A Dataset for Research on Water Sustainability

Pranjol Sen Gupta
The University of Texas at Arlington
pranjolsen.gupta@mavs.uta.edu

Md Rajib Hossen
The University of Texas at Arlington
mdrajib.hossen@mavs.uta.edu

Pengfei Li
University of California, Riverside
pli081@ucr.edu

Shaolei Ren
University of California, Riverside
sren@ece.ucr.edu

Mohammad A. Islam
The University of Texas at Arlington
mislam@uta.edu

ABSTRACT

Freshwater scarcity is a global problem that requires collective efforts across all industry sectors. Nevertheless, a lack of access to operational water footprint data bars many applications from exploring optimization opportunities hidden within the temporal and spatial variations. To break this barrier into research in water sustainability, we build a dataset for operation direct water usage in the cooling systems and indirect water embedded in electricity generation. Our dataset consists of the hourly water efficiency of major U.S. cities and states from 2019 to 2023. We also offer cooling system models that capture the impact of weather on water efficiency. We present a preliminary analysis of our dataset and discuss three potential applications that can benefit from it. Our dataset is publicly available at Open Science Framework (OSF) [1].

CCS CONCEPTS

• **General and reference** → **Metrics**; **Measurement**; • **Hardware** → **Impact on the environment**.

KEYWORDS

Sustainability, Water Consumption, Dataset, Cooling Tower, Electricity Generation

ACM Reference Format:
Pranjol Sen Gupta, Md Rajib Hossen, Pengfei Li, Shaolei Ren, and Mohammad A. Islam. 2024. A Dataset for Research on Water Sustainability. In *The 15th ACM International Conference on Future and Sustainable Energy Systems (E-Energy '24), June 04–07, 2024, Singapore, Singapore.* ACM, New York, NY, USA, 5 pages. https://doi.org/10.1145/3632775.3661962

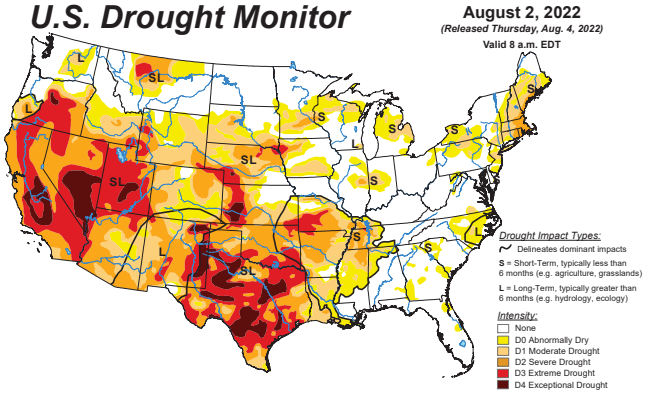


Figure 1: US drought map for August 2, 2022, with 4.47% area under exceptional drought (D4), 18.96% area under extreme drought or worse (D3-D4), and 37.03% area under severe drought or worse (D2-D4) [3]

1 INTRODUCTION

Global freshwater supply is under immense pressure due to the growing population and deteriorating climate conditions, making extended droughts a norm in many parts of the world [2]. For example, Fig. 1 shows that 37.03% of the U.S. area was under severe drought or worse in 2022 [3]. Even in regions not historically prone to drought, such as the eastern U.S., the importance of water conservation persists due to aging public water infrastructure [4]. Therefore, every industry and application sector must scrutinize its water footprint and actively contribute to water sustainability efforts [5–7].

In this paper, we focus on two pervasive components that significantly contribute to the water footprint across many sectors — the water consumption in the cooling system and the water footprint embedded in the electricity consumption. Evaporative cooling has been extensively used for the temperature regulation of buildings, and the water usage in these cooling systems can account for more than 50% of the building's total use [8]. Meanwhile, the water footprint in electricity generation remains high despite a steady gain in water efficiency over the years. In 2021, every megawatt-hour of electricity generation used nearly 12,000 gallons of water [9].

The substantial water footprints in cooling systems and electricity consumption also present considerable opportunities for savings. More importantly, both of these water footprints under scrutiny vary over time — the cooling system's water consumption varies with local weather conditions, whereas the water embedded in electricity changes with variations in electricity generation sources. These temporal variations can be leveraged by applications like EV charging, which benefit from scheduling flexibilities [10]. Moreover, large cloud-scale data center applications can also incorporate in their load balancing the variations in water footprint across geographical locations [11, 12]. However, the general lack of access to operational water footprint data hinders the development of water-sustainable operation strategies across many applications, such as building management and EV charging.

