The water-energy nexus in the United States: Environmental footprints of cities and data centers

by

Md Abu Bakar Siddik

B.S., Bangladesh University of Engineering and Technology, 2016

A THESIS

submitted in partial fulfillment of the
requirements for the degree

MASTER OF SCIENCE

Department of Civil Engineering
Carl R. Ice College of Engineering

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2020

Approved by:

Major Professor
Dr. Landon Marston

Abstract

Rapid changes in regional water cycle, and accelerated water use by the energy sectors have highlighted the need for holistic understanding of "Water-Energy nexus". Along with water consumption, high amount of green house gas (GHG) emission associated with energy generation serves as the polluting strength of any energy consumer. Environmental footprinting methods at sectoral or geographical level provide a means to relate the environmental externalities of electricity production to electricity consumers. Though several methods have been developed to connect the environmental footprint of electricity generation to end users, estimates produced by these methods are inherently uncertain due to the impossibility of actually tracing electricity from the point of generation to utilization. Previous studies rarely quantify this uncertainty, even though it may fundamentally alter their findings and recommendations. Here, we evaluate the sensitivity of water and carbon footprints estimates among seven commonly used methods to attribute electricity production to end users. We assess how sensitive water and carbon electricity footprint estimates are to attribution method, how these estimates change over time, and the main factors contributing to the variability between methods. We evaluate the water and carbon footprints of electricity consumption for every city across the contiguous United States for all assessed methods. We find significant but spatially heterogeneous variability in water and carbon footprint estimates across attribution methods. No method consistently overestimated or underestimated water and carbon footprints for every city. The variation between attribution methods suggest future studies need to consider how the method selected to attribute environmental impacts through the electrical grid may affect their findings.

We have implemented the general understanding and findings of the water-energy nexus at the sectoral level for thorough investigation at industry scale. Spatial dependency of sectoral demands for limited environmental resources require adequate attention as the competition for ever shrinking resources are on the peak. Data centers comprised of computer systems and related components represent one of the largest and fastest growing energy users in the United States. Vigorous effort from the researchers have been able to restrict the energy requirement growth to 6% compared to a sixfold growth in workload and computing demand in the past decade. Predicted a more ferocious bloom in near future, comprehensive study on the environmental stress exerted at a higher spatial resolution by these data centers requires imminent attention. The quantitative analysis found that more than 500 million m^3 water is consumed annually to support the operational stage of data centers with a high dependency on arid south western region. Geographical distribution of the servers coinciding with water stressed subbasins have almost tripled the water scarcity footprint (WSF) of the consumed water. Furthermore, present state of data centers is considerable source of greenhouse gas (GHG) emission, accounting for almost 0.6% of overall emission in the US. The results are validated by sensitivity analysis based on a set of electricity attribution approaches commonly found in existing literature. Finally, a comparative approach optimizes relative environmental stress of a hypothetical data center with no added technological innovation across the US subbasins to support decision making for future installation of server bases.

Table of Contents

List of Figures

List of Tables

Acknowledgments

Christopher M. Chini and Landon Marston conceived the idea of chapter 2. Md A. B. Siddik conducted the analysis. C.M.C. created the figures and visualizations. For chapter 3, L.M. conceived the study. MD conducted the analysis. Arman Shehabi provided data and fundamental concepts regarding the analysis. All authors contributed to the writing of the manuscript. L.M. acknowledges support by the National Science Foundation Grant No. ACI-1639529 (INFEWS/T1: Mesoscale Data Fusion to Map and Model the US Food, Energy, and Water (FEW) system). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation. All data used in this study come from public sources. Data produced through this research can be found in the Supporting Information.

Dedication

My family and friends for their continuous support.

Chapter 1

Introduction and Motivation

Energy and water are intimately interrelated, such that the use of one, by default, initiates the use of the other[2]. ‘Water–energy nexus’ refers to the critical interconnection between energy and water for production and consumption tradeoffs between these two resources to fulfill the human demands[3]. To facilitate the rapid growth of development of our modern society, electricity consumption is expected to experience an increase of 70% by 2035[4,5]. As the competition for electricity experienced a rapid growth and the resources used to generate electricity became scarce, policy makers started to become aware of energy saving, proper management and utilization of it. Also, increasing awareness about the environment motivated the researchers to look for alternative sources of producing clean energy and minimize the environmental impact. "Energy and Water Research Integration Act" was enacted as law in 2010 to ensure efficient, reliable, and sustainable energy–water policy coupling[3]. Power generation sector is the second highest consumer of water, and initiates highest amount of water withdrawal within the US[6]. At the consumption point, utilization process of electricity often initiates further water use. Numerous river basins are experiencing water stress while satisfying the increasing energy demand of human society and buffering the environmental impacts generated from human activity[7]. Almost all the renewable water sources globally are already serving some commodity production or community, and the introduction of new water consumer or increase the intensity of existing users will result in overcommitment of the water resources[8]. As the water basins of the United States are at risk of facing severe water stress in the coming decades[9], evaluation of sectoral and regional water use at greater spatial detail, and potential water savings at sectoral and spatial level will be the first step towards sustainable water allocation[10]. The interlink between water and energy makes it more important to study the sectoral and spatial pattern of energy use in order to manipulate the water-energy tradeoffs and identify the process chain with minimum water footprint. Also, power plants are second highest emitter of greenhouse gas (GHG) emission in the US and fossil fuel operated power plants are the major contributor to this emission[11]. Environmental impact of any sector in form of global warming is proportional to the amount of GHG emission associated with it[12]. Any consumer of electricity initiates embedded water consumption and emission of GHG emission during the generation of that amount of electricity.

Cities are increasingly the focus for sustainable transitions as they are central points of consumption and account for a significant portion of energy use and emissions[13,14]. Due to their importance in sustainable transitions and resource efficiency, it is important to understand material and energy demands of cities[15]. Footprinting methods offer one way of understanding the resource demands of cities through the lens of resource coupling and life-cycle processes. Attributing environmental footprints of electricity production through the electrical grid to the final consumer is challenged by data uncertainty, incongruent scales of production and consumption, and traceability within the electrical grid[16]. Urban footprint studies are growing in the literature, comparing resource demands by evaluating direct and indirect resource consumption of cities[17]. However, when accounting for electricity footprints, studies are often disparate and utilize methods of varying degrees of intensity with different assumptions. What are the differences between these methods and how sensitive are the results of the water footprint to these methods? In this study, we focus on empirical methods that use both simple emission factors and trading models used in water footprint attribution methods from across the literature and their resulting water footprints of electricity for urban areas across the United States. The geographical attribution boundaries selected here for empirical methods ranges from political to electricity management to hydrological boundaries (see Table 1.1).

We have used U.S. metropolitan area (defined by the United States Census Bureau) for environmental stress assessment due to their high energy intensive nature, and importance in US economy. We evaluated the water and carbon footprint associated with electricity consumption in each MSA for all these attribution methods and evaluated each method's unique set of pros and cons based on their considerations, computational requirements, and prioritization of impacts. Using metropolitan statistical areas (MSAs) as our urban boundaries, we investigated the sensitivity of attributed water footprints for each method for the years 2014 to 2017 and carbon footprints for 2014 and 2016, based on available data. The combination of these two footprints provides a unique perspective on the methodologies of attributing environmental footprints of electricity to urban areas. The results of this analysis help understand the data requirements and refine the methods used in environmental footprinting analyses. These methods are essential components in developing a systems understanding of the life-cycle impacts of electricity. Previous literature often focuses on the uncertainty of the data in attributing environmental footprints; however, we illustrate the importance of also considering the method and its inherent uncertainty. We estimate both water and carbon footprints of electricity, evaluating how tradeoffs may exist between the two resources as both are important in understanding the sustainability of consumption patterns. Additionally, we evaluate interannual variability of the electric grid and its resources, while providing all our data to aid future research. The combination of evaluating water and carbon footprints across time, space, and accounting for method variability fills an important knowledge gap in the literature.

While this comparative study provides a general understanding of the uncertainties associated with environmental footprinting, and ranges for water and carbon intensities of

Table 1.1: *We evaluate seven common empirically-based methods to attribute the environmental footprint of electricity to end users. The advantages and disadvantages of each method, as well as studies that have employed each method, are shown below.*

Method	**Advantage**	**Disadvantage**	**Employed by**
Interconnections	Conforms to electricity infrastructure; Minimum data requirements and calculations	Large area; does not prioritize local impacts	Ruddell et al. [18]
Balancing Authority	Geographically smaller than interconnect	Pass-through nodes; non-specific geographical areas	Cohen and Ramaswami [19]
Balancing Authority with Transfers	Conforms to electricity infrastructure; illustrates burden shift of resources	Time consuming; disparate datasets	Kodra et al. [20], Chini et al. [21], Djehdian et al. [15]
EPA eGRID Boundaries	Conforms to data used for emission assessments and electricity infrastructure	Data only available every two years; requires integration with EIA data for water resources	Peer et al. [22]
Basin Scale	Conforms to natural hydrology	Does not consider infrastructure	Kelley and Pasqualetti [23], Tidwell et al. [24]
Radius from City	Accounts for local impacts	Does not consider infrastructure	Chini et al. [25]
State	Policy and regulations often set at state level; EIA aggregates data at state level	Cities in some states are supplied by different providers (i.e., Chicago, IL)	Bartos and Chester [26], Chini et al. [27], DeNooyer et al. [28], Grubert and Webber [29], Stillwell et al. [2]

the urban areas, it also provides scope for consumer level assessment. A wide variety of electricity users at different scale is one of the main features of cities. Estimation of environmental stress at geographical level offers an scope for the policy makers to initiate or reform policies regarding resource use. For example, water discharge permits and allowable thermal pollution are decided at state level. However, sectoral analysis of resource use provide scope for stakeholders in optimal decision making. Environmental footprint associated with the selected technological approach or geographical location of a industry may potentially alter the decision making of the owners and organizations. If we inspect the sectoral consumers of electricity, Information and communication technologies (ICT) will be along the top rows. ICT dominates almost all the sectors of human lives, and shape the economy in many ways. Data centers lie in the core as a backbone to this ICT boom that support the growth and opportunities of an automated and interconnected globalization[30]. It first caused concerned to the policy makers in the early 2000's when Koomey[31] estimated a global electricity consumption of 153 TWh by data centers, representing 1% of total global electricity use. That was only the beginning of the bloom as within a decade, the industry saw a sixfold growth in workload and computing demand[30]. Growing criticism of the formidable energy use forced the technologist and energy analysts look for sustainable solution on the improved energy use. Studies built upon these primal works to decouple the energy uses by data center equipment. Servers are the basis of a data center, while all the auxiliary equipment are installed to support and ensure the workability of servers. Replacing traditional servers with high processing blade servers[32,33], virtualization of servers to improve scalability from 10% to around 50% of maximum processing capacity[34,35,36], introduction of cloud based services[37,38,39,40] have significantly controlled the forecasted growth in number of servers, which defines almost half of total energy requirement by a data center. Cooling infrastructure consumes almost all the electricity use denoted to auxiliary equipment. Improved air management by hot isle- cold isle system[41,42], replacing dry coolers with cooling tower[43], installation of direct contact cold plate or liquid cooling to overcome low thermal capacity of air[44,45], introduction of air-side and water-side economizers[46,47] are among the numerous improvements that significantly reduced the cooling energy requirement in data centers. Improved processing power of servers and efficient cooling system that occurred in parallel to the rapid growth of data centers have been able to throttle the exponential growth in energy use to some extent. Efforts have been paid off as energy use increased by only 6% compared to a 550%increase in workload[30].

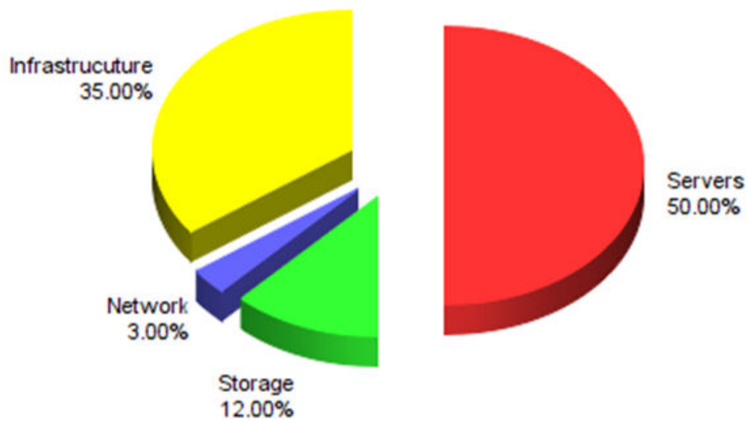


Figure 1.1: *Electricity consumption of data center components (based on the findings of Shehabi et al.[1]).*

However, the energy use by data centers still constitute as a major energy consumer globally, and all the existing efficiency measures are almost utilized to the fullest. The analysts are skeptic of this sustainable energy trend as growth of data center are expected to continue with more ferocity, and next doubling in global computing demand is expected within the next 3-4 years[48,49]. Although the energy use and efficiency of data centers is being analyzed vigorously, studies on environmental impact of this data centers is relatively dearth. In recent years, studies are focused on sectoral impact on the environment to create awareness among both the operators and consumers of goods and services. People are increasingly concerned of their environmental footprint, and look for opportunities to choose option with minimal footprint. Therefore, it is critical that we assess the sectoral contribution of water and carbon intensive sectors, and provide comparative analysis for different options available as a benchmark for the interested communities. United States houses almost 30% of the total server inventory, responsible for almost 2% of the total electricity consumption in the US. Here, we tried to explore the spatial distribution of water consumption and green house gas emission at the operational stage of data centers. We also provided comparative assessment for environmental stress of new server installation at the watersheds of contiguous US.

Onsite water use is initiated by the cooling system to dissipate the immense heat generated by the servers. But a significant amount of water is consumed in form of electricity use by the data center equipment. Electricity consumption by these data centers also contribute to the greenhouse gas (GHG) emission. Assigning generative source of consumed electricity at the data centers introduces uncertainty to this study. Generated electricity gets mixed within the electric grid and reaches the consumers, making it impossible to find the generative source of consumed electricity. Water use and GHG emission of the generated electricity can vary significantly depending on the fuel type, technology, and cooling method implemented at the power plant level. Also, using the national average value will fail to capture the intricacies of local impact for our spatial analysis. Researchers over the years used a wide range of attribution methods and boundaries for electricity consumption (Table 1.1). The boundaries ranges from electricity management to geopolitical to hydrological boundaries. Previous studies[50,51] mainly focused on the uncertainties arising from the underlying data used for water consumption and GHG emission. We have performed a parallel study to assess the sensitivity of environmental stress for unit consumption of electricity aggregated on commonly used geographical attribution boundaries.

In short, this study focuses on one way relationship between water-energy nexus, i.e. embedded water use of the consumed electricity at sectoral and geospatial level. Along with water footprint, we have also attributed the emission associated with with electricity generation to the end users. Uncertainty associated with tracing the generative source of electricity prohibits us using any simple or universal attribution method. We performed sensitivity analysis of the commonly used geographical attribution method to show how the burden shift of environmental stress from electricity generator to consumer may vary depending on the selected attribution method. Our estimated water and carbon footprint of the MSAs will help the policy makers better decision making as majority of the social and economic activities are centered on this MSAs. MSAs are often comprised of a clearly defined and diverse set of electricity end users with varying degree of resource use. Therefore, we chose data centers as an example for sectoral level environmental stress. Data centers are important environmental stressor as they consume significant amount of electricity embedded water, and at the same time a huge amount of direct on site water consumption for cooling purpose. We delineated the dependency of data centers on distant watersheds for electricity embedded water supply, and burden shift of GHG emission. This approach can be replicated for any industry to show their spatial dependency for water consumption and carbon emission. Finally, our comparative analysis shows how choice of location can minimize the environmental footprint of future installation of data centers.

Chapter 2

Water and carbon footprints of electricity are sensitive to geographical attribution method

2.1 Introduction

Electricity production is the largest emitter of greenhouse gases (GHG)[52] and the second largest water consumer[53,54], globally. Environmental footprinting methods, as defined by Hoekstra and Wiedmann[55], offer one way of understanding and quantifying the direct and indirect pressures of electricity[56]. However, data uncertainty, incongruent scale of production and consumption, and traceability within the electric grid challenge robust attribution of environmental footprint of electricity production to the final consumer. Researchers have developed numerous environmental footprint attribution methods to overcome some of these challenges within both the water footprint[18,21,24] and carbon footprint[57,58,59] communities. Yet, there remains a great deal of uncertainty as to how sensitive results are to attribution method and how this sensitivity differs between different footprint indicators.

Here, we conduct a comparative study of common approaches to estimate the environmental footprint of electricity consumption to test how sensitive water and carbon footprints of electricity consumption are to geographical attribution method. While previous studies often focus on the uncertainty of the underlying data used to calculate environmental footprints[50,60], we demonstrate the importance of also considering the impact of the method selected to attribute environmental footprints of electricity production to consumers. We focus on commonly used bottom-up approaches to estimating different footprints (as opposed to top-down approaches, such as environmentally-extended multi-regional input-output models; e.g., Mo et al.[61], Tian et al.[62]) . Environmental footprints associated with electricity production are assigned to end consumers with the same or connected geopolitical, infrastructure, or natural boundaries (e.g., state, electricity grid, or watershed). Henceforth, we refer to geographical attribution boundaries simply as 'attribution boundaries'. We ask and answer the following three questions: i) how sensitive are water and carbon footprints of electricity estimates to attribution method? ii) does variance between attribution methods differ between areas and within an area over time? and iii) what factors contribute to variability between attribution methods and do these factors differ by environmental footprint type?

Attribution methods can be classified into two general types: i) empirical data models and ii) power system optimization models[63]. Empirical models use historical observations to calculate emission factors, trading models, or statistical relationships to connect environmental footprints of electricity production to electricity consumption. Trading models incorporate additional data to account for imports and exports of electricity across specified boundaries[64,65,66]. Power system optimization models determine embedded resources based on power distribution networks and economic optimization. In this study, we compare empirical methods that use both simple emission factors and trading models as these are the most commonly used environmental footprint attribution methods. Moreover, many power system optimization models are proprietary, making comparison of these methods infeasible. The empirical methods evaluated rely on different geographical, infrastructure, and political boundaries, including interconnections, balancing authorities, EPA eGRID, river basins, state, and radius from cities (see Table 3.1). Regardless of attribution boundary, all methods utilize the same underlying data.

We calculate the water and carbon footprint associated with electricity consumption in each metropolitan statistical area (MSA; as defined by the U.S. Census Bureau) within the contiguous United States using the most common empirically based attribution methods. Though any electricity consumer could be used in this study, MSAs provide a clearly defined and diverse set of electricity end users. Further, cities are integral in achieving environmental sustainability and climate change mitigation targets as they are central points of consumption and account for a significant portion of energy use and emissions[67,14]. Urban areas consume around three-fourths of global energy, with electricity being the second largest energy source, as well as the fastest growing energy use[68,49]. Nearly two-thirds of the 43 cities evaluated by Cohen and Ramaswami[19] imported over half of their electricity, demonstrating how metropolitan areas' resource consumption and environmental impact stretch well beyond their geopolitical boundaries. While some studies have evaluated the carbon footprint of a city's electricity consumption[69,70,71], we have a more limited understanding of how cities draw on local and non-local water resources to fulfill their electricity demand[27].

The following section provides background on environmental footprints of electricity production and how these footprints are assigned to end consumers. Next, we describe the methodology employed in this study, followed by our results. Lastly, we discuss our findings and the implications they have on future research, as well as cities, companies, and other groups that want to determine their water and carbon footprints of electricity consumption.

Importantly, all water and carbon footprint estimates for every US MSA is published with this study to support future research and aid electricity consumers in determining the water and carbon footprints of their electricity use.

2.2 Background

The electric grid in the United States is divided into three main interconnects: Western Interconnection, Eastern Interconnection, and the Electric Reliability Council of Texas (ERCOT). The Eastern and Western Interconnect are comprised of 31 and 37 balancing authorities, respectively[72]. ERCOT consists of a single balancing authority. Each balancing authority balances electricity supply and demand in real-time to ensure system reliability. Power plants are distributed across each of these interconnections, supplying electricity to the grid. Depending on the fuel source and technology employed, power plants emit significant amounts of GHGs. Further, power plants impact local water resources through their large water withdrawals. A portion of water withdrawals are evaporated and removed from the local water system, while the rest are returned to the water body at elevated temperatures leading to thermal pollution and ecological damage[73,74]. Attributing these local impacts to end consumption shows the burden shift of electricity demand to production locations.

