Classification-Based Surrogate-Assisted Evolutionary Algorithm for Expensive Many-Objective Optimization. IEEE Trans. Evol. Comput. 23, 74–88. https://doi.org/10.1109/TEVC.2018.2802784

Panepinto, D., Fiore, S., Zappone, M., Genon, G., Meucci, L., 2016. Evaluation of the energy efficiency of a large wastewater treatment plant in Italy. Appl. Energy 161, 404–411. https://doi.org/10.1016/j.apenergy.2015.10.027

Park, S., Ryu, S., Choi, Y., Kim, J., Kim, H., 2015. Data-Driven Baseline Estimation of Residential Buildings for Demand Response. Energies 8, 10239–10259. https://doi.org/10.3390/en80910239

Paschke, M., Spencer, K., Waniarcha, N., Simpson, A.R., Widdop, T., 2001. Genetic algorithms for optimising pumping operations. 19th Fed. Conv. Aust. Water Assoc. Canberra, Aust.

Paterakis, N.G., Erdinç, O., Catalão, J.P.S., 2017. An overview of Demand Response: Key-elements and international experience. Renew. Sustain. Energy Rev. 69, 871–891. https://doi.org/10.1016/j.rser.2016.11.167

Pitzer, E., Affenzeller, M., 2012. A comprehensive survey on fitness landscape analysis. Stud. Comput. Intell. 378, 161–191. https://doi.org/10.1007/978-3-642-23229-9_8/COVER/

Plappally, A.K., Lienhard, V.J.H., 2012. Energy requirements for water production, treatment, end use, reclamation, and disposal. Renew. Sustain. Energy Rev. 16, 4818–4848. https://doi.org/10.1016/j.rser.2012.05.022

Poli, R., Kennedy, J., Blackwell, T., 2007. Particle swarm optimization. Swarm Intell. 2007 11 1, 33–57. https://doi.org/10.1007/S11721-007-0002-0

Pryke, A., Mostaghim, S., Nazemi, A., 2007. Heatmap Visualization of Population Based Multi Objective Algorithms. Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics) 4403 LNCS, 361–375. https://doi.org/10.1007/978-3-540-70928-2_29

Quinn, J.D., Reed, P.M., Giuliani, M., Castelletti, A., 2017. Rival framings: A framework for discovering how problem formulation uncertainties shape risk management trade-offs in water resources

systems. Water Resour. Res. 53, 7208–7233. https://doi.org/10.1002/2017WR020524

Quintiliani, C., Creaco, E., 2019. Using Additional Time Slots for Improving Pump Control Optimization Based on Trigger Levels. Water Resour. Manag. 33, 3175–3186. https://doi.org/10.1007/s11269-019-02297-6

Rao, Z., Salomons, E., 2007. Development of a real-time, near-optimal control process for water-distribution networks. J. Hydroinformatics. https://doi.org/10.2166/hydro.2006.015

Rossman, L., Woo, H., Tryby, M., Shang, F., Janke, R., Haxton., T., 2020. EPANET 2.2 User Manual. U.S. Environ. Prot. Agency Cincinnati.

Santhosh, A., Farid, A.M., Youcef-Toumi, K., 2014. The impact of storage facility capacity and ramping capabilities on the supply side economic dispatch of the energy-water nexus. Energy 66, 363–377. https://doi.org/10.1016/j.energy.2014.01.031

Savic, D.A., Walters, G.A., Schwab, M., 1997. Multiobjective genetic algorithms for pump scheduling in water supply. Evol. Comput. 1305, 227–235. https://doi.org/10.1007/BFB0027177

Schäfer, M., 2019. Short-term flexibility for energy grids provided by wastewater treatment plants with anaerobic sludge digestion. Water Sci. Technol. 81, 1388–1397. https://doi.org/10.2166/wst.2019.365

Schäfer, M., Gretzschel, O., Schmitt, T.G., Knerr, H., 2015. Wastewater Treatment Plants as System Service Provider for Renewable Energy Storage and Control Energy in Virtual Power Plants – A Potential Analysis. Energy Procedia 73, 87–93. https://doi.org/10.1016/j.egypro.2015.07.566

Schäfer, M., Hobus, I., Schmitt, T.G., 2017. Energetic flexibility on wastewater treatment plants. Water Sci. Technol. 76, 1225–1233. https://doi.org/10.2166/wst.2017.308

Seier, M., Schebek, L., 2017. Model-based investigation of residual load smoothing through dynamic electricity purchase: The case of wastewater treatment plants in Germany. Appl. Energy 205, 210–224. https://doi.org/10.1016/j.apenergy.2017.07.116

Siano, P., 2014. Demand response and smart grids—A survey. Renew. Sustain. Energy Rev. 30, 461–478. https://doi.org/10.1016/j.rser.2013.10.022

Sousa, J., Muranho, J., Sá Marques, A., Gomes, R., 2016. Optimal Management of Water Distribution Networks with Simulated Annealing: The C-Town Problem. J. Water Resour. Plan. Manag. 142, C4015010. https://doi.org/10.1061/(asce)wr.1943-5452.0000604

State Water Resource Control Board, 2000. Water Quality Order No. 2000-03 NPDES Permit No. CA0022764 ID NO. lB830990SON Waste Discharge Requirements for the City of Santa Rosa, Laguna Subregional Wastewater Collection, Treatment, Conveyance, Reuse, and Disposal Facilities [WWW Document].