E-Energy '24, June 04–07, 2024, Singapore, Singapore

ACM ISBN 979-8-4007-0480-2/24/06
https://doi.org/10.1145/3632775.3661962

Table 1: Estimated water intensity of different electricity generation sources in the U.S. [13].

Energy Source	Coal	Hydro	Natural Gas	Nuclear	Other	Petroleum	Solar	Wind
Water Intensity (L/kWh)	1.817	22.675	0.795	2.309	0.757	1.363	0.000	0.000

To foster water sustainability research and open up the largely untapped optimization opportunities in the temporal and spatial variations of water footprints, we build a water efficiency dataset that provides the hourly water efficiency of the cooling system and electricity generation across major U.S. cities and states from January 2019 to December 2023. We also present our cooling system models to capture the impact of weather conditions on the water consumption rate of the cooling system. Our dataset, along with all source data, models, and scripts, are made publicly available at Open Science Framework (OSF) [1]. We conduct a preliminary analysis of our dataset, offering insights into temporal and spatial variations of water efficiency. Additionally, we discuss three sample applications that can benefit from our dataset.

2 PRELIMINARIES

2.1 Water Footprint

Withdrawal vs consumption. These are two important terms in the context of water management that warrant clarification [14, 15]. Water withdrawal is the total amount of water taken from a water source. It does not imply permanent removal and can be returned to the source after use, although it may undergo changes in quality or temperature during the process. Water consumption refers to the portion of withdrawn water that is not returned to its source. It represents the amount of water that is either evaporated, incorporated into products, or otherwise not available for immediate reuse in the same water source. While water withdrawal is integral for sustainable water use [14], our focus in this work centers on water consumption, which poses a more imminent threat to available water and also consistent with the water footprint literature [15].

Direct water consumption. It refers to the water, an entity consumes for its own operational processes and activities. It involves water that is physically used on-site and is often directly under the control or management of the entity. Examples of direct water consumption are water consumed in the cooling systems, sanitation and cleaning, irrigation, and fire suppression.

Indirect water consumption. It refers to the water consumed beyond an entity's operational boundaries and direct control. Examples of an entity's indirect water consumption are water usage in the production of energy consumed, supply chain, manufacturing and transportation equipment, and consumer use.

To enable holistic sustainability, we can align the direct and indirect water footprint with the Greenhouse Gas Protocol's (GHG Protocol) widely accepted accounting approach for measuring and managing greenhouse gas emissions [16]. While the GHG Protocol has been primarily associated with greenhouse gas emissions, it can be extended to include water usage as well [13]. Following the GHG Protocol, the direct water footprint can be considered Scope 1, while the indirect water footprint falls under Scopes 2 and 3.

2.2 Water Consumption in Cooling Systems

Commercial high-capacity cooling systems, including those used in office buildings and many data centers, typically use water in their cooling system for heat transport and dissipation into the environment [17]. As illustrated in Fig. 2, these cooling systems consist of two water loops — the inner loop carries the heat from the facility air handlers to the chiller heat exchanger, and the outer loop carries the heat from the chiller heat exchanger to the cooling tower for releasing the heat. The inner loop is closed and does not lose any water. The outer loop, on the other hand, sends hot water to the cooling tower, which cools down the water using water evaporation and, therefore, loses water in the process. This water loss through evaporation is the direct water consumption for the cooling system. Note that the water in the outer loop also requires regular recycling (known as blowdown) to avoid any buildup from concentrated minerals due to evaporation. The water lost through blowdown, however, is not considered water consumption as it is returned to the source as grey water.

2.3 Water Consumption in Electricity Generation

Different electricity generation sources have varying water consumption patterns, and their impact on water resources depends on the technology and processes involved [18, 19]. Thermal power plants such as coal, natural gas, and oil use water for cooling purposes in the generation process. Nuclear power plants also require water for cooling, typically through cooling towers or direct discharges. Solar and wind power generation technologies have minimal water consumption during the electricity generation phase. Water use is mostly associated with the manufacturing and maintenance of the equipment. Hydropower plants' water consumption, on the other hand, mainly comes from expedited surface evaporation in their water reservoirs. Table 1 shows the U.S. average water consumption to generate a kilowatt-hour of electricity [13].