The transmission of electricity through the electric grid creates difficulties associated with attributing water and carbon footprints of electricity generation to end-users. Previous work[16,63] has explored how these inherent challenges may impact the attribution of carbon footprints to different electricity users, but no study has evaluated the impact on water footprint estimates. Further, no studies, to our knowledge, have evaluated carbon and water footprints together to understand the resource demands of all urban areas across an entire nation. Ryan et al.[63] and Weber et al.[16] highlight the variation and assumption of multiple attribution methods with respect to emissions, concluding that the study objective often motivates the method choice. Within the United States, all empirically-based methods rely on power plant level data reported by the Energy Information Administration (EIA). Each power plant is mapped to the particular attribution boundary of interest. The data within the Energy Information Administration are self-reported via Form 923 and come with their own sets of uncertainty (though data quality has improved markedly in recent years[51]). Quantifying the uncertainty of the underlying reported data has been evaluated by others[51,50] and is outside the scope of this study.

Here, we highlight seven different methodologies for attributing electricity-related water consumption and carbon emissions to electricity consumers within each US MSA. The complexity of the electrical grid and the impossibility of tracing an electron through it means there is no 'correct' attribution method, and it is impractical to consider one estimate better than others. Instead, each environmental footprint attribution method of electricity has distinct advantages and disadvantages (Table 3.1). Each empirical method employs different geographic boundaries, which draw on a different collection of power plants (Figure S1). Approaches using interconnections, balancing authorities, or eGRID boundaries consider, to varying degrees, the physical infrastructure of the electrical grid. The interconnect boundary represents the largest geographic scale and is the simplest to calculate, while methods utilizing the balancing authority scale are more computationally intensive and require integration across multiple databases (EIA, Environmental Protection Agency, and the Federal Energy Regulatory Commission). The eGRID scale also offers some smaller scale regional attribution and varies slightly from the boundaries of balancing authorities. The eGRID boundary was designed to promote consumer-scale or regional decision-making capability. The basin scale and geographical radius boundaries attempt to localize impacts of the water footprint of electricity production by evaluating the removal of water resources from the immediate environment. The state scale method has advantages in that it follows policy boundaries for water discharge permits and thermal pollution. However, basin, radius, and state boundary methods can overlook some of physical constraints of electricity distribution through the grid.

In this study, we consider the water consumed and carbon and carbon equivalents emitted (henceforth, denoted simply as 'carbon') during the operational stage of electricity generation. Roughly two-thirds of water consumption in the life cycle of electricity production occurs during the operational stage of electricity generation [22,75]. Similarly, the operational stage of electricity generation constitutes 83-99% of the total GHG emissions associated with fossil fuel-based electricity production [76]. Environmental footprint assessments use physical or monetary units to normalize the footprint in terms of production (e.g., Marston et al. [77] uses both units). When determining the water or carbon footprint of electricity, water consumption or GHG emissions are most often normalized by energy units, which we adopt in this study. Our analysis evaluates how sensitive our results are to temporal dynamics by using available water consumption data (years 2014 to 2017) and GHGs emissions data (years 2014 and 2016).

2.3 Materials and Methods

Attributing water and carbon footprints of electricity requires two steps. First, it is necessary to determine the water or carbon footprint per kWh of delivered electricity, i.e., volume of water per kWh and mass of carbon per kWh (intensity). The water and carbon footprint per unit of delivered electricity is largely a function of the power plants assumed to service the area of interest. Second, one must determine the electricity demand of the city or entity of interest. In this study, we focus on the first step and the various methods to estimate water and carbon footprints per unit of electricity generation. The following sections describe the methods and data needed to replicate each of the seven attribution approaches most commonly employed in the literature.

2.3.1 Electricity generation and environmental footprint data

Electricity generation and water consumption data were taken from self-reported generator observations, which are collected and tabulated by the Department of Energy's Energy Information Administration (EIA)[78]. While the quality of EIA data has been questioned[79,51], it provides detailed data at a fine spatial resolution and is the data set most commonly used in studies aiming to estimate the environmental footprint of electricity production and consumption. Besides, the purpose of this study is to compare different attribution methods, meaning it is of greater importance that each attribution method utilize the same data across all methods. Power plants with generation capacity greater than 100 MW are required to report their water consumption to EIA[80]. These large power plants contribute almost 75% of the United States total electricity generation[81]. Smaller power plants (generation capacity less than 100 MW) are required to report their energy production but not their water consumption to the EIA. These smaller power plants are included within our study by assigning the median value of water consumption calculated from the reporting power plants to all small power plants with similar fuel type and generation capacity less than 100 MW. EIA does not have water consumption data for renewable energy sources, such as wind, solar, or hydropower. Average water consumption values based on detailed engineering studies were used for solar and wind operated renewable power plants[82,75]. Water consumption attributed to hydroelectric power is related to reservoir evaporation and is often many times the magnitude of other types of power plants[75]. Water footprints of hydroelectric power plants are taken from Grubert[83], which considers the multiple users of a reservoir (e.g., irrigation, flood control, hydropower) and allocates the evaporative losses across these users so to avoid the overestimation of hydropower water consumption.

We utilize the most recent versions of the EIA Form 923 (annual values from 2014-2017) and the Environmental Protection Agency (EPA) tabulated emissions from power generating facilities[84] to analyze the temporal variability of water consumption and carbon emissions within each metropolitan area for a given attribution method. We utilize EPA's

Clean Air Markets Division (CAMD) data[84] on observed emissions from stack monitoring, as opposed to EIA's modeled emission estimates[11], to estimate carbon footprints. Carbon footprints are calculated using equivalent carbon dioxide weights, CO_{2e}.

With respect to water resources, we take a water footprinting approach to assess water intensity of electricity based on attribution method. We recognize that there are other approaches to assess the environmental impacts of water resources, specifically with respect to water scarcity (e.g., International Organization of Standards ISO 14046). This life-cycle assessment method is outside the scope of the current study.

2.3.2 Attribution of electricity source to consumers

Following Kodra et al.[20], we aggregate power flow among the electricity generating units within the attribution boundaries under analysis. In general, there are two different types of data-driven attribution methods: (1) those based on grid infrastructure and (2) those based on geographical boundaries. Attribution methods based on grid infrastructure better constrain the production, transfer, and consumption of electricity to the underlying grid infrastructure and the companies that operate them, but these methods are limited by data and require a higher-order of computation. Boundaries defined by grid infrastructure, including interconnections, balancing authorities, and eGRID, are defined by EPA.

Geopolitical or geophysical boundaries do not match the actual flow of electricity along the grid, but national and state regulations and policies concerning water and GHG emissions are often mandated based on these boundaries. This makes geopolitical and geophysical boundaries particularly important when analyzing the burden they exert on the environment. For geopolitical and geophysical boundary-based attribution methods, an attribution boundary may have few or no power plants within its border. The electricity demand within that attribution boundary may well exceed the generation. To overcome this issue, we used an energy balance approach to match excess electricity generation to unmet electricity demand following the approach of Ruddell et al.[18]. Areas with electricity generation exceeding demand will make their excess electricity available to a 'collective pool' of surplus electricity that deficit areas can pull from the grid.

Both the grid-based and geographical boundary methods utilize the same generalizable equations to estimate the environmental resources or emissions intensity of electricity production (EIP).

$$EIP_i = \frac{\sum_x E_x}{\sum_x P_x} \tag{2.1}$$

$$EIP_{i*} = (EIP_i \times \alpha_i) + EIP_{i-interconnect} \times (1 - \alpha_i) \tag{2.2}$$

Here, EIP_i is the weighted averaged embedded environmental resources or emissions (E) of electricity production (P) of the power plants (x) within attribution boundary i. EIP_{i*} recalculates the embedded environmental resource or emissions intensity of electricity production within a geographical attribution boundary (e.g., state boundaries) when electricity transfers between attribution boundaries are considered. Since it is infeasible to consider actual electricity transfers across the grid with geographical attribution boundaries, electricity demand that cannot be supplied by power plants within the specified boundary will be fulfilled from excess electricity produced within the interconnect to which the attribution boundary i is nested within ($i - interconnect$). α_i is the ratio of electricity generation and consumption within attribution boundary i. α_i is capped at 1, which signifies that power plants within the attribution boundary are capable of fully meeting electricity demand within the attribution boundary (i.e., no electricity transfers occur). If electricity transfers across grid-based attribution boundaries are considered, the previous equation can be updated as follows:

$$EIP_{i*} = (EIP_i \times \beta_i) + \sum_j EIP_j \times (1 - \beta_{i,j}) \tag{2.3}$$

where β_i is the fraction of electricity produced within attribution boundary i to total production plus net imports of attribution boundary i. $\beta_{i,j}$ is the fraction of electricity imported into attribution boundary i from j to total production and net imports of attribution boundary i.

Finally, the embedded environmental resources or emissions of electricity consumption of MSA m (EIC_m) is determined by summing the product of each overlapping attribution boundary's EIP_i* and the proportion of MSA m geographical area (A_i) covered by the area of the attribution boundary ($A_{i,m}$).

$$EIC_m = \frac{\sum_i (EIP_i \times A_{i,m})}{A_m} \tag{2.4}$$

We used this general approach to estimate both water and carbon footprints and intensities of each MSA for all attribution methods. Further discussion on the individual methods and their underlying assumptions and data can be found in the supporting information.

Due to data limitations, our study focuses on the annual scale to assess both carbon and water footprints. While the EIA provides data at a monthly scale for several environmental impacts, we are limited in our study by datasets from the EPA (eGRID) and the Federal Energy Regulatory Commission (FERC). These datasets are only at the annual scale. We recognize that there are variations in renewables intra-annually which might affect the results, to an extent; however, for uniform comparison across methods, we aggregate EIA data and conduct the study on the annual scale.

2.4 Results

2.4.1 Sensitivity of carbon and water intensities to attribution method

Each metropolitan area demonstrates different levels of sensitivity to the attribution method for water and carbon footprints of electricity. The sensitivity of each metropolitan area to water and carbon attribution methods is quantified by the coefficient of variation (CV) and presented in Figures 2.1a and 2.1b, respectively. For water intensities, higher CV in urban areas of the Southwestern United States suggest water intensity of delivered electricity is highly sensitive to the attribution method (denoted by an orange or red color in Figure 2.1). This variation indicates diverse electricity production technologies (e.g. presence of large number of hydro and solar power plants in the same region) in the surrounding areas. Although water consumption of nuclear power generation is higher than other thermometric generations, the difference is insignificant when compared with hydroelectricity. The amount of water consumed in the production of electricity can vary based on several factors, including fuel type, combustion method, and type of cooling technology. Macknick et al. [79], Peer and Sanders [85], and others provide breakdowns of the water intensity based on these factors. Variability arises from change in energy generation mix portfolios of an MSA for different geographic attribution boundaries (shown in Table S4 of SI) based on the geographic location of the generating units.

The Mid-Atlantic and Northwestern regions of the United States have smaller CV, indicating that the water intensity values are not as sensitive to the attribution method. The relative consistency between estimates produced by different methods in these regions is due to a largely homogeneous electricity generation portfolio across all the attribution boundaries. Hydroelectric power plants are ubiquitous in the Northwestern US, resulting in a high but consistent water intensity for MSAs in the region. Large amounts of electricity, and therefore embedded water, are transferred between the states of California, Arizona, Colorado, Utah, and New Mexico within the Western Interconnect creating geographical dispersed dependencies on water resources [18,21].

Using the same attribution methods, we substitute water consumption for greenhouse gas emissions, represented by CO_2 equivalents (Figure 2.1b). These carbon intensity equivalents provide another way to evaluate these attribution methods. When evaluating the coefficient of variation of each attribution method across the country, there are localized areas of high variation between methods in the Southwestern and Northwestern United States. The Mid-Atlantic region's sensitivity of emission intensity, like water intensity calculations, has a

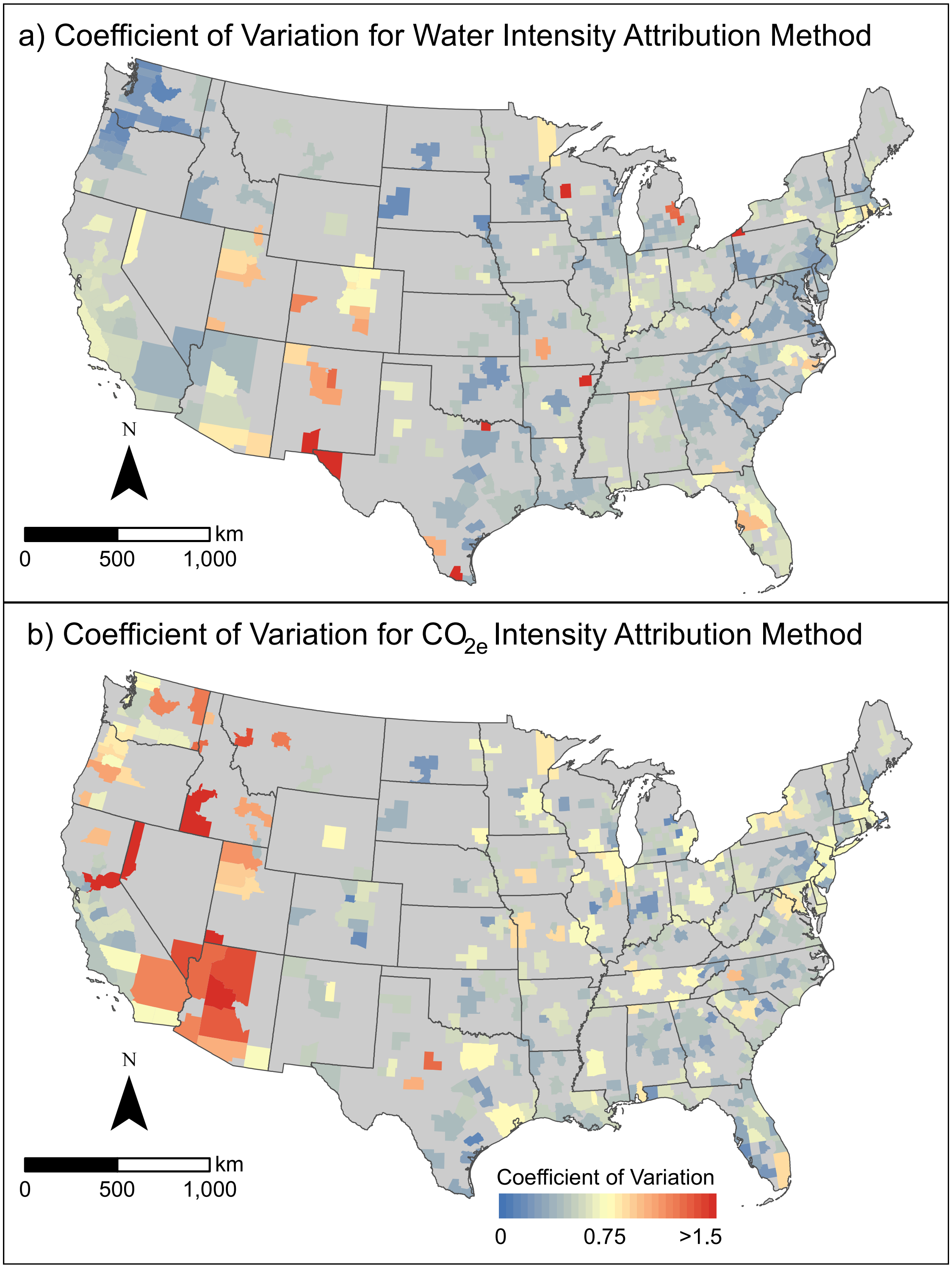


Figure 2.1: *The variability between attribution methods (represented here as the coefficient of variation, computed based on the results from each of the methods) is not constant across the country. A lower coefficient of variation (represented by blue) signifies agreement in estimates among the attribution methods, while a higher coefficient of variation (represented by orange or red) signifies a divergence between estimates produced by each method. The Mid-Atlantic and Northwestern regions show greater homogeneity in water and carbon footprints are not as sensitive to the attribution method as other regions of the country.*

relatively low coefficient of variation. However, in general there was no correlation between the coefficient of variation of water intensities and emission intensities. Further comparison of the coefficients of variation between water and carbon intensities can be found in the supporting information (Figure S2).

To further compare the attribution methods and their impact on water and carbon intensity calculations, we investigate the effect of electricity transfers between attribution boundaries on environmental footprints. We evaluated the impact of electricity transfers for three attribution methods: HUC–4, PCA/balancing authority, and state scales. Balancing the electricity demand from the surrounding interconnect changes the embedded resource intensity. In general, water intensity remained constant or increased for each of the three methods when including energy balancing (Figure S3). Conversely, carbon intensity of MSA's electricity use demonstrated a much wider range of change, with no clear increasing or decreasing trend when electricity transfers were considered. Carbon intensities vary more widely across power plants and attribution boundaries than water intensities, and this greater variation is the primary reason carbon intensities exhibit greater heterogeneity in response to electricity transfers than water intensities (Figure S3). Moreover, we found the carbon footprint of electricity consumption are more sensitive to the attribution method selected compared to the water intensities (95% confidence level). In general, the CV of carbon intensities are larger than water intensities. The CV of water and carbon follow a gamma distribution, with a long right tail signifying some MSAs exhibit much greater sensitivity to attribution method than their peers (Figure S2).

2.4.2 Trends across attribution methods

Analysis of the water intensity for the top 50 most populous MSAs shows significant variation across different MSAs for the same attribution method and within the same MSA with different attribution methods (Figure 2.2). Table S3 in the SI provides a list of the top 50 MSAs by population. For many of the most populous cities, the majority of the attribution methods produce similar results. However, for some of these cities there is a much wider spread of the estimated water intensity values. For example, the mean estimated water intensity of Buffalo, NY is approximately 40 m^3/MWh – nearly 7 times the average US city – and ranged from approximately 5–80 m^3/MWh, which is the second largest spread of water intensities across all MSAs. Numerous MSAs have one or more attribution methods that produce water or carbon intesnity estimates that are much higher than the average, though there is no singular method or set of methods that consistently results in larger or smaller water or carbon intensity estimations. The 50–km radius attribution method has the smallest water footprint for about one-third of MSAs, while one-third of MSAs had the interconnection as the largest water footprint. Interestingly, the HUC–4 boundary method produced the largest carbon intensity value for nearly half (48.7%) of all MSAs.

Although the selected urban areas show high sensitivity to the attribution method selected, the temporal variation of water intensity of delivered electricity is relatively constant across all urban areas (Figure 2.3 shows the 50 most populous US cities). In general, there is no significant difference between the four years within each MSA. This finding supports previous research[22] showing temporal variability of regional water intensity is minimal compared to change in fuel and technology mixes. Therefore, any changes seen are most likely due to an addition or retirement of a power plant included in the spatial boundary.

2.4.3 Factors contributing to variability between attribution methods

To further illustrate why different attribution methods may produce variation in environmental footprint estimates, we reexamine Buffalo, NY, which has a large spread in water intensities estimates by different attribution methods (approximately 5–80 m^3/MWh). We also investigate Chicago, IL, which has a relatively small spread of water intensities across attribution methods (approximately 1–5 m^3/MWh). Each of these MSAs are located on the borders of their respective state and at the intersection of multiple hydrologic boundaries.