Sterling, M., Coulbeck, B., 1975. Technical Note. A dynamic programming solution to optimization of pumping costs. Proc. Inst. Civ. Eng. 59, 813–818. https://doi.org/10.1680/iicep.1975.3642

Stokes, C.S., Maier, H.R., Simpson, A.R., 2015. Water Distribution System Pumping Operational Greenhouse Gas Emissions Minimization by Considering Time-Dependent Emissions Factors. J. Water Resour. Plan. Manag. 141, 04014088. https://doi.org/10.1061/(ASCE)WR.1943-5452.0000484

Thompson, L., Lekov, A., Mckane, A., Piette, M.A., 2010. Opportunities for Open Automated Demand Response in Wastewater Treatment Facilities in California – Phase II Report: San Luis Rey Wastewater Treatment Plant Case Study. California, USA.

Thompson, L., Song, K., Lekov, A., McKane, A., 2008. Automated demand response opportunities in wastewater treatment facilities. California, USA.

Valdés, J.J., Barton, A.J., 2007. Visualizing high dimensional objective spaces for multi-objective optimization: A virtual reality approach. 2007 IEEE Congr. Evol. Comput. CEC 2007 4199–4206. https://doi.org/10.1109/CEC.2007.4425019

van der Maaten, L., Hinton, G., 2008. Visualizing Data using t-SNE. J. Mach. Learn. Res. 9, 2579–2605.

Van Staden, A.J., Zhang, J., Xia, X., 2011. A model predictive control strategy for load shifting in a water pumping scheme with maximum demand charges. Appl. Energy 88, 4785–4794. https://doi.org/10.1016/j.apenergy.2011.06.054

Van Zyl, J.E., Clayton, C.R.I., 2007. The effect of pressure on leakage in water distribution systems. Proc. Inst. Civ. Eng. Water Manag. 160, 109–114. https://doi.org/10.1680/WAMA.2007.160.2.109/ASSET/IMAGES/SMALL/WAMA160-109-F3.GIF

Van Zyl, J.E., Savic, D.A., Walters, G.A., 2004. Operational Optimization of Water Distribution Systems Using a Hybrid Genetic Algorithm. J. Water Resour. Plan. Manag. 130, 160–170. https://doi.org/10.1061/(ASCE)0733-9496(2004)130:2(160)

Verzijlbergh, R.A., De Vries, L.J., Dijkema, G.P.J., Herder, P.M., 2017. Institutional challenges caused by the integration of renewable energy sources in the European electricity sector. Renew. Sustain. Energy Rev. 75, 660–667. https://doi.org/10.1016/j.rser.2016.11.039

Wakeel, M., Chen, B., Hayat, T., Alsaedi, A., Ahmad, B., 2016. Energy consumption for water use cycles in different countries: A review. Appl. Energy 178, 868–885. https://doi.org/10.1016/j.apenergy.2016.06.114

Wallenius, J., 1975. Comparative Evaluation of Some Interactive Approaches to Multicriterion Optimization. http://dx.doi.org/10.1287/mnsc.21.12.1387 21, 1387–1396. https://doi.org/10.1287/MNSC.21.12.1387

Wang, H., Yang, Y., Keller, A.A., Li, X., Feng, S., Dong, Y. nan, Li, F., 2016. Comparative analysis of energy intensity and carbon emissions in wastewater treatment in USA, Germany, China and South Africa. Appl. Energy 184, 873–881. https://doi.org/10.1016/j.apenergy.2016.07.061

Wegley, C., Eusuff, M., Lansey, K., 2004. Determining pump operations using particle swarm optimization, in: Joint Conference on Water Resource Engineering and Water Resources Planning and Management 2000: Building Partnerships. American Society of Civil Engineers, Reston, VA, pp. 1–6. https://doi.org/10.1061/40517(2000)206

Wi, Y.M., Kim, J.H., Joo, S.K., Park, J.B., Oh, J.C., 2009. Customer baseline load (CBL) Calculation using exponential smoothing model with weather adjustment, in: Transmission and Distribution Conference and Exposition: Asia and Pacific, T and D Asia 2009. https://doi.org/10.1109/TD-ASIA.2009.5356984

Williams, A.A., 1996. Pumps as turbines for low cost micro hydro power. Renew. Energy 9, 1227–1234. https://doi.org/10.1016/0960-1481(96)88498-9

Xin, B., Chen, L., Chen, J., Ishibuchi, H., Hirota, K., Liu, B., 2018. Interactive multiobjective optimization: A review of the state-of-the-art. IEEE Access 6, 41256–41279. https://doi.org/10.1109/ACCESS.2018.2856832

Yu, G., Powell, R.S., Sterling, M.J.H., 1994. Optimized pump scheduling in water distribution systems. J. Optim. Theory Appl. 83, 463–488. https://doi.org/10.1007/BF02207638

Zairi, M., Aydi, A., Dhia, H. Ben, 2014. Leachate generation and biogas energy recovery in the Jebel Chakir municipal solid waste landfill, Tunisia. J. Mater. Cycles Waste Manag. 16, 141–150. https://doi.org/10.1007/s10163-013-0164-3

Zhang, Q., Li, J., 2012. Demand response in electricity markets: A review, in: 9th International Conference on the European Energy Market, EEM 12.

https://doi.org/10.1109/EEM.2012.6254817

Zhang, S., Li, S., Harley, R.G., Habetler, T.G., 2018. Visualization and Data Mining of Multi-Objective Electric Machine Optimizations with Self-Organizing Maps: A Case Study on Switched Reluctance Machines. 2018 IEEE Energy Convers. Congr. Expo. ECCE 2018 4296–4302. https://doi.org/10.1109/ECCE.2018.8558399

Zhou, K., Fu, C., Yang, S., 2016. Big data driven smart energy management: From big data to big insights. Renew. Sustain. Energy Rev. https://doi.org/10.1016/j.rser.2015.11.050