2.4 Water Usage Effectiveness (WUE)

Water Usage Effectiveness (WUE) is an operational water metric that quantifies the water efficiency of a system, as defined below

$$\text{WUE} = \frac{\text{Water Consumption}}{\text{Energy Processed/Generated}} \tag{1}$$

For a cooling system, WUE is the ratio of water consumption to the amount of heat dissipated, whereas, for electricity, it is the ratio of water consumption and electric energy generation.

3 WATER EFFICIENCY DATASET

3.1 Methodology

Estimation of direct WUE. A cooling tower's water consumption varies with the outside air temperature and humidity. More specifically, the cooling tower consumes more water (i.e., water

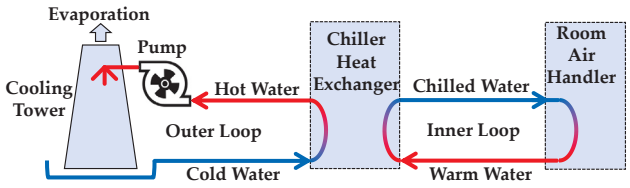
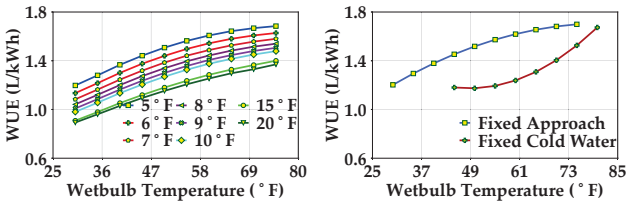


Figure 2: Cooling system with evaporative cooling.



(a) (b)

Figure 3: (a) Change in WUE with wet bulb temperatures with different approaches (i.e., the difference between cold water temperature and wet bulb temperature). (b) Model of weather impact on direct WUE. For the "Fixed Approach" model, the approach is set at 5°F. For the "Fixed Cold Water" temperature model, cold water temperature is set at 85°F.

evaporates away) for the same cooling load when the outside air temperature is higher or more humid. While the precise relationship between WUE and weather conditions for a specific cooling system can vary, we here offer a generic model that captures the impact of the weather conditions on commercial cooling towers. We utilize the Water Calculator tool offered by SPX Technologies [20] to derive our model.

Before introducing our model, we briefly discuss the operational set points that are crucial to understanding cooling towers and the SPX water calculator, as illustrated in Fig. 2. The "range" denotes the temperature difference between hot and cold water in the outer loop. The range and flow rate determine the system's cooling load. SPX water calculator captures the impact of weather conditions by including the wet bulb temperature. Wet bulb temperature indicates the temperature of air saturated with water and can be measured from a wet cloth exposed to airflow. Wet bulb temperature can be measured from a wet cloth exposed to airflow. For cooling tower operation, there must be a temperature difference between the wet-bulb temperature and the cold water temperature. This temperature difference is called the "approach".

Next, to analyze the impact of weather conditions on cooling tower water consumption using the SPX water calculator, first we set the flow rate to 1000 gallons per minute and a range of 10 Fahrenheit. Note that the flow rate set point here is chosen to be an arbitrary round number and does not affect the WUE. This configuration results in a cooling capacity of 1466 kW or 417 tons. Meanwhile, the drift rate and concentration are maintained at their SPX default values of 0.005% and 3, respectively. The drift rate refers to water droplets carried away from the cooling tower by airflow, while concentration indicates how many times water circulates in the outer loop before being discarded. In the SPX water calculator, drift rate and concentration values do not influence the water evaporation rate.

Fig. 3(a) shows the water efficiency of the cooling tower at different wet-bulb temperatures and different approaches. We see that at any given wet-bulb temperature, the WUE goes up with a decreasing approach. While this indicates a higher approach will result in lower water consumption, it also leads to a higher temperature of the cold water in the chiller heat exchanger, requiring the heat exchanger to work harder (and consume more energy) to transfer the heat from the inner loop to the outer loop. Here, we offer two different models, considering two operation strategies. In the first one, we *fix the approach* to 5°F to maximize the heat exchanger efficiency. In the second one, we *fix the cold water temperature* to 85°F (the typical maximum allowed temperature for a heat exchanger) and, therefore have a variable approach with changing wet bulb temperature. Using these two strategies on our data points in Fig. 3(b), we derive the following two models

$$W_{direct}^{FixedApproach} = -0.0001896 \cdot T_w^2 + 0.03095 \cdot T_w + 0.4442 \quad (2)$$