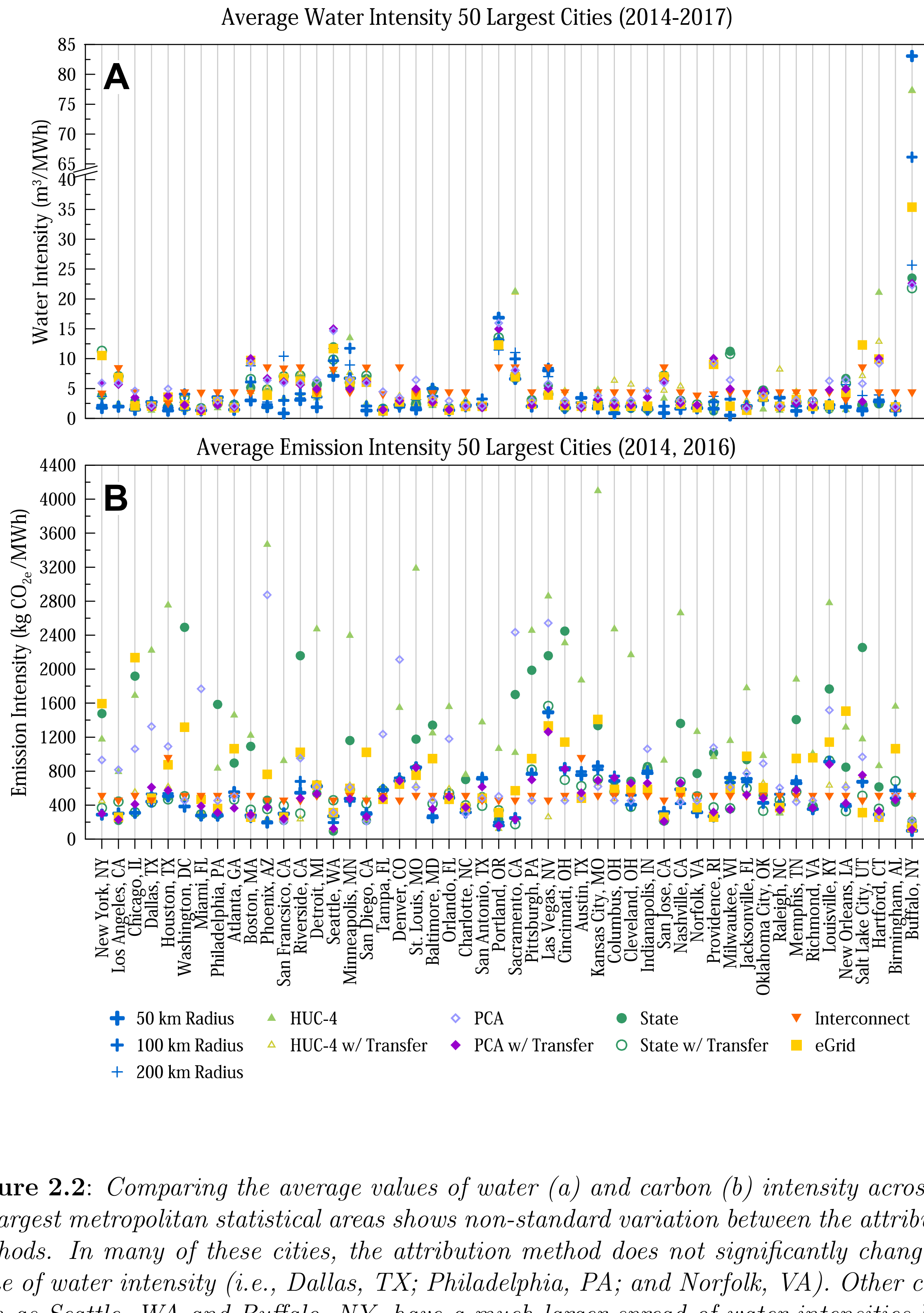


Figure 2.2: *Comparing the average values of water (a) and carbon (b) intensity across the 50 largest metropolitan statistical areas shows non-standard variation between the attribution methods. In many of these cities, the attribution method does not significantly change the value of water intensity (i.e., Dallas, TX; Philadelphia, PA; and Norfolk, VA). Other cities, such as Seattle, WA and Buffalo, NY, have a much larger spread of water intensities based on methods. The emissions intensity for the 50 largest cities varies widely depending on the attribution method. Additionally, methods that utilize HUC–4, PCA, and State boundaries generally produce larger estimates than other attribution methods.*

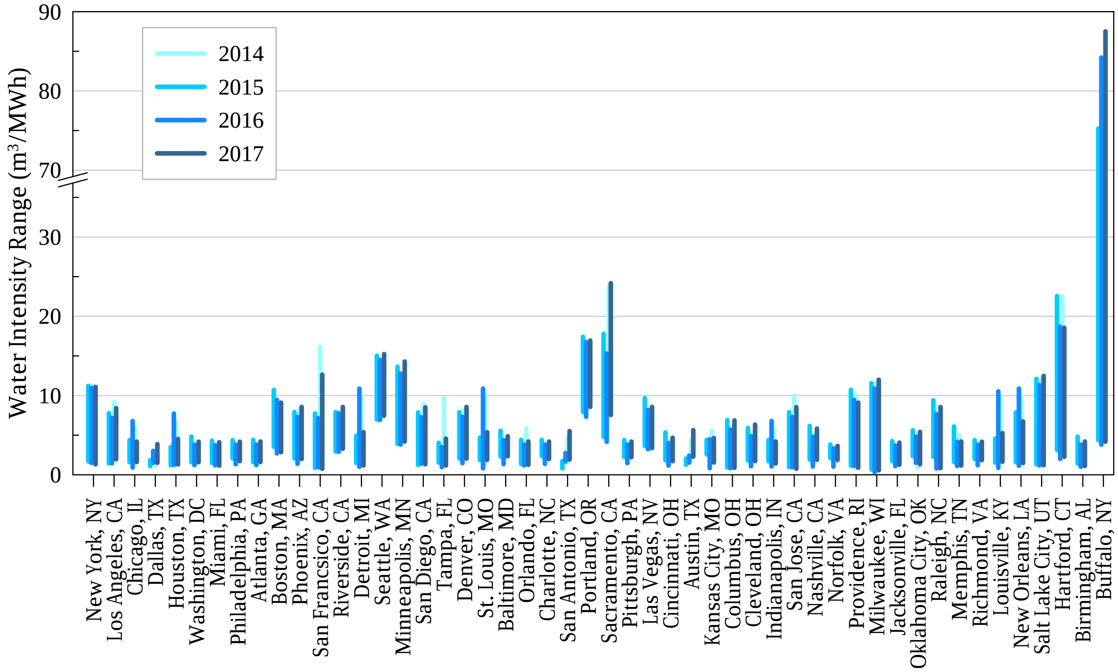


Figure 2.3: *Water intensities for the largest 50 metropolitan areas in the United States show little variation between years.*

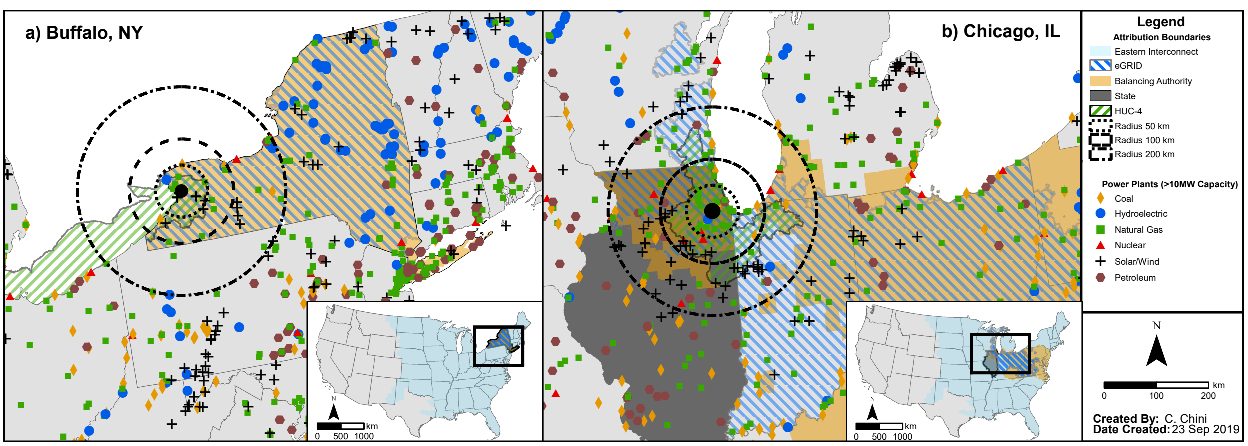


Figure 2.4: *Buffalo, NY (a) and Chicago, IL (b) demonstrate among the greatest and least variance in water intensities of electricity deliveries between attribution methods, respectively. Water intensity estimates for Buffalo are more sensitive to the attribution method due to the misalignment of attribution boundaries and the clustering of certain power plant types (namely, hydropower) within some attribution boundaries but not others. Conversely, attribution boundaries used to determine Chicago's water intensity have largely the same collection of power plant types, all producing similar water intensity estimates.*

Buffalo is located on Lake Erie on the western edge of New York, while Chicago is on Lake Michigan at the northeast edge of Illinois.

The large variance in water intensity for Buffalo comes from the diverging attribution boundaries and the diverse forms of power generation types clustered throughout the state (Figure 2.4a). In more general terms, when attribution boundaries do not significantly overlap it means a different set of power plants are assumed to supply a MSA's electricity. This assumption is particularly consequential in places like western New York, where a clustering of electricity generation technologies can dramatically shift estimates of water intensities depending on the set of power plants within the respective attribution boundary. For example, Figure 2.4a depicts solar/wind (low water intensity) and hydroelectric (very high water intensity) power facilities in northern New York, which are excluded in the HUC–4 boundary and the radii attribution methods but captured by other boundaries. Chicago shares many similarities to Buffalo (it also lies on the boundary of its state at the edge of the Great Lakes, with greatly diverging attribution boundaries); yet, Chicago has a much smaller variation in estimated water intensities across all attribution methods. Chicago's small variation can largely be explained by the relatively uniform distribution of different power plant types throughout the surrounding area (Figure 2.4b). Unlike Buffalo, there is not a clustering of particular types of power production that might sharply skew water intensity estimates upon inclusion of this area within an attribution method.

2.5 Discussion

We do not suggest a 'best' or 'correct' attribution method for environmental footprints of electricity. Instead, we contend it is important to understand the inherent assumptions associated with each attribution method and the degree that these methods produce different estimates. We suggest that the chosen method for attributing environmental footprints of electricity production to end users be selected based on the research problem posed. If the study, for example, aims to assess the impacts of state regulations or grid operation, a state or grid-based attribution boundary may be most appropriate. However, if the study is focused on local hydrologic impacts of electricity consumption or the opportunity cost of local water withdrawal, then the radius or HUC-4 attribution boundaries provide a better localized context of analysis. With that said, the methods most commonly employed in the literature to relate environmental footprints of electricity production to consumers do not explicitly consider the environmental impacts of freshwater appropriations (i.e., they do not follow the LCA approach set forth by ISO 14046). Future studies would benefit from assessing the environmental consequences of water consumption and GHG emissions, including water scarcity[86,87,88].

Regardless of what attribution method is deemed the most appropriate for a particular study, the potential large variations in environmental footprint estimates (as demonstrated in this study) highlight the need to use multiple attribution methods so to quantify the sensitivity associated with the primary attribution method selected. In areas that have a high sensitivity to attribution method, it is particularly important to characterize this variability and the assumptions associated with the chosen attribution method. Data uncertainty and sensitivity have previously been shown to have a non-trivial impact on estimates of environmental footprints[77]. Here, we demonstrate that the method selected to attribute the footprint of electricity generation to end users can also significantly shape estimates of a consumer's water and carbon footprints. Therefore, future studies relating the environmental impacts of electricity production to end users should incorporate some measure of variability associated with the selected attribution method. The differences in water and carbon intensity calculations produced by each method demonstrates the difficulty in formulating sound policy and decision-making based on one attribution method, as each can yield very different conclusions. An ensemble approach that balances these tradeoffs presents an opportunity to avoid bias associated with a selection of one methodology over another.

As urbanization and overexploitation of natural resources intensifies in the future, assessing and attributing the environmental footprint of electricity generation to cities will be critical to understand the telecoupling between production and consumption of electricity within the water-energy-carbon nexus. However, it is important that the scientific community converges on a means to attribute the environmental impacts of electricity production to end users so comparisons can be made across different studies and decision-making is based on robust findings. For example, a standardized approach for determining the carbon footprint of electricity use that quantifies uncertainty or variability of the estimates will be important as voluntary and mandatory carbon offset markets become more common. Cities, corporations, and other groups aiming to determine the environmental footprint of their electricity consumption should present sound reasoning for the attribution method they select and this methodology should be consistently applied across all environmental footprint types, regions, and industries that the entity operates so that different attribution methods are not selected merely to produce the most favorable results. While we do not settle the debate on which method is 'best', we do make it clear that future studies should assess the sensitivity of their key conclusions to their selection of attribution methodology.

Chapter 3

The hidden water and energy dependency of our digital society

3.1 Introduction

Data centers underpin our digital lives. Though relatively obscure just a couple of decades prior, data centers are now critical to nearly every business, university, and government, as well as those that rely on these organizations. Data centers utilize servers, digital storage equipment, and network infrastructure for the purpose of large-scale data processing and data storage[89]. Increasing demand for data creation, processing, and storage from existing and emerging technologies, such as online platforms/social media, video streaming, smart and connected infrastructure, autonomous vehicles, and artificial intelligence, has led to exponential growth in data center workloads and compute instances[48].

The global electricity demand of data centers was 205 TWh in 2018, which represents about 1% of total global electricity demand[30]. The cumulative electricity usage of all data centers is equivalent to the electricity usage of South Africa, the 22$^{\text{nd}}$ highest electricity consuming country in the world. The United States houses nearly 30% of data center servers, more than any other country[90,91,30]. Nearly 1.8% of US electricity consumption can be attributed to data centers[89]. Though the amount of data center computing instances has increased nearly 550% between 2010 and 2018, data center energy consumption has only risen by 6% due to dramatic improvements in energy efficiency and storage-drive density cross the industry[89,30]. However, it is unclear whether energy efficiency improvements can continue to offset the energy demand of data centers as the industry is expected to continue its rapid expansion over the next decade[89].

The growing energy demand of data centers has attracted the attention of researchers and policymakers not only due to scale of the industry's energy use but because the implications the industry's energy consumption has on green house gas (GHG) emissions and water use. Data centers directly and indirectly consume water and energy in their operation. Energy and water are interrelated, such that the use of one, implies the use of the other[92]. Most data centers' energy demands are supplied by the electricity grid, which distributes electricity from connected power plants. Power generation is the second largest water consumer[6] and the second largest emitter of GHGs in the US[93]. These environmental externalities can be attributed to the place of energy demand using several existing approaches[57,27].

In addition to the electricity consumed directly by data centers, it takes energy to supply and treat water used by data centers. Water is also used indirectly in the operation of data centers in the form of electricity utilization at the data center and through the electricity used in the treatment and distribution of water cycled through the data center. Like data centers, water and wastewater facilities are major electricity consumers, responsible for almost 1.8% of total electricity consumption in the US[94]. Beyond indirect water use, water is used directly within a data center to dissipate the immense amount of heat that is produced during its operation. For example, onsite water consumption to cool a 15 MW data center is around 300-550 thousand cubic meters annually[89].

Researchers have developed novel models and methods to estimate national and global energy use of data centers (e.g.,[95,96,30,89]). However, much less is known about the environmental footprint of data centers. The carbon footprint of a data center, expressed as equivalent CO_2, defines the global warming potential of a data center. A series of studies have focused on technological innovation and benchmarking for minimizing the carbon footprint of data centers based on case studies of select data centers[97,98]. Otherwise, research detailing the carbon footprint of data centers have been limited to national or global scale estimates using global average emission factors[99,100,90]. The geographic location[101,102] and the local electricity mix[103] are strong determinants of a data center's carbon footprint, though these spatial details are seldom considered in most studies. Systematic analyses of data centers' water footprint (i.e., consumptive water use) are even more scarce. A preliminary water footprint assessment of data centers by Ristic et al.[104] provided a range of water footprints associated with data center operation. Although Ristic et al. provided general estimates based on global average water intensity factors, their study highlights the importance of considering both direct and indirect water consumption associated with data center operation. Moreover, Ristic et al. highlights the importance of considering the type of power plants supplying electricity to a data center and the type/size of a data center, as each of these factors can significantly impact energy use and indirect water footprint estimates.

In this study we utilize spatially-detailed records of data center operations to provide the first sub-national estimates of data center water and carbon footprints. We focus on data center operations in the US since a plurality of the world's data centers are in the US and relevant data is more readily available. Figure 3.1 illustrates the scope of this study, which includes the operation of the data center, as well as the power plant(s), water supplier, and wastewater treatment plant servicing the data center. The environmental footprint of the non-operational stages of a data center's life cycle (e.g. manufacturing of servers) is negligible[99]. The spatial detail afforded by our approach enables us to provide more accurate estimates of water consumption and GHG emissions associated with data centers than previous studies, since both water and carbon footprints exhibit significant spatial variation[105] not captured by previous studies. Moreover, we evaluate the impact of data

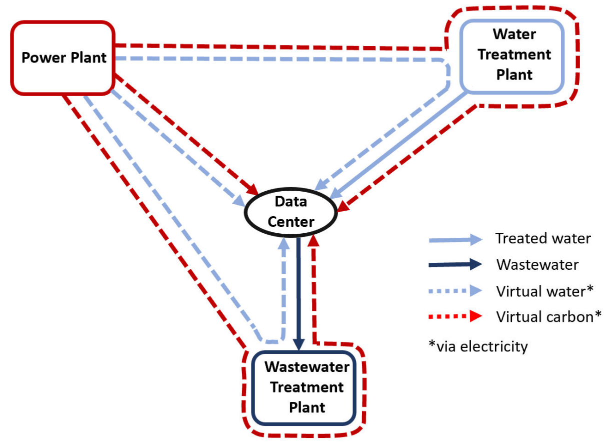


Figure 3.1: *The system boundaries and interlinkages defining the operational water and carbon footprints of data centers. Specific power plants, water utilities, and wastewater utilities are connected to each data center through their provisioning of electricity and water. Power plants emit GHGs and consume water in the production of electricity. These environmental impacts are attributed to data centers in proportion to how much electricity the data center uses (red and blue dashed lines connecting facilities). The GHG emissions and water consumption associated with the provisioning of treated water and disposal of wastewater, including the GHGs and water consumed in the generation of the electricity supplied to these facilities, are also attributed to data centers in proportion to their use of these utilities. Data centers do not directly emit GHGs but they do directly consume water to dissipate heat. All these facilities work together to keep data centers operational and contribute to the water and carbon footprint of data centers.*

centers on the local water balance and identify data centers located in already water stressed watersheds. In doing so, this study aims to answer the following questions: (i) What is the direct and indirect water footprint of US data centers (ii) Which watersheds support each data center's water demand and what portion of these watersheds are water stressed? (iii) How much GHG emissions are associated with the operation of data centers?

3.2 Methods

We utilize spatially detailed records on data centers, electricity generation, GHG emissions, and water consumption to determine the carbon footprint and water footprint of data centers in the US. Our approach connects specific power plants, water utilities, and wastewater treatment plants to each data center within the US. The unparalleled spatial detail provided by our methodology greatly improves previous estimates of water and carbon footprints of data centers since each footprint can vary greatly depending on the fuel source and cooling technology employed by the electricity provider.

3.2.1 Data

All data used in this study is for the year 2018, the most recent year where all data is available.

Data Centers

Information availability on data center location and size varies depending on its type and owner. For instance, there is little detailed public information on closet or small data centers since these are primarily used for organizations' internal computational and storage needs. Ganeshalingam et al.[90] performed a systematic investigation based on the Commercial Building Energy Consumption Survey (CBECS) and Commercial Building Stock Assessment (CBSA) report on commercial building stock for estimated locations of in-house small and midsize data centers. The survey results, which we utilize in our study, showed that approximately 40% of the installed servers in the US are located in small and midsize data centers. Alternatively, more information on colocation data centers typically exists since these data centers aim to make their services available to the public. We utilized detailed information on colocation and hyperscale data centers from commercial compilations[106,107,107] that get direct support and input from data center service providers to identify these type of data centers.

We have reclassified our collected data based on the International Data Corporation (IDC) classification of data centers (summarized in Table 3.1) and estimated the electricity use based on the available data on floor spaces of data centers. We used IT load intensity values (IT_S in watt/ft^2) for different data center types (s) as suggested by Shehabi et al.[108]

to estimate the total energy requirements (DC_E_{total}) of colocation and hyperscale data centers as follows:

$$DC\_E_{Total} = IT_s \times PUE_s \times A \tag{3.1}$$

where PUE_s is the power usage efficiency of space type s, and A is the floor area of data center in ft^2. Respondents may overstate data center capacity[90], and reported data lacks clear distinction between gross and raised floor area. Also, the colocation data centers provide rack spaces for rent that are not yet filled to the fullest capacity. We account for these data limitations by scaling our server counts to match the 2018 estimate of servers by data center type[30] and spatially distribute these servers in proportion to the the spatial distribution of installed server bases in existing records.

Power usage effectiveness (PUE) is a key metric of data center energy efficiency[109]. A value close to 1.0 is ideal as it indicates all energy consumed by a data center is used to power computing devices. Any energy used to power non-computing components, such as lighting and cooling, increases the PUE above 1.0 (see Equation 3.2). Generally, a data center's PUE is inversely proportionate to its size since larger data centers are better able to optimize their energy usage. Average PUE values by data center size were taken from Shehabi et al.[1] and shown in Table 3.1.

$$PUE = \frac{Total\ power\ supplied\ to\ the\ data\ center}{Power\ consumed\ by\ the\ IT\ equipment} \tag{3.2}$$

Electricity Generation, Water Consumption, and GHG Emissions

The US Energy Information Administration (EIA) publishes plant-specific, self-reported data for power plant electricity generation and water consumption[?]. Approximately 9,000 power plants are supplying electricity to the US electric grid, with 63% and 19% of the total electricity produced from 3,200 fossil fuel and 60 nuclear operated power plants, re-

Table 3.1: *Breakdown of electricity consumption by space type.*

Space type	Typical Size	IT Load	PUE
Closet	$< 100\ ft^2$		2
Room	$100 - 999\ ft^2$	$40\ W/ft^2$	2.35
Localized	$500 - 1,999\ ft^2$	$60\ W/ft^2$	1.88
Mid-tier	$2000 - 19,999\ ft^2$	$80\ W/ft^2$	1.79
High-end	$> 20,000\ ft^2$	$100\ W/ft^2$	1.6
Hyperscale	*	*	1.13

** Hyperscale data centers are not defined by their floor area but by their mode of operation. However, hyperscale data centers are typically among the largest data centers by floor area.*

spectively[110]. Though EIA reports of electricity generation have long been viewed as reliable, a 2014 change in data collection methodology significantly improved the quality of reported water consumption[51], which had previously been viewed as inconsistent[111]. The EIA only requires power plants with generation capacity greater than 100 MW (representing three-fourth of the total generation) report water consumption. We assigned national average values of water consumption per unit of electricity generation (i.e., water intensity; m^3/MWh) to all the power plants with generation capacity less than 100 MW according to their fuel type.

The operational water footprint of renewable energy sources can vary by several orders of magnitude. The operational water footprint of wind and solar are inconsequential compared to the other electricity generation types. Here, we use operational water footprints of solar and wind power from Macknick et al.[112]. The scale and variance of the water footprint of hydropower is significant. Estimates of water consumption associated with hydropower are not straightforward due to difficulties arising from attributing reservoir evaporation to different users of a multi-purpose reservoir. Following Grubert[113], we assign all reservoir evaporation to the dam's primary purpose (e.g., hydropower, irrigation, water supply, flood control). We connected hydroelectric dams with their respective power plants using data from Grubert[114]. Reservoir specific evaporation comes from Reitz et al.[115].

We use the U.S. Environmental Protection Agency's eGRID database[116] on GHG emissions associated with each power plant. GHG emissions are converted to an equivalent amount of carbon dioxide ($CO_2 - eq$) with the same global warming potential so to derive a single carbon footprint metric[117]. Direct GHG emission during the operation of data centers are negligible[99] and therefore not considered in this study.

Data centers, water suppliers, and wastewater treatment plants typically utilize electricity generated from a mix of power plants connected to the electricity grid. Within the electricity grid, electricity supply matches electricity demand by balancing electricity generation within and transferred into/out of a power control area (PCA). Though it is infeasible to trace an electron generated by a particular power plant to the final electricity consumer, there are several approaches to relate electricity generation to electricity consumption (Siddik et al.[105] summarizes the most common approaches). Here, we primarily rely on the approach used by Colett et al.[118] and Chini et al.[119] to identify the generative source of electricity supplied to any given data center. This approach assesses electricity generation and distribution at the same level which it is primarily managed, the PCA. PCA boundaries are derived from the Homeland Infrastructure Foundation level data (HIFLD)[120] and cross-checked against EIA-861 form[121], which identifies the PCAs operating in each state. Annual inter-PCA electricity transfers reported by the Federal Energy Regulatory Commission[122] are also represented within this approach. A data center (as well as water and wastewater utilities) draws on electricity produced within its PCA, unless the total demand of all energy consumers within the PCA exceeds local generation, in which case electricity imports from other PCAs are utilized. If a PCA's electricity production equals or exceeds the PCA's electricity demand, it is assumed all electricity imports pass through the PCA and are re-exported for utilization in other PCAs. Siddick et al.[105] notes that water and carbon footprints are sensitive to the attribution method used to connect power plants to energy consumers. Therefore, we conduct a sensitivity analysis (see the SI and results section for additional details) to test the degree to which our electricity attribution method affects our results. Additionally, we also test different assumptions regarding the water footprint of hydropower generation, as this too is a key source of uncertainty (again, consult the SI and results section for additional details).

We focus on the annual temporal resolution and assume an average electricity mix proportional to the relative annual generation of each contributing power plant. Though the electricity mix within a PCA can fluctuate hourly depending on balancing measures, these intra-annual variations will not significantly impact our annual-level results. While it is infeasible to determine the precise amount of electricity each power plant provides to each data center, water utility, and wastewater treatment plant, our approach will enable us to estimate where each facility is most likely to draw its electricity. The dependency of a data center on local and imported electricity from other PCAs was calculated using equation 3.3 and 3.4.

$$DC\_E_{p,l} = DC\_E_p \times (1 - \sum_i r_i) \tag{3.3}$$

$$DC\_E_{p,i} = DC\_E_p \times \sum_i r_i \tag{3.4}$$

where $DC\_E_{p,l}$ and $DC\_E_{p,i}$ is the local and imported electricity (MWh) to a data center from PCA p, respectively. DC_E_p is the total electricity consumption of the data center, whereas r_i represents the electricity contribution of each PCA to PCA p as follows:

$$r_i = \begin{cases} \frac{Import_{adj}}{Generation_p + \sum Import_p - \sum Export_p}, & \text{if PCA } p \text{ is net importer} \\ 0, & \text{if PCA } p \text{ is net exporter} \end{cases}$$

where $Import_{adj}$ is defined as the electricity from a linked PCA that was consumed while bypassing the node PCA p.

Adjusted electricity consumption from the PCAs were assigned to the power plants using equation 3.5.

$$DC\_E_{p,k} = DC\_E_{p,adj} \times \frac{PP_k}{\sum_{k=1}^{n} PP_k} \tag{3.5}$$

Where, $DC\_E_{p,k}$ is the total energy directly consumed [MWh/year] by data centers from power plant k that is attributed to PCA p, PP_k is the net generation by a specific power plant in MWh/year, and n is the number of power plants within the PCA p. A similar approach was taken to connect power plants to water and wastewater utilities, with their electricity usage (and associated environmental footprints) then linked to the data center they service.

3.2.2 Water Consumption of Data Centers

Indirect Water Consumption of Data Centers

The indirect water footprint of each data center consist of water consumption associated with the generation of i) electricity utilized during data center operation, ii) electricity used by water treatment plants for treatment and supply of cooling water to data centers, and iii) electricity used by wastewater treatment plants to treat the wastewater generated by a data center. Water consumed by power plants in generating electricity directly supplied to a data center is attributed as a portion of the indirect water footprint of a data center operation as follows:

$$DC\_IW_k = DC\_E_k \times BWF_k \tag{3.6}$$

Where DC_IW_k is the indirect water footprint associated with electricity consumed during the operation of a data center from power plant k in m^3/y, BWF_k is the water consumption per unit generation of electricity for power plant k in m^3/MWh (i.e. blue water footprint per unit of production).

The EPA Safe Drinking Water Information System contains information on the location, system type, and source of water for each public water and wastewater utility [93]. We assume the nearest public water system and wastewater treatment plant services a data center's water demand and wastewater management, respectively. After calculating the water supply requirement of a data center (discussed later in this section), the electricity needed for treatment and distribution of cooling water can be calculated using the data from Pabi et al.[94]. Water and wastewater treatment plants were linked to power plants (as described in Section 2.1.2) to estimate the indirect water footprint associated with electricity required to distribute and treat water and wastewater used by a data center. We then sum the water consumed by each power plant to directly or indirectly service a data center to determine the total indirect water footprint of that data center. The indirect water footprint associated with each power plant was also aggregated within watershed boundaries to determine which water sources each data center was reliant upon.

Direct Water Consumption of Data Centers

Water consumed for dissipating heat is the only source of data center direct water consumption. Data centers generate significant amounts of heat in their operation, which if not removed will compromise its continued operation. The ambient temperature inside a data center needs to be maintained between $15-32°C$ and a relative humidity needs to be between 20-80% to decrease the risk of IT equipment failure[109]. For medium and larger data centers, conventional air conditioning system are unable to meet these conditions. A modern data center cooling system dissipates heat in the range of $40-80\,W/ft^2$, almost 10 times the capacity of a conventional HVAC system $(4-8\,W/ft^2)$[123]. Most data centers utilize water cooled chillers to maintain temperatures within the prescribed range[124]. Most large-scale data centers use airside or waterside economizers (depending on the geographic location) to assist in cooling. Economizers significantly decreases the required water used by the cooling tower, as well as facility energy use (note the lower PUE values for larger data centers in Table 3.1). Use of a dry cooler in place of a cooling tower is uncommon but can minimize or even eliminate direct water use; however, energy consumption increases significantly.

Direct water consumption of a data center can be estimated from the heat generation capacity of a data center[125], which is related to amount of electricity used[126]. Records of data center specific cooling system are not available. Therefore, the average water use rate of a water-cooled chiller cooling tower (the most common type of cooling system) was used (approximately 1.8 m^3 of water per MWh[89]). Together, we use this information to estimate the direct water footprint of each data center as follows:

$$DC\_DW = DC\_E \times 1.8 \tag{3.7}$$

where DC_DW is the direct water consumption (m^3) of a data center. The direct water consumption is assigned to the watershed where the water utility supplying the data center withdraws its water. While the data center may be located within the same watershed as the water utility, this is not always the case.

3.2.3 Water Scarcity Footprint

While the water footprint metric assesses the direct and indirect volumetric consumptive water use of a data center, the water scarcity footprint (WSF) puts the water use of a data center in the context of local water availability. The water scarcity footprint (as defined by ISO 14046 and Boulay et al.[127]) indicates the pressure exerted by consumptive water use on available freshwater within a river basin and determines the potential to deprive other societal and environmental water users from meeting their water demands. We quantified the WSF of data centers using the AWARE method set forth by Boulay et al.[127]. Other societal and environmental water use data come from Marston et al.[77] and Richter et al.[128], whereas data on natural water availability within each US watershed come from Richter et al.[128].

In effect, WSF can be interpreted as the volume of water used by a data center after considering the relative abundance of unused water in a subbasin compared to all other subbasins across the US. The WSF for each subbasin i (Hydrologic Unit Code 8) in the conterminous US was calculated as

$$WSF_i = WC_i \times CF_i \tag{3.8}$$

where WC_i is the water consumption within subbasin i attributed to data centers direct and indirect water requirements. The characterization factor, CF_i, normalizes water consumption within subbasin i against the national average monthly water availability after all demands are satisfied per unit land area. The characterization factor (CF) is bounded between 0.1 - 100. A subbasin with a CF equal to 1 indicates that it has the same amount of unused water per contributing area over a certain time period as the national average. Subbasins with CF less than 1 exhibit less water scarcity than average, whereas subbasins with CF greater than 1 experience greater levels of water scarcity. The following equation was used to calculate CF.

$$CF_i = \frac{AMD_{USavg}}{AMD_i} \tag{3.9}$$

AMD_i is the difference between water availability per area and water demand per area in subbasin i, as shown in Equation 3.10. AMD_USavg represents the national average AMD value, which is used to normalize the AMD of each subbasin.

$$AMD_i = \frac{WA_i - HWC_i - EWR_i}{A_i} \tag{3.10}$$

where WA_i is the water availability defined as the summation of streamflow and available storage within subbasin i, HWC_i and EWR_i are human water consumption and the environmental water requirements, respectively, in subbasin i, and A_i is the area of subbasin i.

Similar approach was used to compare environmental impact of new data center installation across the US. We assumed a hypothetical 1 MW identical data center across the US subbasins. Each hypothetical data center was connected to PCAs, and subsequent subbaisns for their dependency of water supply using the approach discussed above. Water supply portfolio data centers across the US subbasins can be found in [$Subbasin\_DC$] tab of the XLS SI. Volumetric WSF also can serve as comparative suitability matrix of selected locations for installing new server bases. We have placed an identical 1 MW data center in all the US subbasin, traced it's quantitative WSF on all the subbasins. WULCA proposed quantification of WSF with relative available water remaining (AWARE) per area after satisfying the anthropogenic and environmental requirements. Similar to their global assumption, current state of US ecosystems served the benchmark of fair condition. This method compares individual subbasins with the overall US condition.

3.2.4 GHG Emissions Associated With Data Centers

Data centers rely on energy directly for their operation and indirectly for the treatment and distribution of water and wastewater associated with their operations. Depending on the energy source and technology employed by each power plant, GHGs are emitted into the atmosphere at different quantities. We were able to represent the heterogeneity in GHG emissions and attribute the GHGs emitted (represented as $CO_2 - eq$) by each power plant contributing electricity to the data center by using our aforementioned approach to link data centers, as well as water and wastewater utilities, to individual power plants. Emissions attributed to each data center were estimated using a similar approach used for calculating indirect water consumption of data centers.

$$DC\_GHG_k = DC\_E_k \times GHG_k \tag{3.11}$$

Where DC_GHG_k is $CO_2 - eq$ emissions (tonnes) from power plant k associated with electricity consumed during the operation of a data center, DC_E_k is the total energy consumed [MWh/y] by a data center from power plant k (from equation 3.5), and GHG_k is the $CO_2 - eq$ per unit of generated electricity for power plant k in tonnes/MWh.

3.3 Results

This study divides the results into groups to present water consumption, and GHG emissions. Each section identifies the direct and indirect portion of contribution to partition the consumption and emission of the operation stages alongside the spatial distribution.

3.3.1 Water footprint of data centers

Water transmitted from power plants to data centers in form of electricity is estimated to be 383 x 10^6 m^3/y. In medium to larger size data centers that employ cooling tower-based chillers to improve energy efficiency, water is consumed at the data center site itself. Electricity required to supply water from public utilities and treat the generated wastewater at wastewater treatment plants initiate a virtual water flow of approximately 450 x 10^3 m^3/y. Cooling towers use water evaporation to reject heat from the data center causing losses approximately equal to the latent heat of vaporization for water, along with some additional losses for drift and blowdown except for closet and room data centers, which are assumed to use direct expansion (air-cooled chillers). On-site water consumption is estimated at 130 million m^3 for all data centers. Adding up all these, the total water consumed for the data centers operation stage in the US for our collected data is about 513 million $m^3/year$. Our attribution of nearest community public water system as cooling water source introduces inter boundary physical water transfer. Approximately 17 x 10^6 m^3 of water directly consumed for cooling of data centers are sourced at a different subbasin than the location of the installed server bases. Majority of the indirect water is contributed by the eastern subbasins (figure 3.2(B)), while a significant portion of the servers are located in the southern and southwestern US. Also, balancing authorities in the eastern US are net exporter of electricity in most of the cases. A net flow of virtual water in form electricity, therefore, occurs from eastern to western US for supporting the data center activity. Water footprint trend of the existing data centers requires understanding of three scenarios; i) how much electricity is consumed by the data centers located within a region, ii) how much electricity is supplied by power plants within a region to data centers in all regions, and iii) how much water is consumed from the watersheds within a region. Indirect water contribution from the southwest subbasins are high as hydropower reservoirs located located along the arid southwest belt evaporate highest amount water per area compared to any other region. Although electricity consumption by the data centers located in west and southwest region are little less than 20%, and only 17% of the overall data centers electricity is supplied from the power plants within this region, embedded water consumption of electricity from this region is above 30% of the total indirect water footprint of data centers. On other extreme, 20% of the total indirect water consumption occurs from the subbasins of southeastern region, oppose to the 25% of total electricity consumption by the server bases hosted in this regions.

Water consumption for 1 MWh energy consumption of a data center from our study (7.1 m^3/MWh) differs from the approximate value used by Shehabi et al.[89]. They used an average value 7.6 m^3/MWh for indirect water consumption as estimated by the National Renewable Energy Laboratory[112]. Use of power plant specific water consumption, partitioning of water footprint of hydropower, and thermal heat output to estimate direct water requirement for cooling are the main reasons of discrepancy of our study from the available literature. However, the estimated water consumption may vary slightly depending on the shares of different electricity sources, especially renewables, from the total energy production, which are hard to determine due to the uncertainties that exist with regard to the complex movement of electricity through the electric grid[129]. Water consumption was the secondary concern of the earlier studies, while main focus being the direct energy consumption. Also, previous studies did not consider the energy required for treatment of supplied cooling water and treatment of generated wastewater.

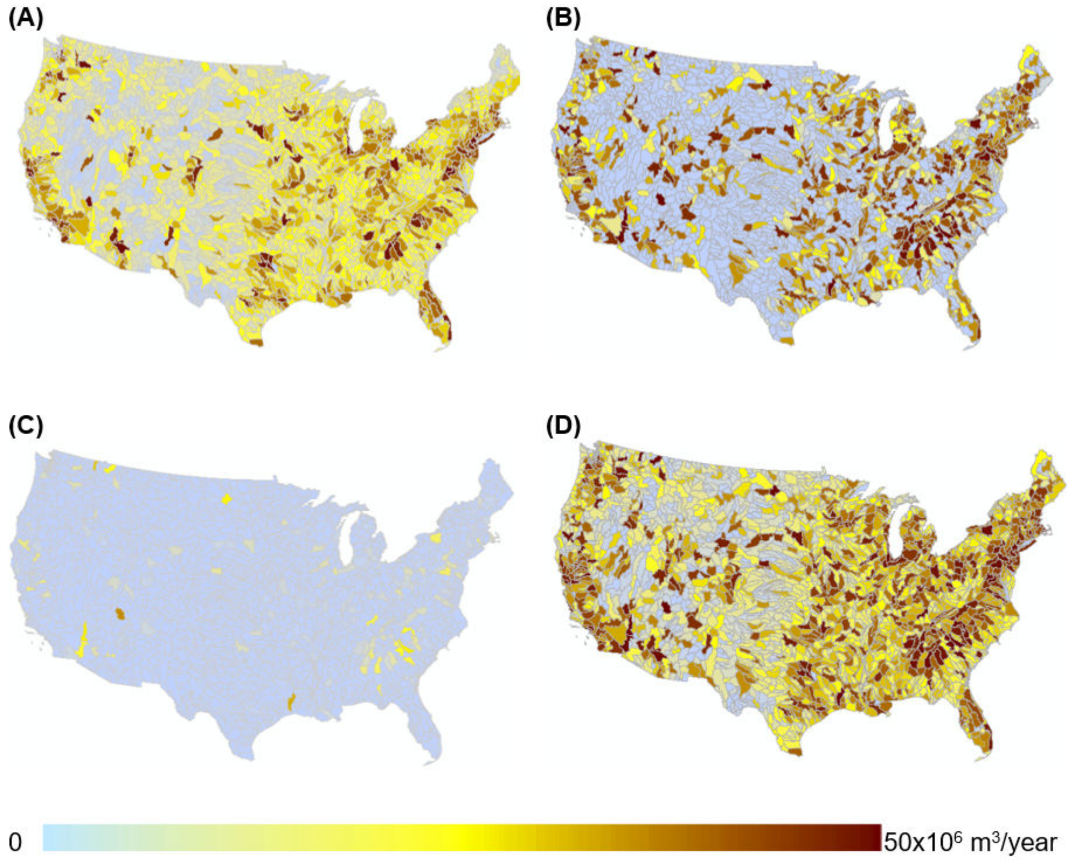


Figure 3.2: *Blue water footprint* ($m^3/year$) *of US data centers, resolved to each subbasin (HUC-8). Direct water footprint of data centers (A), indirect water footprints associated with direct electricity consumption of data center equipment (B), indirect water footprints associated with treatment of supplied cooling water and treatment of generated wastewater (C), and total water footprints of data centers operation stage (D).*

3.3.2 Water scarcity footprints

To holistically assess the impact of data centers on freshwater vulnerability, we need to highlight the results of this study from the perspective of interactions of data centers with power plants, and water and wastewater utilities. It is important to visualize the flow of water from the distant interacting sectors to data centers, instead of showing the results as a whole, or at the locations of data centers. The AWARE method used in this study measures the water deprivation strength of a consumer based on AMD (availability minus demand) factors of a watershed. Data centers account for 1.59 x 10^9 m^3/year of US equivalent water consumption from the subbasins. This resulted in a water scarcity footprint of 22 m^3 US eq water/MWh. This value represents combined impact of water intensity of the electricity consumed and existing water condition data center locations. But this total and average values vastly misleads the actual water stress distribution of the data centers. Almost all the watersheds in eastern US has a CF value less than one, suggesting per unit of water use has far less depriving capacity compared to overall US water condition. On the contrary, many of the watersheds in the western US have CF values much higher than 1, some with the limiting value of hundred. Water users from these watersheds have much higher depriving potential. The relative distribution of our CF values follows a similar trend as the county based CF values suggested by Lee et al. [130]. One primary difference is that we have considered return flow to better represent the existing condition of the US watersheds. Although the west and southwestern watersheds supply only 20% and 30% of direct and indirect water supply to data centers, 70% of overall water scarcity footprint are exerted in these subbasins (figure 3.3 (A)).