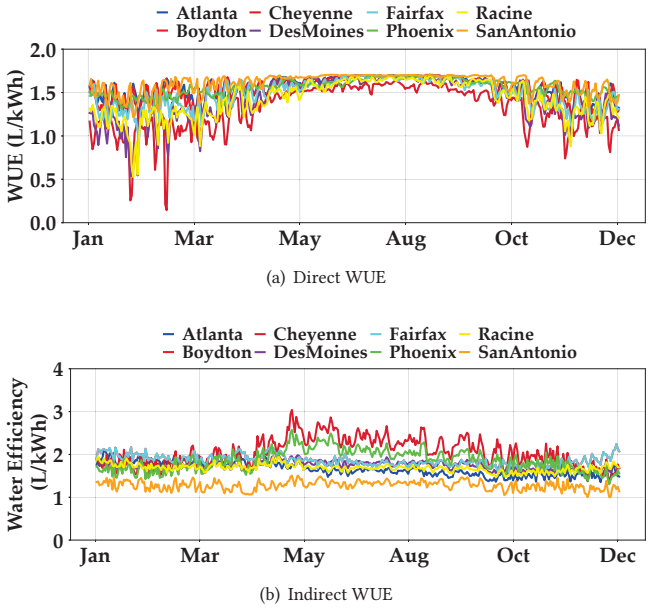
$$W_{direct}^{FixedColdWater} = 0.0005112 \cdot T_w^2 - 0.04982 \cdot T_w + 2.387 \quad (3)$$

where T_w is the wet bulb temperature in Fahrenheit. Note here that, due to incompatible operation set points in cold weather conditions, in Eqns. (2) and (3), the lower limits for T_w are 30°F and 45°F, respectively. We show our models in Fig. 3(b).

In deriving Eqns. (2) and (3) above, we consider 100% heat-transfer efficiency at the chiller heat exchanger and perfect thermal isolation in the water loops. That is, the cooling load is solely determined by the flow rate and temperature difference between hot and cold water. Moreover, we use SPX's own estimation of water evaporation rates in their cooling towers.

In practice, the cooling system efficiency can vary depending on the specific installation and manufacturer. This may increase or decrease the tower's water evaporation rate. To capture such variability, we can introduce a cooling tower efficiency multiplier $\lambda > 0$ and update Eqns. (2) and (3) as $W_{direct}^{FixedApproach} = \lambda \cdot (-0.0001896 \cdot T_w^2 + 0.03095 \cdot T_w + 0.4442)$ and $W_{direct}^{FixedColdWater} = \lambda \cdot (0.0005112 \cdot T_w^2 - 0.04982 \cdot T_w + 2.387)$. The value of λ for a particular cooling system can be estimated by measuring its cooling load, water flow rate, range, and evaporation rate for a few different operational set points. For the rest of this paper, we consider $\lambda = 1$ unless otherwise specified.

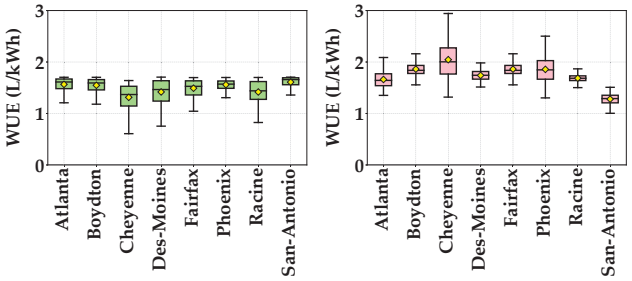
Estimation of indirect water. The electricity consumed from the power grid comes from various generation sources such as coal, oil, hydro, natural gas, nuclear, wind, and solar. These various electricity generation sources have a varying degree of water footprint associated with their electricity generation. However, once the electricity enters the power grid, it is difficult to separate the generation sources. Hence, the water footprint embedded in the electricity from the power grid is a mix of water footprints from multiple sources. Following prior literature and common practice [13], we calculate indirect water footprint using a weighted mix of water footprints from energy sources where each source's weight corresponds to the fraction of electricity generation it is responsible



(a) Direct WUE

(b) Indirect WUE

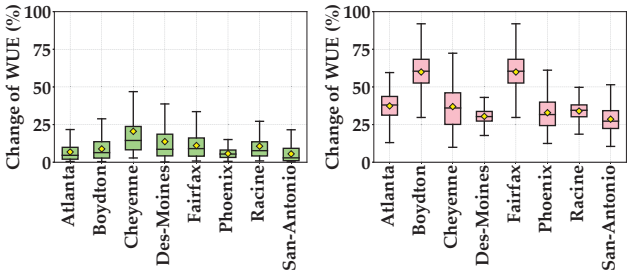
Figure 4: Water efficiency across different U.S. locations in 2023.