3.3.3 Emission of GHG from data centers

No direct emission is associated with data center operation, and electricity consumed by data centers is the only source of emission. Total emission of GHG was found to be 31.5 million tons per year. The result indicates that data centers are responsible for almost 0.6%

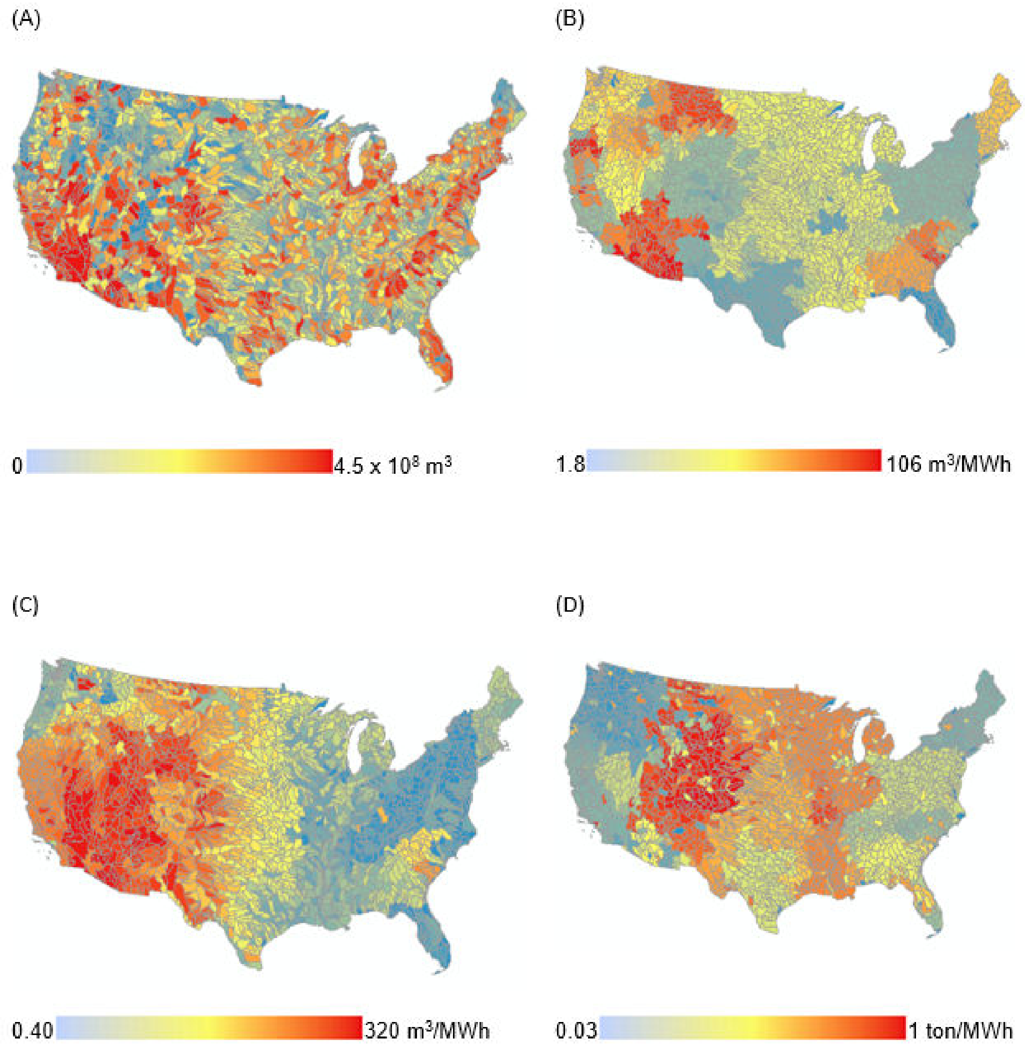


Figure 3.3: *Current water scarcity footprint of the existing data centers (A), WC footprint (m^3/MWh) for installation of a 1 MW data center across the US (B), WSF footprint (m^3-eq/MWh) for installation of a 1 MW data center across the US, and (D) Emission intensity ($CO_2 - eq/MWh$) for installation of a 1 MW data center across the 2110 watersheds in the continental United States.*

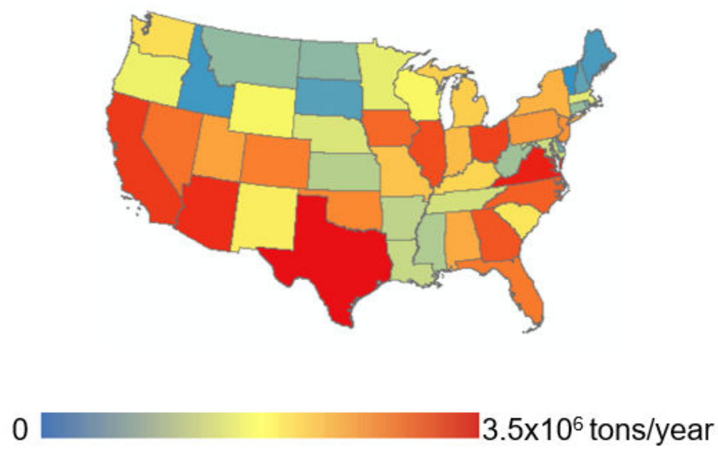


Figure 3.4: *Maps of Greenhouse gas emissions (Metric tons of $CO2-eq/y$) from data centers operation in the US from data centers' operation at the State level spatial resolution.*

of total $CO2-eq$ emissions in the US. Thermometric power plants generate almost all the emissions associated with electricity generation. As the eastern US has high concentration of thermoelectric power plants, along with large number of data centers, almost half of the total emission of data centers operation are attributed to these regions. Weightage in data centers power consumption may be significantly different for a region compared to GHG emission from that region as the energy may be sourced to a distant power plant or low carbon intensive power plant or both. Interesting feature was observed at the Ohio Valley where almost 30% of the emissions occur oppose to only 10% power consumption and 9% water consumption associated with data centers operation. One reason is the high density of thermoelectric power plants in this region with emission intensity of 0.65 ton/MWh. But when we discuss from electricity supply perspective, power plants in Ohio valley constitute 20% of total electricity consumed by overall data centers. A large portion of this electricity supports the data centers located in northeast and southeast region, where almost one-third of total servers are located. Whereas, southeast region, where 25% of the electricity consumption occurs, is attributed to little over 10% emission from the power plants within this region.

3.3.4 Where to locate data centers centers to minimize water and carbon footprints

This study enlightens on the choice of location for environmental impact mitigation as a supplementary to the constant effort on energy efficient technological innovations. Quantitative water consumption by the hypothetical data centers placed accoross the US subbasins ranges from 1.8 - 106 m^3/MWh. Relative water scarcity should also be considered, along with quantitative water use for decision making. WSF can potentially alter the selection of location for management and purchasing decision of consumers. Comparative suitability analysis shows that the WSF of a 1 MW data center can vary from 0.5 to 305 m^3-eq/MWh depending on the choice of location. Emission potential of a data centers spatial profile can range from 0.02-1 ton/MWh of electricity consumption by data center equipment. Data center located at a watershed draws water from distant subbasins as virtual water embedded electricity use. Characterization factors (CF) provides a mean to quantify the relative water availability at a watershed. Although watersheds with low CF feels like an impromptus choice in terms of water use, it may not be always the best choice. Data centers located at low CF watersheds ensure a lower scarcity footprint of the direct water use. But indirect water footprint extends far beyond the boundary of the local watersheds for supply of electricity. Around 10% of the hypothetical data centers placed in the watersheds within 25^{th} percentile CF values showed a total WSF above 75^{th} percentile. Low WSF of direct water consumption was diminished by indirect water flow from power plants that are highly water intensive or located in watersheds with high CF or both. On the other hand, almost 20% of watersheds that proved to be within the 25th percentile in term of total water consumption for a hypothetical data center, resulted in 75th percentile or above US equivalent water consumption. Finally, less 5% of the hypothetical data centers had both their emission and WSF within the first quantile. This imposes barrier on simplistic decision making as policy makers have to consider both CF of all the inteconnected watersheds and potential tradeoff between the environmental factors. Optimal location of data center based on an environ-

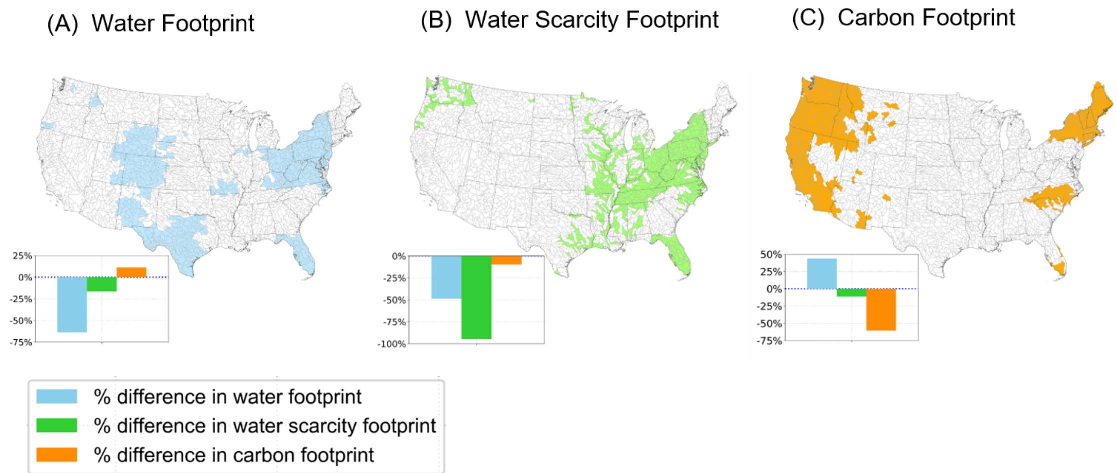


Figure 3.5: *Choice of optimum location for new data center is complicated due to dependency on multiple watersheds for environmental resources. If 25th percentile is set as a benchmark based on an environmental factor for all future installation of hypothetical data centers, instead of uniform distribution over the US, potential tradeoffs between environmental resources arise.*

mental stress often do not spatially overlap with the other stressing factors (figure 3.3). Western US seems particularly water sensitive to introduction of new installation of server bases, whereas mid western regions show higher emission intensity. As the data centers are almost saturated with technological innovations, choice of location provides new scope for minimizing environmental footprint of future data centers. Existing condition of watersheds in lower northeast and upper central valley regions make those preferable sites in terms of water footprint and water stress. Almost opposite trend is seen when the choice of location is made based on minimizing carbon emission that identifies west and northwest regions as potential location for future data centers (figure 3.5. This finding imply that a tradeoffs between water and carbon needed to be in consideration for the policy makers.

3.3.5 Sensitivity analysis

Indirect water consumption in form of electricity dominates in the total water consumption for data centers operation. So, electricity portfolio of the selected attribution boundary plays a vital role in the estimated water footprint of data centers operation. Considering primary purpose water allocation of hydroelectricity, water intensity for data centers operation ranges from 6.0-7.9 m^3/MWh. Although hydroelectricity constitutes only 7% of the total generation from electric grid, it is often the driving force in water consumption estimation due its very high-water intensity compared to other methods of electricity generation. Based on the assumption of water allocation to the hydroelectricity, water consumption estimation varies from 3.25 m^3/MWh to 12 m^3/MWh, where the lower extreme assumes no water consumption for hydroelectricity and higher extreme assumes all the water evaporated from hydroelectric reservoirs as consumptive water use for hydroelectricity. Annual estimation of water use seems to be more sensitive to attribution method compared to the carbon emission (figure 3.6), which is contrary to the findings of Siddik et al.[105]. But local analysis shows that almost 350 watersheds have coefficient of variation (CV) for carbon emission greater than 1, whereas less than 250 watersheds have CV values of water use greater than 1. Average CV of the watersheds for carbon emission always exceed the average CV of water consumption for all the regions. Watersheds located in the southwest region have highest level of sensitivity (CV = 0.61) to attribution method in terms of water use, and watersheds within the northeast region are highly sensitive to the carbon emission values (CV = 0.80) for attributing data centers to potential energy sources.

3.4 Discussion

We are aware of the two drawbacks of our study. First, there is no consensus in assigning the generative source of consumed electricity. Our model is based on the concept of attributing carbon cost to the consumer from their local supplier. Second, further data availability of the installed server bases would significantly improve our assessment of local environmental impacts. Despite the limitations, this study has important implications for both stakeholders and policymakers. The novelty of this study is analyzing the spatial distribution of water

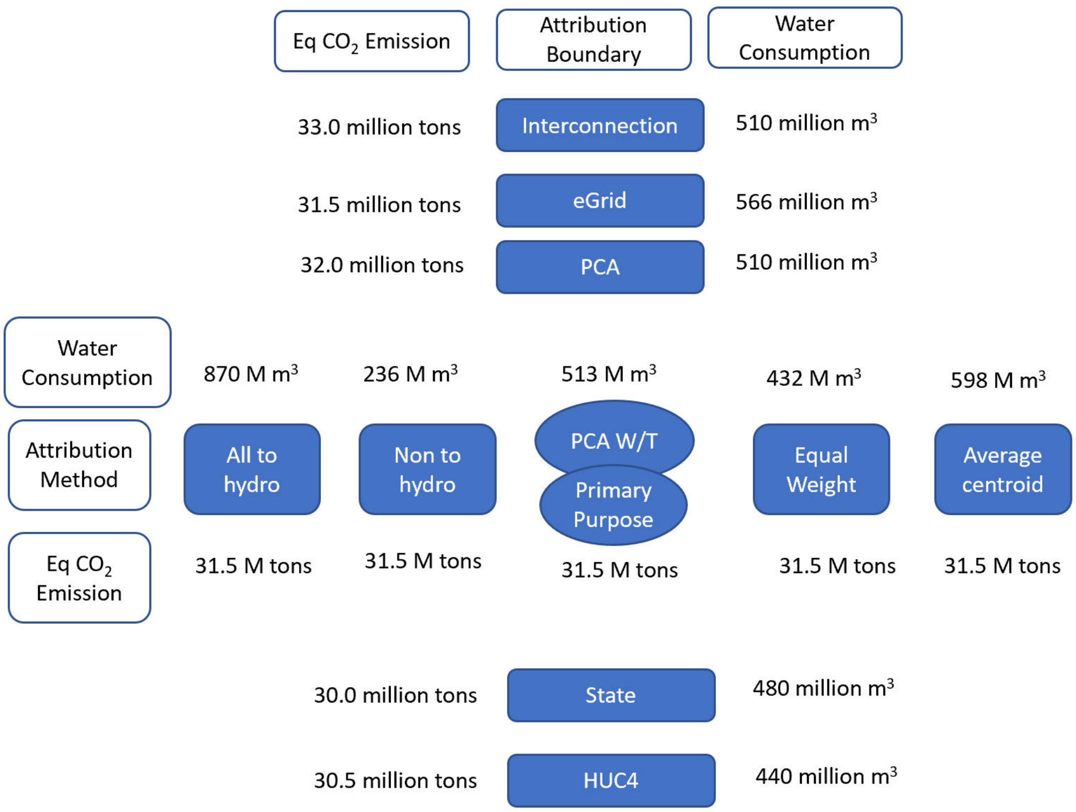


Figure 3.6: *Variability in consumptive water use estimation based on different assumptions.*

and carbon footprint of data centers within the US. Also, it quantifies the local impact of data centers operation as water scarcity footprint. Derived outcomes from water consumption and GHG emission pattern of data centers can be interpreted as: (i) major share of water footprint occurs in form of indirect water consumption from electricity generation, (ii) coincident of water consumption with the water stressed sub basins should be put into considerations by the policy makers, (iii) Choice of location can dramatically change the WSF and carbon intensity of a proposed server base installation. Spatial variation of water and carbon footprint, and existing stress level of the sub basins will enable the stakeholders choosing appropriate location for future installation of data centers. We believe that the water scarcity footprint assessment shown in this study can serve a deciding factor for stakeholders in choice of location for new data centers. This approach enables comparing the water scarcity profile of a data center as a function of watershed and electricity provider. Virtual service providers connect consumers far from the location of water consumption and GHG emission. A transparent data availability on the energy and water source will play a vital role in decision making of the consumers to connect with a cloud service provider.

Bibliography

[1] Arman Shehabi, Sarah J Smith, Eric Masanet, and Jonathan Koomey. Data center growth in the united states: decoupling the demand for services from electricity use. *Environmental Research Letters*, 13(12):124030, 2018.

[2] Ashlynn S Stillwell, Carey W King, Michael E Webber, Ian J Duncan, and Amy Hardberger. The energy-water nexus in Texas. *Ecology and Society*, 16(1), 2011.

[3] Christopher A Scott, Suzanne A Pierce, Martin J Pasqualetti, Alice L Jones, Burrell E Montz, and Joseph H Hoover. Policy and institutional dimensions of the water–energy nexus. *Energy Policy*, 39(10):6622–6630, 2011.

[4] Lu Liu, Mohamad Hejazi, Hongyi Li, Barton Forman, and Xiao Zhang. Vulnerability of us thermoelectric power generation to climate change when incorporating state-level environmental regulations. *NATURE*, 2(17109):1, 2017.

[5] PA Behrens, Michelle TH van Vliet, T Walsh B Nanninga, Rodriques JF Dias, et al. Climate change and the vulnerability of electricity generation to water stress in the european union. *Nature*, 2:17114, 2017.

[6] total water use in the united states. URL https://www.usgs.gov/special-topic/water-science-school/science/total-water-use-united-states?qt-science_center_objects=0#qt-science_center_objects.

[7] Ashlynn S Stillwell, Mary E Clayton, and Michael E Webber. Technical analysis of a river basin-based model of advanced power plant cooling technologies for mitigating water management challenges. *Environmental Research Letters*, 6(3):034015, 2011. ISSN 1748-9326.

[8] Juliet Christian-Smith, Peter H Gleick, Heather Cooley, Lucy Allen, Amy Vanderwarker, and Kate A Berry. *A twenty-first century US water policy*. Oxford University Press, 2012.

[9] Naresh Devineni, Upmanu Lall, Elius Etienne, Daniel Shi, and Chen Xi. America's water risk: Current demand and climate variability. *Geophysical Research Letters*, 42 (7):2285–2293, 2015.

[10] Arjen Y Hoekstra, Ashok K Chapagain, and Guoping Zhang. Water footprints and sustainable water allocation, 2016.

[11] K R Gurney, J Huang, and K Coltin. Bias present in us federal agency power plant co2emissions data and implications for the us clean power plan. *Environmental Research Letters*, 11(6):064005, Jan 2016. doi: 10.1088/1748-9326/11/6/064005.

[12] Nana Yaw Amponsah, Mads Troldborg, Bethany Kington, Inge Aalders, and Rupert Lloyd Hough. Greenhouse gas emissions from renewable energy sources: A review of lifecycle considerations. *Renewable and Sustainable Energy Reviews*, 39:461–475, 2014.

[13] Richard Dobbs, Sven Smit, Jaana Remes, James Manyika, Charles Roxburgh, and Alejandra Restrepo. *Urban world: Mapping the economic power of cities*. McKinsey Global Institute, 2011.

[14] Konrad Otto-Zimmermann. From rio to rio+ 20: The changing role of local governments in the context of current global governance. *Local Environment*, 17(5):511–516, 2012.

[15] Lucas A Djehdian, Christopher M Chini, Landon Marston, Megan Konar, and Ashlynn S Stillwell. Exposure of urban food–energy–water (few) systems to water scarcity. *Sustainable Cities and Society*, 50:101621, 2019.

[16] Christopher L Weber, Paulina Jaramillo, Joe Marriott, and Constantine Samaras. Life cycle assessment and grid electricity: What do we know and what can we know? *Environmental Science & Technology*, 44(6):1895–1901, 2010.

[17] Willa Paterson, Richard Rushforth, Benjamin L Ruddell, Megan Konar, Ikechukwu C Ahams, Jorge Gironás, Ana Mijic, and Alfonso Mejia. Water footprint of cities: A review and suggestions for future research. *Sustainability*, 7(7):8461–8490, 2015.

[18] Benjamin L Ruddell, Elizabeth A Adams, Richard Rushforth, and Vincent C Tidwell. Embedded resource accounting for coupled natural-human systems: An application to water resource impacts of the western us electrical energy trade. *Water Resources Research*, 50(10):7957–7972, 2014.

[19] Elliot Cohen and Anu Ramaswami. The water withdrawal footprint of energy supply to cities. *Journal of Industrial Ecology*, 18(1):26–39, 2014.

[20] Evan Kodra, Seth Sheldon, Ryan Dolen, and Ory Zik. The North American electric grid as an exchange network: An approach for evaluating energy resource composition and greenhouse gas mitigation. *Environmental Science & Technology*, 49(22):13692–13698, 2015.

[21] Christopher M Chini, Lucas A Djehdian, William N Lubega, and Ashlynn S Stillwell. Virtual water transfers of the us electric grid. *Nature Energy*, 3(12):1115, 2018.

[22] Rebecca Allyson Marie Peer, Emily Grubert, and Kelly T Sanders. A regional assessment of the water embedded in the us electricity system. *Environmental Research Letters*, 2019.

[23] Scott Kelley and Martin Pasqualetti. Virtual water from a vanishing river. *Journal-American Water Works Association*, 105(9):E471–E479, 2013.

[24] Vincent C Tidwell, Michael Bailey, Katie M Zemlick, and Barbara D Moreland. Water supply as a constraint on transmission expansion planning in the western interconnection. *Environmental Research Letters*, 11(12):124001, 2016.

[25] Christopher M Chini, Kelsey L Schreiber, Zachary A Barker, and Ashlynn A Stillwell. Quantifying energy and water savings in the us residential sector. *Environmental Science & Technology*, 50(17):9003–9012, 2016.

[26] Matthew D Bartos and Mikhail V Chester. The conservation nexus: Valuing interdependent water and energy savings in arizona. *Environmental Science & Technology*, 48(4):2139–2149, 2014. ISSN 0013-936X.

[27] Christopher M. Chini, Megan Konar, and Ashlynn S. Stillwell. Direct and indirect urban water footprints of the united states. *Water Resources Research*, 53(1):316–327, 2017. ISSN 1944-7973. doi: 10.1002/2016WR019473.

[28] T. A. DeNooyer, J. M. Peschel, Z. Zhang, and A. S. Stillwell. Integrating water resources and power generation: the energy-water nexus in Illinois. *Applied Energy*, 162:363–371, 2016.

[29] Emily A Grubert and Michael E Webber. Energy for water and water for energy on maui island, hawaii. *Environmental Research Letters*, 10(6):064009, 2015.

[30] Eric Masanet, Arman Shehabi, Nuoa Lei, Sarah Smith, and Jonathan Koomey. Recalibrating global data center energy-use estimates. *Science*, 367(6481):984–986, 2020.

[31] Jonathan G Koomey. Worldwide electricity used in data centers. *Environmental research letters*, 3(3):034008, 2008.

[32] James Hamilton. Cooperative expendable micro-slice servers (cems): low cost, low power servers for internet-scale services. In *Conference on Innovative Data Systems Research (CIDR'09)(January 2009)*. Citeseer, 2009.

[33] Neil Rasmussen. Strategies for deploying blade servers in existing data centers. *White Paper*, 125:1–14, 2006.

[34] Jyothi Sekhar, Getzi Jeba, and Sai Durga. A survey on energy efficient server consolidation through vm live migration. 2012.

[35] Xiaoqiao Meng, Vasileios Pappas, and Li Zhang. Improving the scalability of data center networks with traffic-aware virtual machine placement. In *2010 Proceedings IEEE INFOCOM*, pages 1–9. IEEE, 2010.

[36] Mueen Uddin and Azizah Abdul Rahman. Server consolidation: An approach to make data centers energy efficient and green. *arXiv preprint arXiv:1010.5037*, 2010.

[37] Mohammad Shojafar, Nicola Cordeschi, and Enzo Baccarelli. Energy-efficient adaptive resource management for real-time vehicular cloud services. *IEEE Transactions on Cloud computing*, 7(1):196–209, 2016.

[38] Seung-Min Han, Mohammad Mehedi Hassan, Chang-Woo Yoon, and Eui-Nam Huh. Efficient service recommendation system for cloud computing market. In *Proceedings of the 2nd international conference on interaction sciences: information technology, culture and human*, pages 839–845, 2009.

[39] Junaid Shuja, Kashif Bilal, Sajjad A Madani, Mazliza Othman, Rajiv Ranjan, Pavan Balaji, and Samee U Khan. Survey of techniques and architectures for designing energy-efficient data centers. *IEEE Systems Journal*, 10(2):507–519, 2014.

[40] Michele Mazzucco and Dmytro Dyachuk. Optimizing cloud providers revenues via energy efficient server allocation. *Sustainable Computing: Informatics and Systems*, 2 (1):1–12, 2012.

[41] Michael Tresh, Brian Jackson, Edward Bednarcik, John Prunier, Martin Olsen, and

Mark Germagian. Systems and methods for closed loop heat containment with cold aisle isolation for data center cooling, July 1 2014. US Patent 8,764,528.

[42] John Niemann, Kevin Brown, and Victor Avelar. Impact of hot and cold aisle containment on data center temperature and efficiency. *Schneider Electric Data Center Science Center, White Paper*, 135:1–14, 2011.

[43] Steve Greenberg, Evan Mills, Bill Tschudi, Peter Rumsey, and Bruce Myatt. Best practices for data centers: Lessons learned from benchmarking 22 data centers. *Proceedings of the ACEEE Summer Study on Energy Efficiency in Buildings in Asilomar, CA. ACEEE, August*, 3:76–87, 2006.

[44] MM Ohadi, SV Dessiatoun, K Choo, M Pecht, and John V Lawler. A comparison analysis of air, liquid, and two-phase cooling of data centers. In *2012 28th Annual IEEE Semiconductor Thermal Measurement and Management Symposium (SEMI-THERM)*, pages 58–63. IEEE, 2012.

[45] Levi A Campbell, Richard C Chu, Michael J Ellsworth Jr, Madhusudan K Iyengar, and Robert E Simons. Open flow cold plate for liquid cooled electronic packages, March 29 2011. US Patent 7,916,483.

[46] Hainan Zhang, Shuangquan Shao, Hongbo Xu, Huiming Zou, and Changqing Tian. Free cooling of data centers: A review. *Renewable and Sustainable Energy Reviews*, 35:171–182, 2014.

[47] Hafiz M Daraghmeh and Chi-Chuan Wang. A review of current status of free cooling in datacenters. *Applied Thermal Engineering*, 114:1224–1239, 2017.

[48] CVN Index. Cisco global cloud index: Forecast and methodology, 2016–2021 white paper.

[49] World energy outlook, 2017. URL https://www.iea.org/Textbase/npsum/weo2017SUM.pdf.

[50] K Averyt, J Macknick, J Rogers, N Madden, J Fisher, J Meldrum, and R Newmark. Water use for electricity in the united states: an analysis of reported and calculated water use information for 2008. *Environmental Research Letters*, 8(1):015001, Sep 2013. doi: 10.1088/1748-9326/8/1/015001.

[51] Rebecca AM Peer and Kelly T Sanders. Characterizing cooling water source and usage patterns across us thermoelectric power plants: A comprehensive assessment of self-reported cooling water data. *Environmental Research Letters*, 11(12):124030, 2016.

[52] Global greenhouse gas emissions data (us environmental protection agency, 2019). URL https://www.epa.gov/ghgemissions/global-greenhouse-gas-emissions-data.

[53] A. Y. Hoekstra and M. M. Mekonnen. The water footprint of humanity. *Proceedings of the National Academy of Sciences*, 109(9):3232–3237, 2012. doi: 10.1073/pnas.1109936109.

[54] Mesfin M. Mekonnen, P. W. Gerbens-Leenes, and Arjen Y. Hoekstra. The consumptive water footprint of electricity and heat: a global assessment. *Environmental Science: Water Research Technology*, 1(3):285–297, 2015. doi: 10.1039/c5ew00026b.

[55] Arjen Y Hoekstra and Thomas O Wiedmann. Humanity's unsustainable environmental footprint. *Science*, 344(6188):1114–1117, 2014.

[56] Davy Vanham, Adrian Leip, Alessandro Galli, Thomas Kastner, Martin Bruckner, Aimable Uwizeye, Kimo Van Dijk, Ertug Ercin, Carole Dalin, Miguel Brandão, and et al. Environmental footprint family to address local to planetary sustainability

and deliver on the sdgs. *Science of The Total Environment*, 693:133642, 2019. doi: 10.1016/j.scitotenv.2019.133642.

[57] Joseph S. Colett, Jarod C. Kelly, and Gregory A. Keoleian. Using nested average electricity allocation protocols to characterize electrical grids in life cycle assessment. *Journal of Industrial Ecology*, 20(1):29–41, Aug 2015. doi: 10.1111/jiec.12268.

[58] Joe Marriott and H. Scott Matthews. Environmental effects of interstate power trading on electricity consumption mixes. *Environmental Science Technology*, 39(22): 8584–8590, 2005. doi: 10.1021/es0506859.

[59] state-level energy-related carbon dioxide emissions, 2005-2016 (energy information administration, 2019). URL https://www.eia.gov/environment/emissions/state/analysis/.

[60] Emily Grubert and Kelly T. Sanders. Water use in the united states energy system: A national assessment and unit process inventory of water consumption and withdrawals. *Environmental Science Technology*, 52(11):6695–6703, Aug 2018. doi: 10.1021/acs.est.8b00139.

[61] Weiwei Mo, Ranran Wang, and Julie B. Zimmerman. Energy–water nexus analysis of enhanced water supply scenarios: A regional comparison of tampa bay, florida, and san diego, california. *Environmental Science Technology*, 48(10):5883–5891, 2014. doi: 10.1021/es405648x.

[62] Xu Tian, Rui Wu, Yong Geng, Raimund Bleischwitz, and Yihui Chen. Environmental and resources footprints between china and eu countries. *Journal of Cleaner Production*, 168:322–330, 2017. doi: 10.1016/j.jclepro.2017.09.009.

[63] Nicole A Ryan, Jeremiah X Johnson, and Gregory A Keoleian. Comparative assessment of models and methods to calculate grid electricity emissions. *Environmental Science & Technology*, 50(17):8937–8953, 2016.

[64] Joe Marriott and H Scott Matthews. Environmental effects of interstate power trading on electricity consumption mixes. *Environmental Science & Technology*, 39(22):8584, 2005.

[65] Sampo Soimakallio and Laura Saikku. Co2 emissions attributed to annual average electricity consumption in oecd (the organisation for economic co-operation and development) countries. *Energy*, 38(1):13–20, 2012.

[66] Christopher M Chini and Ashlynn S Stillwell. The changing virtual water trade network of the european electric grid. *Applied Energy*, in press.

[67] Richard Dobbs, Sven Smit, Jaana Remes, James Manyika, Charles Roxburgh, and Alejandra Restrepo. Urban world: Mapping the economic power of cities. *McKinsey Global Institute*, 62, 2011.

[68] Felix Creutzig, Giovanni Baiocchi, Robert Bierkandt, Peter-Paul Pichler, and Karen C. Seto. Global typology of urban energy use and potentials for an urbanization mitigation wedge. *Proceedings of the National Academy of Sciences*, 112(20):6283–6288, Dec 2015. doi: 10.1073/pnas.1315545112.

[69] Jonn Axsen, Kenneth S. Kurani, Ryan Mccarthy, and Christopher Yang. Plug-in hybrid vehicle ghg impacts in california: Integrating consumer-informed recharge profiles with an electricity-dispatch model. *Energy Policy*, 39(3):1617–1629, 2011. doi: 10.1016/j.enpol.2010.12.038.

[70] Jae D. Kim and Mansour Rahimi. Future energy loads for a large-scale adoption of electric vehicles in the city of los angeles: Impacts on greenhouse gas (ghg) emissions. *Energy Policy*, 73:620–630, 2014. doi: 10.1016/j.enpol.2014.06.004.

[71] T.v. Ramachandra, Bharath H. Aithal, and K. Sreejith. Ghg footprint of major cities in india. *Renewable and Sustainable Energy Reviews*, 44:473–495, 2015. doi: 10.1016/j.rser.2014.12.036.

[72] Energy information administration (eia). u.s. electric system is made up of interconnections and balancing authorities, 2016. URL https://www.eia.gov/todayinenergy/detail.php?id=27152.

[73] Lauren H Logan and Ashlynn S Stillwell. Probabilistic assessment of aquatic species risk from thermoelectric power plant effluent: Incorporating biology into the energy-water nexus. *Applied energy*, 210:434–450, 2018.

[74] Lauren H Logan and Ashlynn S Stillwell. Water temperature duration curves for thermoelectric power plant mixing zone analysis. *Journal of Water Resources Planning and Management*, 144(9):04018058, 2018.

[75] Mesfin M Mekonnen, PW Gerbens-Leenes, and Arjen Y Hoekstra. The consumptive water footprint of electricity and heat: A global assessment. *Environmental Science: Water Research & Technology*, 1(3):285–297, 2015.

[76] Roberto Turconi, Alessio Boldrin, and Thomas Astrup. Life cycle assessment (lca) of electricity generation technologies: Overview, comparability and limitations. *Renewable and Sustainable Energy Reviews*, 28:555–565, 2013. doi: 10.1016/j.rser.2013.08.013.

[77] Landon Marston, Yufei Ao, Megan Konar, Mesfin M Mekonnen, and Arjen Y Hoekstra. High-resolution water footprints of production of the united states. *Water Resources Research*, 54(3):2288–2316, 2018.

[78] Form eia-923 detailed data with previous form data (thermoelectric cooling water data, 2019). URL https://www.eia.gov/electricity/data/eia923/.

[79] Jordan Macknick, Robin Newmark, Garvin Heath, and KC Hallett. Operational water consumption and withdrawal factors for electricity generating technologies: A review of existing literature. *Environmental Research Letters*, 7(4):045802, 2012.

[80] Thermoelectric cooling water data (u.s. energy information administration, 2019), . URL https://www.eia.gov/electricity/data/water/.

[81] Ranran Wang, Julie B Zimmerman, Chunyan Wang, David Font Vivanco, and Edgar G Hertwich. Freshwater vulnerability beyond local water stress: Heterogeneous effects of water-electricity nexus across the continental united states. *Environmental Science & Technology*, 51(17):9899–9910, 2017.

[82] Leipzig T. Wilson, W. and B. Griffiths-Sattenspiel. Burning our rivers: The water footprint of electricity. *River Network, Portland*, 2012. doi: PublicationNo.96âĂŞ170462p.

[83] Emily A Grubert. Water consumption from hydroelectricity in the united states. *Advances in Water Resources*, 96:88–94, 2016.

[84] Emissions generation resource integrated database (egrid) (environmental protection agency, 2019), Jul 2019. URL https://www.epa.gov/energy/emissions-generation-resource-integrated-database-egrid.

[85] Rebecca AM Peer and Kelly T Sanders. The water consequences of a transitioning us power sector. *Applied energy*, 210:613–622, 2018.

[86] Ranran Wang and Julie Zimmerman. Hybrid analysis of blue water consumption and water scarcity implications at the global, national, and basin levels in an increasingly globalized world. *Environmental Science Technology*, 50(10):5143–5153, Apr 2016. doi: 10.1021/acs.est.6b00571.

[87] Stephan Pfister, Dominik Saner, and Annette Koehler. The environmental relevance of freshwater consumption in global power production. *The International Journal of Life Cycle Assessment*, 16(6):580–591, Mar 2011. doi: 10.1007/s11367-011-0284-8.

[88] Laura Scherer and Stephan Pfister. Global water footprint assessment of hydropower. *Renewable Energy*, 99:711–720, 2016. doi: 10.1016/j.renene.2016.07.021.

[89] Arman Shehabi, Sarah Smith, Dale Sartor, Richard Brown, Magnus Herrlin, Jonathan Koomey, Eric Masanet, Nathaniel Horner, Inês Azevedo, and William Lintner. United states data center energy usage report. Technical report, Lawrence Berkeley National Lab.(LBNL), Berkeley, CA (United States), 2016.

[90] Mohan Ganeshalingam, Arman Shehabi, and Louis-Benoit Desroches. Shining a light on small data centers in the us. 2017.

[91] Ward Van Heddeghem, Sofie Lambert, Bart Lannoo, Didier Colle, Mario Pickavet, and Piet Demeester. Trends in worldwide ict electricity consumption from 2007 to 2012. *Computer Communications*, 50:64–76, 2014.

[92] Ashlynn S Stillwell, Carey W King, Michael E Webber, Ian J Duncan, and Amy Hardberger. The energy-water nexus in texas. *Ecology and Society*, 16(1), 2011.

[93] Sources of greenhouse gas emissions, Apr 2020. URL https://www.epa.gov/ghgemissions/sources-greenhouse-gas-emissions.

[94] S Pabi, A Amarnath, R Goldstein, and L Reekie. Electricity use and management in the municipal water supply and wastewater industries. *Electric Power Research Institute, Palo Alto*, 194, 2013.

[95] Li Ling, Quan Zhang, and Liping Zeng. Performance and energy efficiency analysis of data center cooling plant by using lake water source. *Procedia Engineering*, 205: 3096–3103, 2017.

[96] MJ Ellsworth, LA Campbell, RE Simons, MK Iyengar, RR Schmidt, and RC Chu. The evolution of water cooling for ibm large server systems: Back to the future. In *2008 11th Intersociety Conference on Thermal and Thermomechanical Phenomena in Electronic Systems*, pages 266–274. IEEE, 2008.

[97] Lizhe Wang and Samee U Khan. Review of performance metrics for green data centers: a taxonomy study. *The journal of supercomputing*, 63(3):639–656, 2013.

[98] Eric Masanet, Arman Shehabi, and Jonathan Koomey. Characteristics of low-carbon data centres. *Nature Climate Change*, 3(7):627–630, 2013.

[99] Lotfi Belkhir and Ahmed Elmeligi. Assessing ict global emissions footprint: Trends to 2040 & recommendations. *Journal of Cleaner Production*, 177:448–463, 2018.

[100] Jens Malmodin, Åsa Moberg, Dag Lundén, Göran Finnveden, and Nina Lövehagen. Greenhouse gas emissions and operational electricity use in the ict and entertainment & media sectors. *Journal of Industrial Ecology*, 14(5):770–790, 2010.

[101] Victor Depoorter, Eduard Oró, and Jaume Salom. The location as an energy efficiency and renewable energy supply measure for data centres in europe. *Applied Energy*, 140: 338–349, 2015.

[102] Inigo Goiri, Kien Le, Jordi Guitart, Jordi Torres, and Ricardo Bianchini. Intelligent placement of datacenters for internet services. In *2011 31st International Conference on Distributed Computing Systems*, pages 131–142. IEEE, 2011.

[103] Elsa Maurice, Thomas Dandres, Reza Farrahi Moghaddam, Kim Nguyen, Yves Lemieux, Mohamed Cherriet, and Réjean Samson. Modelling of electricity mix in temporal differentiated life-cycle-assessment to minimize carbon footprint of a cloud computing service. In *ICT for Sustainability 2014 (ICT4S-14)*. Atlantis Press, 2014.

[104] Bora Ristic, Kaveh Madani, and Zen Makuch. The water footprint of data centers. *Sustainability*, 7(8):11260–11284, 2015.

[105] Md AB Siddik, Christopher Matthew Chini, and Landon Marston. Water and carbon footprints of electricity are sensitive to geographical attribution method. *Environmental Science & Technology*, 2020.

[106] data center map: Colocation usa. URL https://www.datacentermap.com/usa/.

[107] datacenterhawk: Search for data centers. URL https://www.datacenterhawk.com/providers.

[108] Arman Shehabi, Eric Masanet, Hillary Price, Arpad Horvath, and William W Nazaroff. Data center design and location: Consequences for electricity use and greenhouse-gas emissions. *Building and Environment*, 46(5):990–998, 2011.

[109] R Steinbrecher and Roger Schmidt. Data center environments. *ASHRAE Journal*, 53: 42–49, 2011.