(a) Direct water

(b) Indirect water

Figure 5: Average WUE across different locations.



(a) Direct water

(b) Indirect water

Figure 6: Change of WUE in any 24-hour period.

for. We use the following formula

$$W_{indirect}(t) = \frac{\sum_k e_k(t) \cdot w_k}{\sum_k e_k(t)} \tag{4}$$

where $e_k(t)$ is the electricity generation from energy source k at time t and w_k is water footprint of energy source k.

3.2 Scope and Source of the Data

We now provide some details of our dataset.

Temporal resolution and duration. In our dataset, we present five-year data from January 2019 to December 2023. We collect our data at a temporal resolution of one hour.

Locations. We incorporate 58 major US cities with at least one city from each state (except Hawaii).

Weather data. To capture the temporal variation in direct WUE due to weather conditions, we collect the weather data from Weather Underground [21]. It offers location-wise hourly air temperature and relative humidity. We estimate the wet bulb temperature from the air temperature and relative humidity using the Stull formula [22].

Electricity data. We collect our hourly electricity data from EIA OpenData [23], which provides hourly electricity generation from different energy sources, as reported by the 76 different balancing authorities that operate in the U.S. Each balancing authority covers a certain geographical area. However, many of these balancing authorities interconnect with each other and, therefore, share the generation source of electricity. Hence, we follow the EPA eGrids zones mapping used for emission data [24]. eGrid divides the U.S. into 25 subregions and maps the balancing authorities to these subregions. We aggregate the balancing authority-level data into the eGrid subregion to determine electricity generation from different energy sources for a particular subregion.

4 PRELIMINARY ANALYSIS OF THE DATASET

Time series visualization. Fig. 4 shows the time series daily average direct and indirect WUE for several U.S. cities. For the indirect WUE, we use the fixed approach model presented in Eqn. (2). We see a seasonal impact on the time series data where winter months show a larger day-to-day variation than summer months for direct WUE. The indirect WUE, on the other hand, does not reveal such seasonal impact prominently.

Spatial variation in WUE. Fig. 5 shows the direct and indirect WUE for several U.S. cities in 2023. This box plot extends from the first quartile (Q1) to the third quartile (Q3) of the data, with a black line at the median and the mean represented using a diamond marker. The whiskers extend from the box by 1.5x the inter-quartile range (IQR) We see differences in both direct and indirect WUE due to differences in weather and the local mix of energy sources. We see particularly larger variations in indirect WUE across different locations.

Daily variations. In Fig. 6, we show the maximum daily variations (difference between the maximum and minimum WUE over a single day) of WUEs in different locations. We see that, on average, the direct WUE varies as much as 20% in places like Cheyenne. Meanwhile, the indirect WUE routinely varies more than 25% across most locations shown in Fig. 6.

5 EXAMPLE APPLICATIONS

We provide three example applications that may benefit from our dataset: EV charging, building load management, and geographical load balancing in data centers.

EV charging. EV charging can benefit from integrating water footprint into their operation sustainability. EV charging mainly involves indirect water consumption due to its electricity consumption. EV charging activities also typically offer greater scheduling flexibilities [10]. Given the scheduling flexibility inherent in EV charging activities, implementing a water-aware charging schedule can optimize EV charging during periods of low indirect WUE, resulting in significant indirect water consumption savings.

Building load management. Commercial buildings commonly feature sizable centralized cooling systems employing cooling towers, resulting in both direct and indirect water consumption. Consequently, managing such buildings' cooling load and energy consumption can integrate the building's varying water efficiencies to reduce its overall water consumption.

Geographical load balancing in data centers Water-aware scheduling and geographical load balancing offer substantial benefits to data centers and cloud applications [11, 12]. Data center workloads and cloud applications can exploit both temporal and spatial variation in the direct and indirect WUE. Unlike building load or EV charging, data center workloads can be moved around to more water-efficient locations. Many data center workloads, such as machine learning training, also offer great temporal scheduling flexibility and can be executed during water-efficient hours.

6 CONCLUDING REMARKS

In this paper, we introduced an hourly operational water efficiency dataset that captures the direct water consumption in the cooling system and indirect water embedded in the electricity generation. Additionally, we presented cooling system models capturing the impact of the weather conditions. We presented a preliminary analysis of our dataset highlighting the inherent temporal and spatial variation in direct and indirect WUEs. Furthermore, we discussed three potential applications that can benefit from utilizing our water efficiency dataset.