[110] What is u.s. electricity generation by energy source?, . URL https://www.eia.gov/tools/faqs/faq.php?id=427&t=3.

[111] Kristen Averyt, Jordan Macknick, J Rogers, N Madden, J Fisher, J Meldrum, and R Newmark. Water use for electricity in the united states: an analysis of reported and calculated water use information for 2008. *Environmental Research Letters*, 8(1): 015001, 2013.

[112] Jordan Macknick, Robin Newmark, Garvin Heath, and KC Hallett. Review of operational water consumption and withdrawal factors for electricity generating technologies. Technical report, National Renewable Energy Lab.(NREL), Golden, CO (United States), 2011.

[113] Emily A Grubert. Water consumption from hydroelectricity in the united states. *Advances in water resources*, 96:88–94, 2016.

[114] Emily Grubert. Conventional hydroelectricity and the future of energy: Linking national inventory of dams and energy information administration data to facilitate analysis of hydroelectricity. *The Electricity Journal*, 33(1):106692, 2020.

[115] Meredith Reitz, Ward E Sanford, GB Senay, and Jeffrey Cazenas. Annual estimates of recharge, quick-flow runoff, and evapotranspiration for the contiguous us using empirical regression equations. *JAWRA Journal of the American Water Resources Association*, 53(4):961–983, 2017.

[116] Emissions & generation resource integrated database (eGRID), 2017. URL https://www.epa.gov/energy/emissions-generation-resource-integrated-database-egrid.

[117] Bradley G Ridoutt and Stephan Pfister. A revised approach to water footprinting to make transparent the impacts of consumption and production on global freshwater scarcity. *Global Environmental Change*, 20(1):113–120, 2010.

[118] Joseph S Colett, Jarod C Kelly, and Gregory A Keoleian. Using nested average electricity allocation protocols to characterize electrical grids in life cycle assessment. *Journal of Industrial Ecology*, 20(1):29–41, 2016.

[119] Christopher M Chini, Lucas A Djehdian, William N Lubega, and Ashlynn S Stillwell. Virtual water transfers of the us electric grid. *Nature Energy*, 3(12):1115–1123, 2018.

[120] Hifld: Control areas.

[121] Annual electric power industry report, form eia-861 detailed data files, .

[122] Form no. 714 - annual electric balancing authority area and planning area report, 2017. URL https://www.ferc.gov/docs-filing/forms/form-714/data.asp.

[123] Khosrow Ebrahimi, Gerard F Jones, and Amy S Fleischer. A review of data center cooling technology, operating conditions and the corresponding low-grade waste heat recovery opportunities. *Renewable and Sustainable Energy Reviews*, 31:622–638, 2014.

[124] Tony Evans. The different technologies for cooling data centers. *APC white paper*, 59, 2012.

[125] SP Fisenko, AA Brin, and AI Petruchik. Evaporative cooling of water in a mechanical draft cooling tower. *International Journal of Heat and Mass Transfer*, 47(1):165–177, 2004.

[126] Alfonso Capozzoli and Giulio Primiceri. Cooling systems in data centers: State of art and emerging technologies. *Energy Procedia*, 83:484–493, 2015. doi: 10.1016/j.egypro.2015.12.168.

[127] Anne-Marie Boulay, Jane Bare, Lorenzo Benini, Markus Berger, Michael J Lathuillière, Alessandro Manzardo, Manuele Margni, Masaharu Motoshita, Montserrat Núñez, Amandine Valerie Pastor, et al. The wulca consensus characterization model for water scarcity footprints: assessing impacts of water consumption based on available water remaining (aware). *The International Journal of Life Cycle Assessment*, 23(2):368–378, 2018.

[128] Brian D Richter, Dominique Bartak, Peter Caldwell, Kyle Frankel Davis, Peter Debaere, Arjen Y Hoekstra, Tianshu Li, Landon Marston, Ryan McManamay, Mesfin M Mekonnen, et al. Water scarcity and fish imperilment driven by beef production. *Nature Sustainability*, pages 1–10, 2020.

[129] Evan Kodra, Seth Sheldon, Ryan Dolen, and Ory Zik. The north american electric grid as an exchange network: An approach for evaluating energy resource composition and greenhouse gas mitigation. *Environmental science & technology*, 49(22):13692–13698, 2015.

[130] Uisung Lee, Hui Xu, Jesse Daystar, Amgad Elgowainy, and Michael Wang. Aware-us: Quantifying water stress impacts of energy systems in the united states. *Science of the total environment*, 648:1313–1322, 2019.

Appendix A

Supplementary Information for Chapter 2

Introduction The supporting information contains more detailed information on the implementation of each attribution method. Text S1 contains sections pertaining to each of the seven different attribution methods detailed in Table 1 in the main document. Additionally, Figure S1 shows the different boundaries for each of the attribution methods. Figure S2 compares water and carbon intensities for each of the attribution methods. Figure S3 compares the changes in water and carbon intensity with the balancing of electricity demand in the defined boundary. Finally, we include two data sets that contain all the water and carbon intensities for each metropolitan statistical area (MSA) across all study years. These data are in units of m3/MWh for water intensity and kg CO2e/MWh for carbon intensity.

Text S1: Attribution Method Description and Assumptions 1. Interconnections Water consumption of each power plant is aggregated to their corresponding interconnect (Eastern Interconnection, Western Interconnection, and ERCOT). The Eastern Interconnection covers most of North America, extending from the Rocky Mountains to the Atlantic Ocean, excluding most of Texas. The Western Interconnection extends from the Rocky Mountains to the Pacific Ocean. ERCOT covers most of Texas (see Figure S1.A). The contributions to each of these interconnections can be determined using EIA form 923 and its NERC Region. Once electricity is generated and delivered to the grid it flows everywhere within an AC-interconnection (Hoffman et al., 2015). Between interconnections, there are only DC-ties so electricity has to be converted to cross the border from one interconnection into another (Hoffman et al., 2015), which can be considered negligible. We estimate water intensity from a complete mix of electricity contributed by all the power plants. 2. Balancing authority with and without transfers Balancing authority maintains the demand and supply of electricity within a portion of electric grid under the control of a single dispatcher (EPA, 2018). There are 31 balancing authorities in the Eastern Interconnection and 37 balancing authorities in the Western Interconnection. The Texas interconnection consists of a single balancing authority. Balancing authorities are comprised of power control areas (PCAs; Figure S1.B). The specific PCAs and balancing authorities for each power plant can be found using the U.S. Environmental Protection Agency's (EPA) eGRID database. Chini et al. (2018) combined this database with the power plant specific information from the Energy Information Administration (EIA) to create a water footprint for each balancing authority. If an MSA draws from powerplants within multiple PCAs, all of these PCAs are assumed to supply electricity to that MSA in proportion to their generation capacity. The Federal Energy Regulatory Commission reports annual electricity transfers between balancing authorities (or power control areas) in Form 714. Applying water and carbon intensities to distributed electricity creates a transfer of virtual water and emissions between balancing authorities. Therefore, importing electricity from a balancing authority with more water (carbon) intensive electricity production effectively raises the importing balancing authority's water (carbon) footprint. While the EIA provides monthly evaluations of water consumption for each power plant, electricity transfers are currently only available at the annual timescale within the United States. Therefore, there is a loss of temporal resolution when including transfers of electricity between balancing authorities. 3. eGRID subregions The EPA eGRID subregions (Figure S1.C) are compromised territories between NERC region and balancing authorities that limit the import and export of electricity (EPA, 2018). The EPA provides a representative map with approximate boundaries of the eGRID subregions. eGRID boundaries are not a rigid geographical feature; rather, eGRID's assign power plants to specific groups to better represent and report the environmental impacts. We aggregated the generation, water consumption, and carbon emission within each eGRID subregion and then calculated the water and carbon intensities associated with the eGRID subregion, similar to Peer et al. (2019). We then determined the weighted average water and carbon intensity of each MSA based on eGRID subregions supplying the MSA's electricity. 4. Basin scale The United States Geological Survey (USGS) developed a hierarchical system to classify different scales of hydrologic units called the Hydrologic Unite Code (HUC). Here, we analysis hydrologic subregions (HUC-4; Figure S1.D), which is between the finer basin delineation (HUC-8) used by Tidwell et al. (2016) and the coarser spatial resolution (HUC-2) used by Kelley and Pasqualetti (2013). 5. Radius from city Chini et al., 2016 used the distance from the urban core to connect the water consumption embedded within nearby electricity production to the urban area where the electricity was consumed. Each MSA is depicted in Figure S1.E. We set attribution boundaries as the distance from the centroid of an MSA (evaluated at 50 km, 100 km, and 200 km). All power plants within these boundaries were assumed to provide electricity to the MSA and the associated environmental impacts were assigned to the MSA. 6. State State-level analysis often focus on a single state and do not consider the embedded resources or emissions imported from outside the state boundary, leading to under accounting of embedded resources or emissions (Ruddell et al., 2014). Following the work of Ruddell et al. (2014), we assumed electricity generated within a state first satisfies in-state demand and then balances the excess or deficiency by trade among the states. State boundaries within CONUS are shown in Figure S1.F.

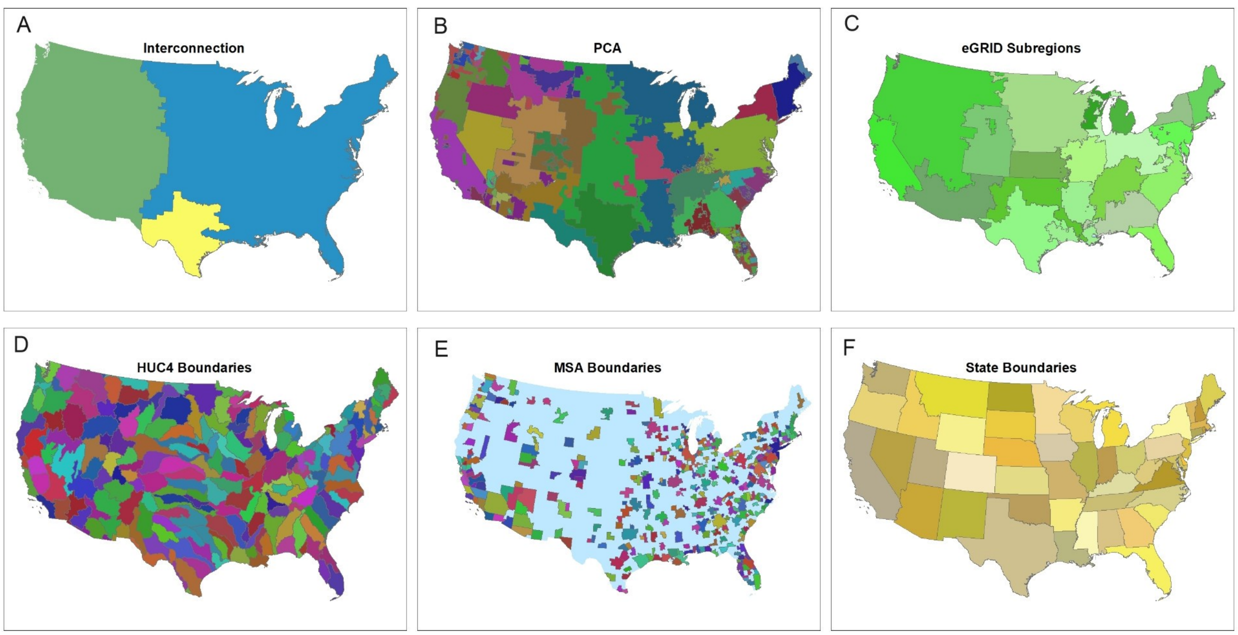


Figure A.1: *The geographic extent of decision boundaries for water and carbon attribution vary across the country.*

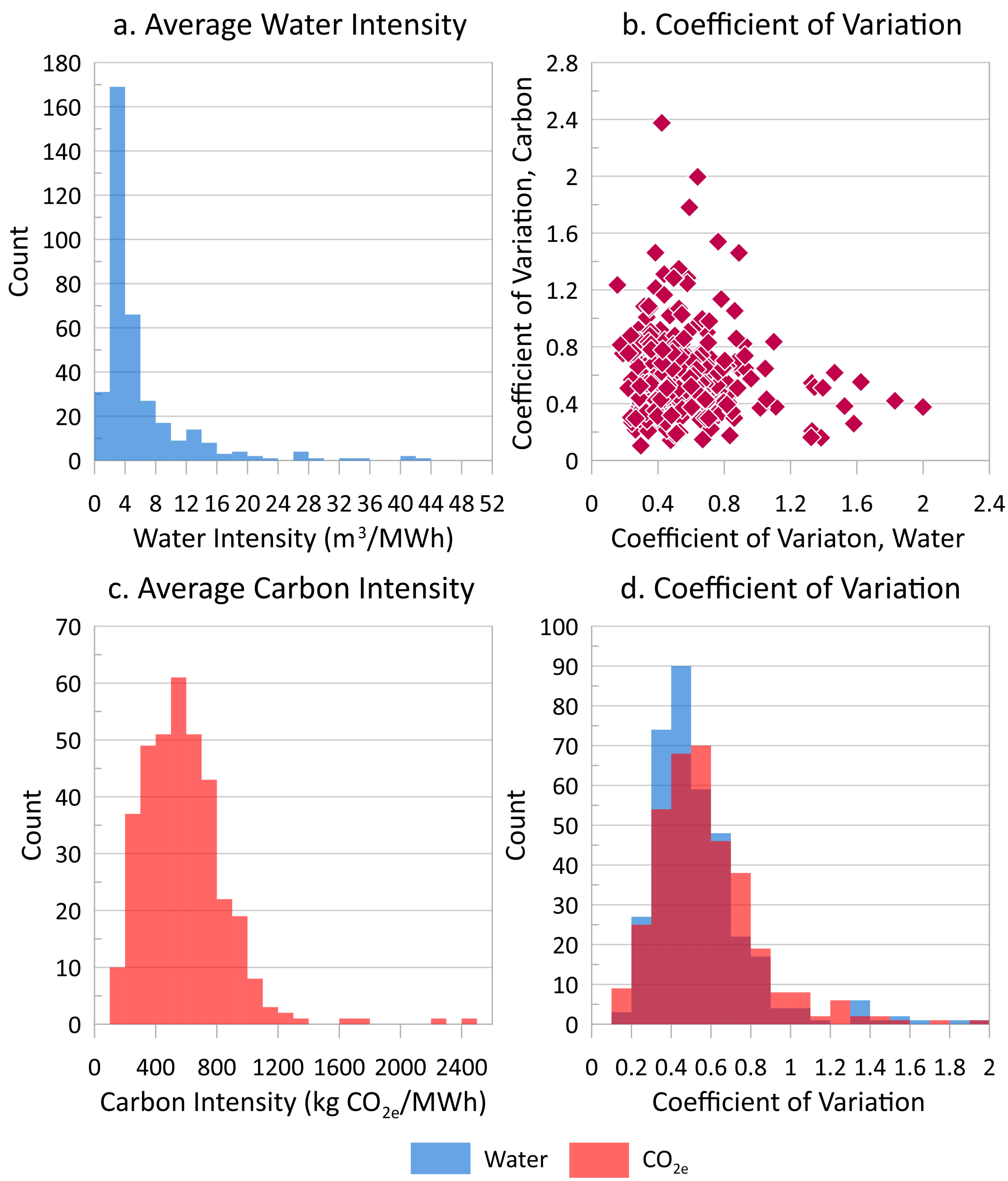


Figure A.2: *There are tradeoffs between water intensity and carbon intensity based on the attribution method. However, these tradeoffs are location specific as the distributions of coefficient of variation are relatively similar (d) and there is no correlation between variability between water and carbon (b). Here, the number of MSAs with a given (a) water intensity and (c) carbon intensity are depicted.*

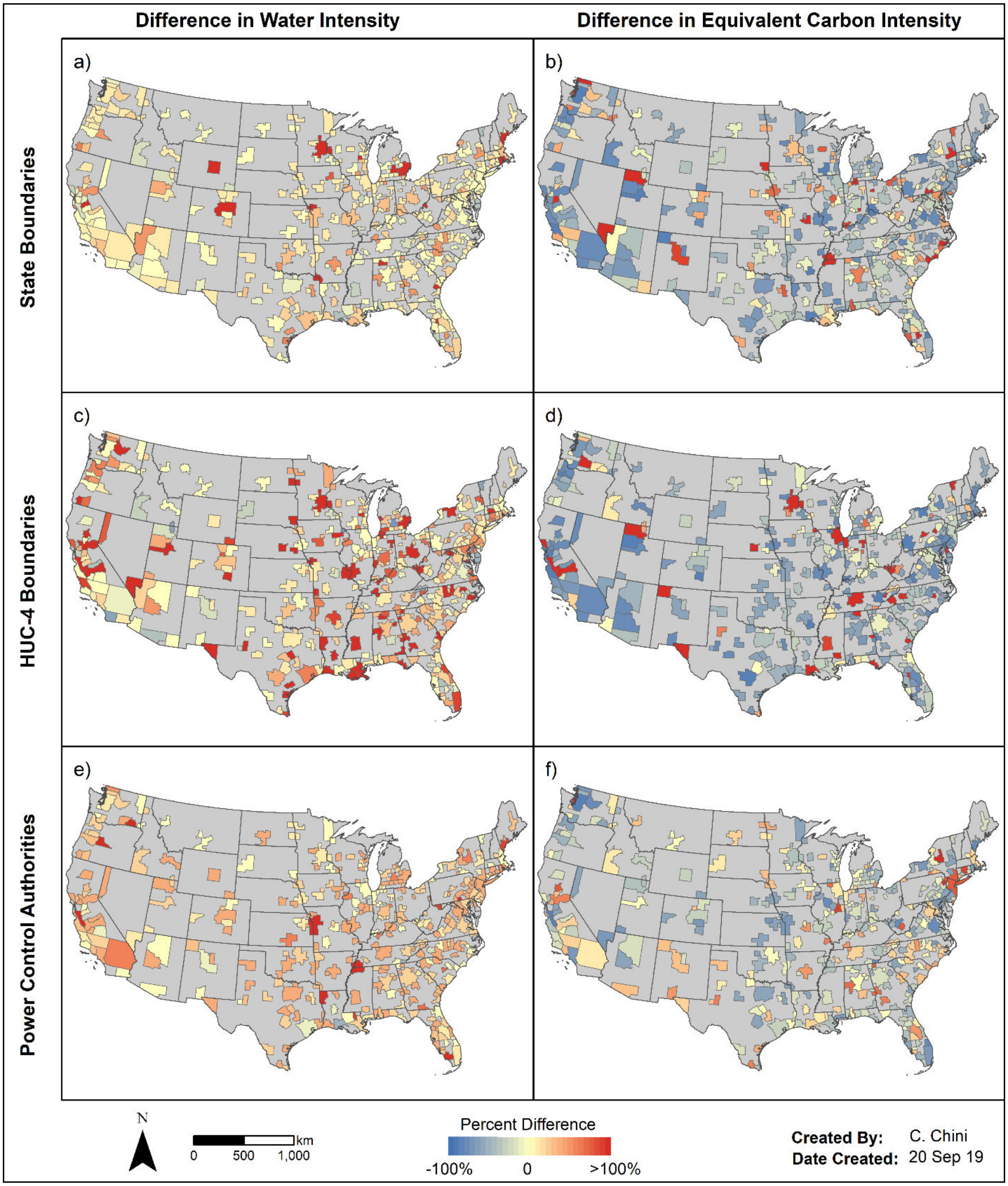


Figure A.3: *Differences in water and equivalent carbon intensity for each MSA was assessed with and without the inclusion of electricity transfers across state (a,b), HUC-4 (c,d), and power control authorities (PCAs; e,f). For most MSAs, the balancing of electricity through inter-boundary transfers led to minimal change in estimated water. The percent change in water intensities was the most significant for the HUC-4 attribution method. In general, the inclusion of electricity balancing across attribution boundaries led to an increase in an MSA's water intensity of electricity consumption. Conversely, the emission intensity with balancing power across state, HUC-4, and PCA boundaries generally decreased compared to the original boundaries.*

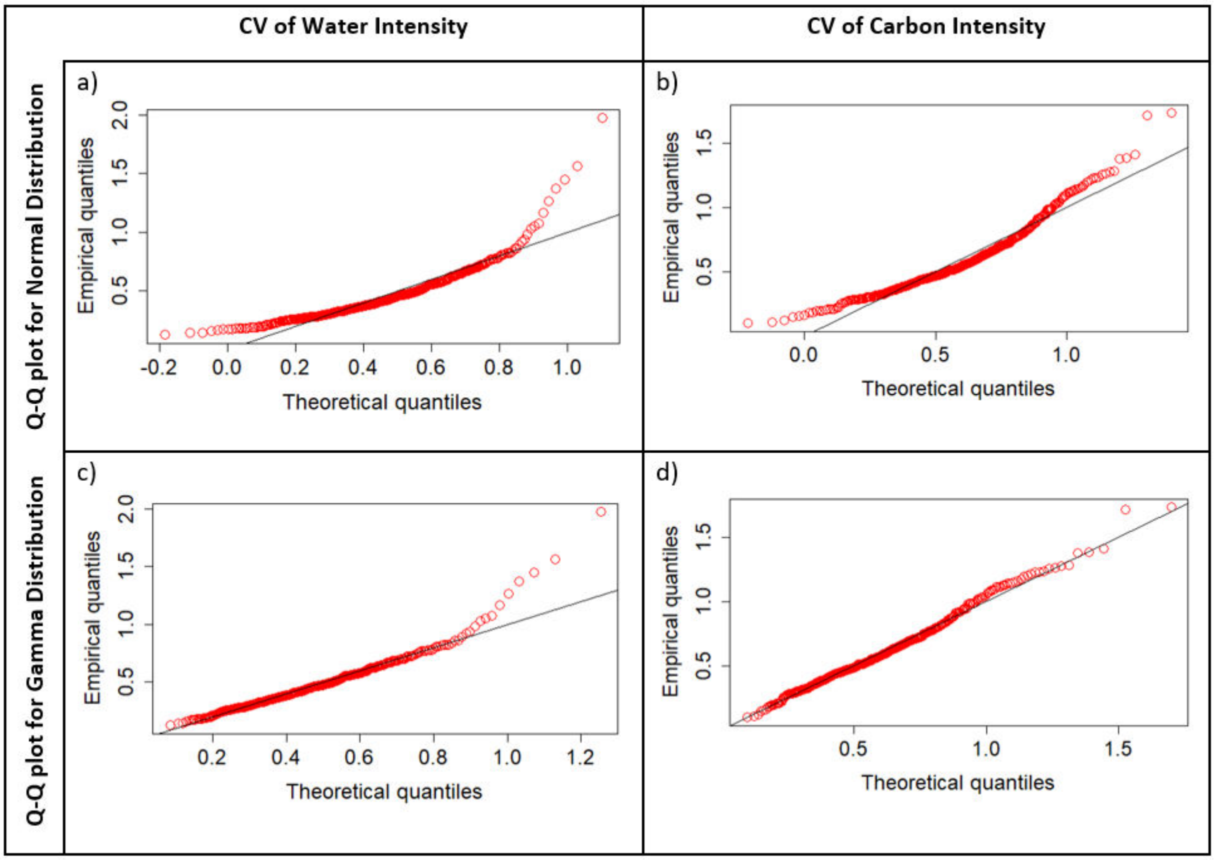


Figure A.4: *Co-efficient of variation (CV) for both water and carbon intensity follow gamma distribution.*

Appendix B

Supplementary Information for Chapter 3

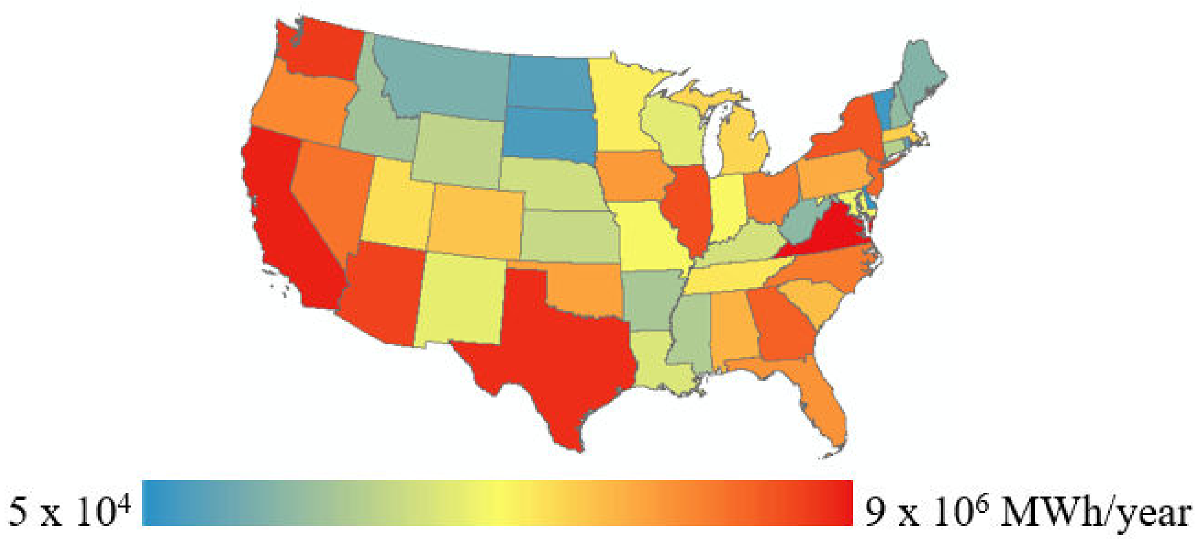
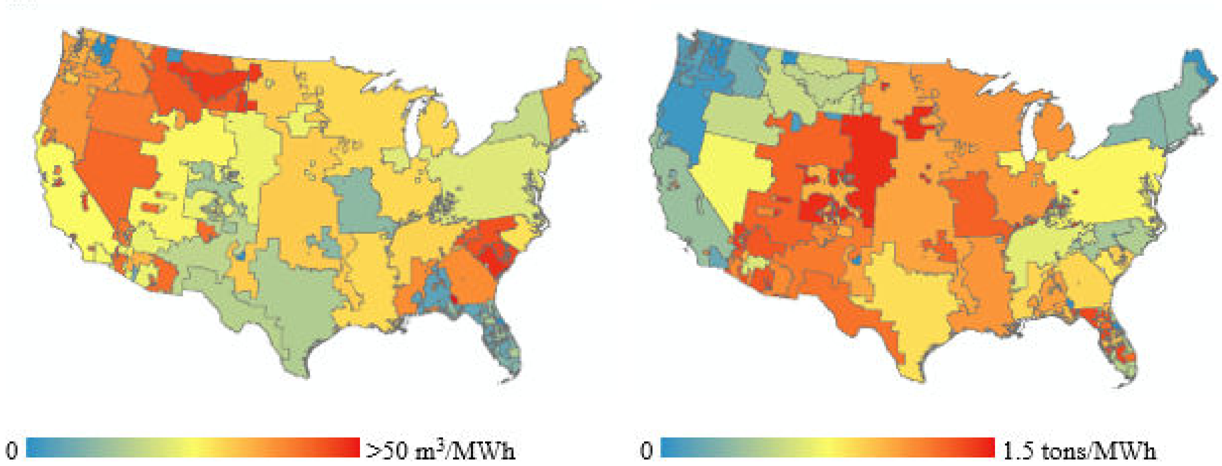
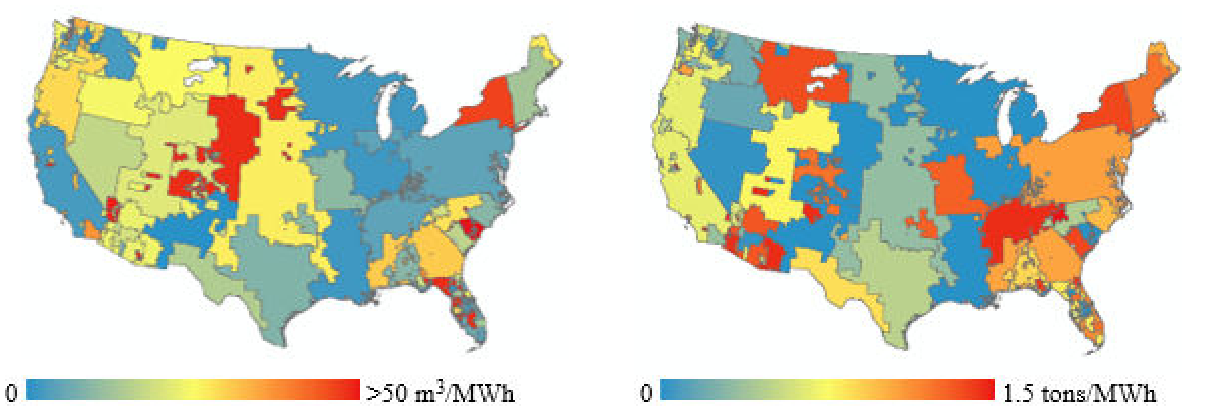


Figure B.1: *Data centers electricity consumption in 2018 for each state.*

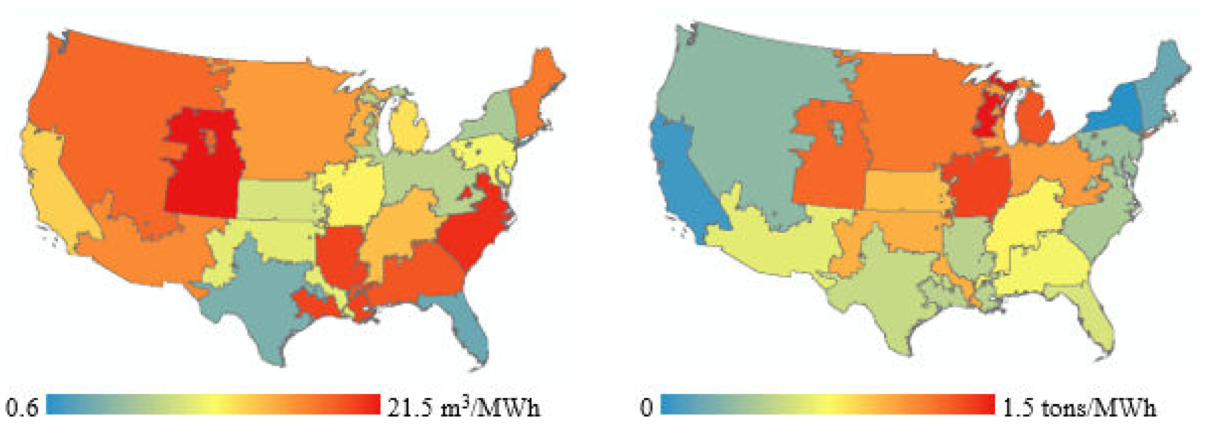
(a) PCA with transfer



(b) PCA without transfer



(c) eGRID subregions



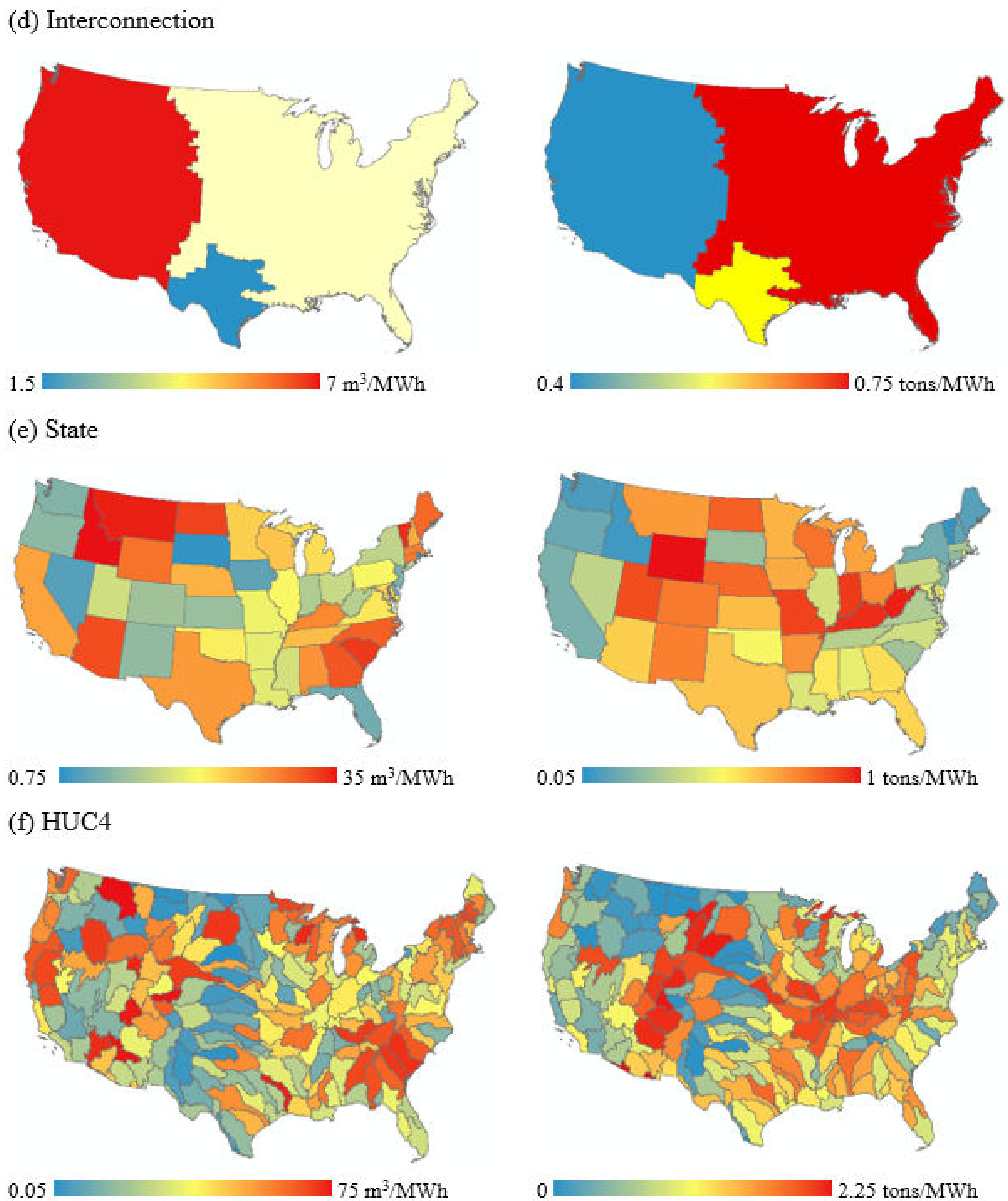


Figure B.2: *Spatial variability of water intensity (m3/MWh) and carbon intensity (tons of eCO2/MWh) of electricity for the attribution boundaries across the US.*

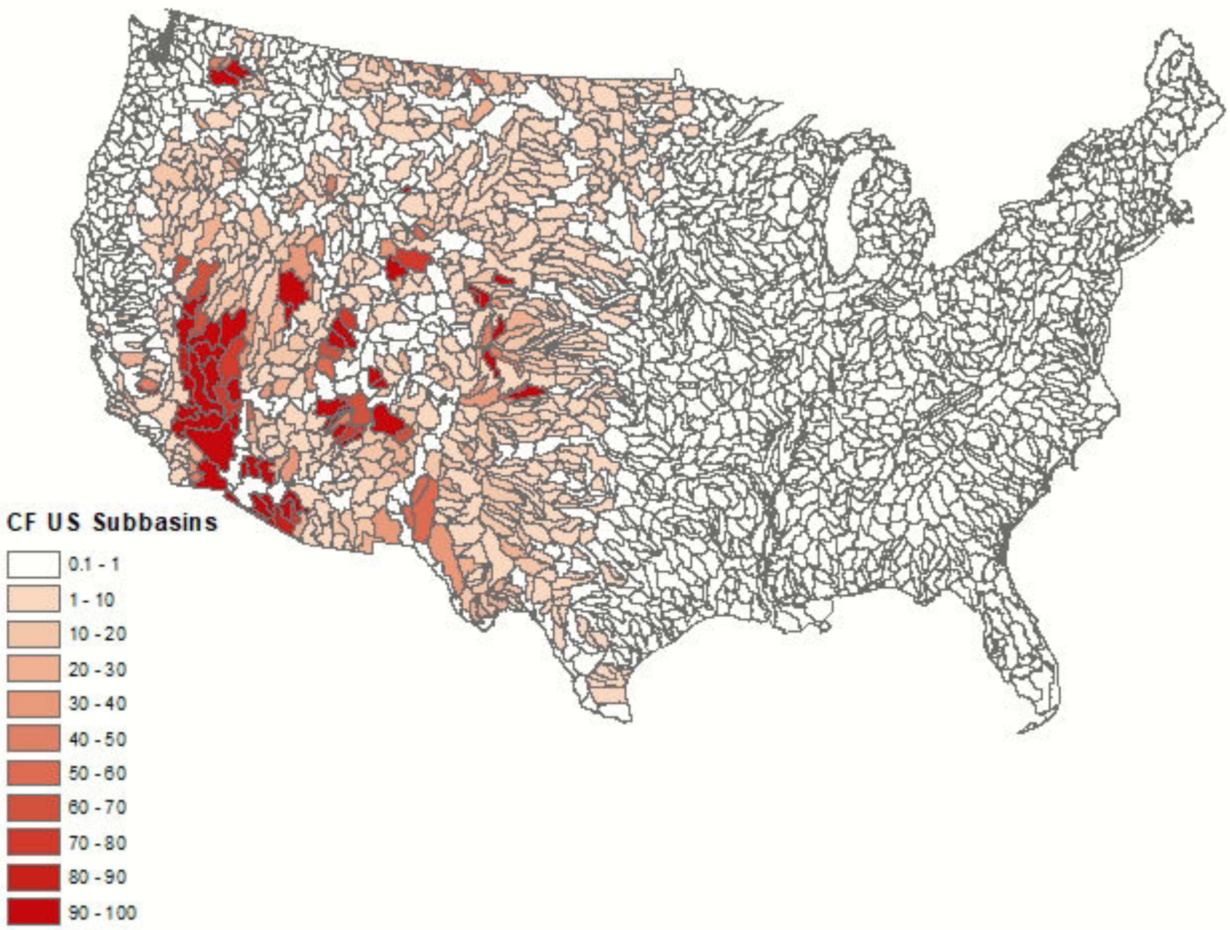


Figure B.3: *Characterization factors for US subbasins at HUC-8 level.*

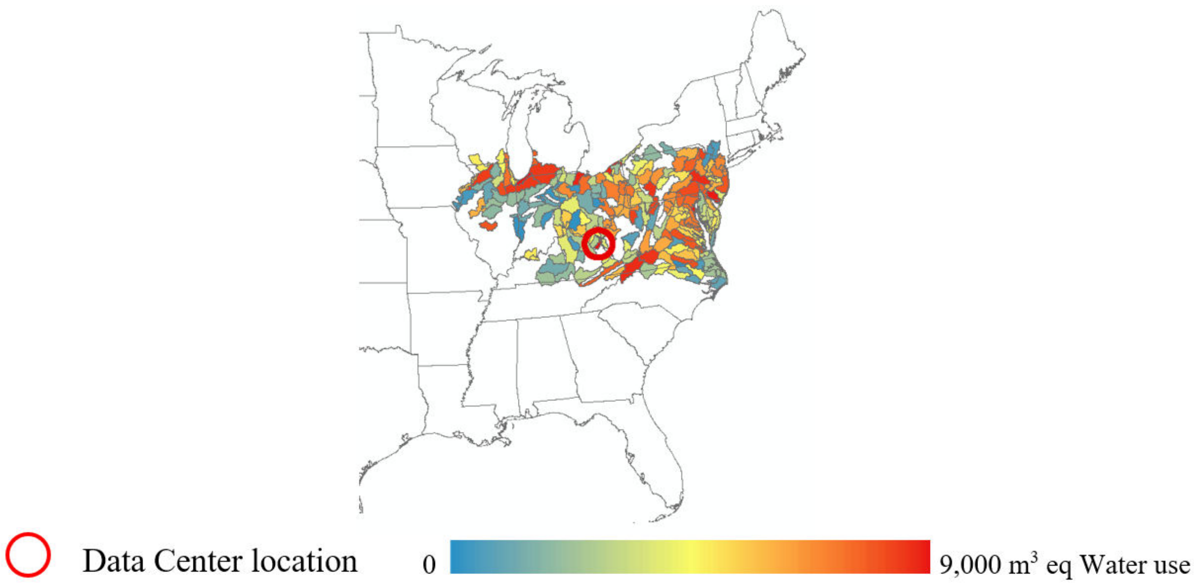


Figure B.4: *Water scarcity footprint (WSF) of a 1 MW data center located at a specific watershed in the Eastern US.*