ACKNOWLEDGMENTS

This work is supported in part by the US National Science Foundation under grants ECCS-2152357, CCF-2324915, and CCF-2324916.

REFERENCES

[1] P. Gupta, M. Hossen, P. Li, S. Ren, and M. Islam, "A dataset for research on water sustainability," *Open Science Framework (OSF)*, https://osf.io/g3zvd/?view_only=63e9c2f0cdf547d792bdd8e93045f89e.
[2] "The water resources group: Background, impact and the way forward," http://www3.weforum.org/docs/WEF/WRG_Background_Impact_and_Way_Forward.pdf.
[3] U.S. Drought Monitor, "Map Archive." https://droughtmonitor.unl.edu/Maps/MapArchive.aspx.
[4] Palm Beach Daily News, "Drought dilemma: Palm Beach leaders struggle to cap water use," https://www.palmbeachpost.com/story/weather/2011/06/09/water-managers-shortage-near-crisis/7179447007/.
[5] EPA, "Watersense," https://www.epa.gov/watersense.
[6] Google, "Environmental report," 2023, https://sustainability.google/reports/.
[7] G. Ramos, M. Squicciarini, and E. Lamm, "Making ai ethical by design: The unesco perspective," *Computer*, vol. 57, pp. 33–43, feb 2024.
[8] Chiller & Cooling Best Practices, "Reducing cooling tower water consumption through advanced water treatment technology," https://coolingbestpractices.com/technology/cooling-towers/reducing-cooling-tower-water-consumption-through-advanced-water-treatment.
[9] EIA, "U.s. electric power sector continues water efficiency gains," https://www.eia.gov/todayinenergy/detail.php?id=56820.
[10] B. Alinia, M. H. Hajiesmaili, and N. Crespi, "Online ev charging scheduling with on-arrival commitment," *IEEE Transactions on Intelligent Transportation Systems*, vol. 20, no. 12, pp. 4524–4537, 2019.
[11] Z. Liu, M. Lin, A. Wierman, S. Low, and L. L. Andrew, "Greening geographical load balancing," *IEEE/ACM Transactions on Networking*, vol. 23, no. 2, pp. 657–671, 2014.
[12] M. A. Islam, K. Ahmed, H. Xu, N. H. Tran, G. Quan, and S. Ren, "Exploiting spatio-temporal diversity for water saving in geo-distributed data centers," *IEEE Transactions on Cloud Computing*, vol. 6, no. 3, pp. 734–746, 2016.
[13] P. Reig, T. Luo, E. Christensen, and J. Sinistore, "Guidance for calculating water use embedded in purchased electricity," *World Resources Institute*, 2020.
[14] P. Reig, "What's the difference between water use and water consumption?," *World Resources Institute Commentary*, 2013, https://www.wri.org/insights/whats-difference-between-water-use-and-water-consumption.
[15] J. Macknick, R. Newmark, G. Heath, and K. Hallett, "A review of operational water consumption and withdrawal factors for electricity generating technologies," *NREL Tech. Report: NREL/TP-6A20-50900*, 2011.
[16] G. G. Protocol, "Greenhouse gas protocol," *Sector Toolsets for Iron and Steel-Guidance Document*, 2011.
[17] U. Hölzle, "Our commitment to climate-conscious data center cooling," 2022, https://blog.google/outreach-initiatives/sustainability/our-commitment-to-climate-conscious-data-center-cooling/.
[18] J. Macknick, R. Newmark, G. Heath, and K. C. Hallett, "Operational water consumption and withdrawal factors for electricity generating technologies: a review of existing literature," *Environmental Research Letters*, vol. 7, no. 4, p. 045802, 2012.
[19] E. Spang, W. Moomaw, K. Gallagher, P. Kirshen, and D. Marks, "The water consumption of energy production: an international comparison," *Environmental Research Letters*, vol. 9, no. 10, p. 105002, 2014.
[20] SPX Cooling, "Water usage calculator," http://spxcooling.com/green/leed/water-usage-calculator/.
[21] "Historical Weather," http://www.wunderground.com/.
[22] R. Stull, "Wet-bulb temperature from relative humidity and air temperature," *Journal of applied meteorology and climatology*, vol. 50, no. 11, pp. 2267–2269, 2011.
[23] U.S. Energy Information Administration, "Open data," https://www.eia.gov/opendata/.
[24] U.S. Environmental Protection Agency, "eGRID data explorer," https://www.epa.gov/egrid/data-explorer.