Applied Geography 121 (2020) 102264

Contents lists available at ScienceDirect

Applied Geography

journal homepage: http://www.elsevier.com/locate/apgeog

Disparities in community water system compliance with the Safe Drinking Water Act

Zoe Statman-Weil, Leora Nanus *, Nancy Wilkinson

Department of Geography & Environment, San Francisco State University, 1600 Holloway Ave., San Francisco, CA, 94132, USA

ARTICLE INFO

Keywords:
Environmental justice
Safe Drinking Water Act
Spatial analysis
GIS
Water resources

ABSTRACT

Understanding potential disparities in community compliance with the Safe Drinking Water Act (SDWA) will help managers effectively and fairly allocate funding for improving drinking water systems. Environmental justice is experienced at the scale of the individual, the household, the neighborhood, or the water service provider and is too fine-grained a scale for spatial analysis with most available state datasets. This research represents an effort to ascertain which spatial analysis method selected to estimate the demographics of water service areas best supports environmental justice analysis at the state scale. To understand whether there are disparities in compliance with the SDWA, this study specifically investigated the relationship between socioeconomic status and race as well as other water system variables, and violations to the SDWA of community water system (CWS) in the state of Pennsylvania. The sociodemographic characteristics of the water systems were estimated using three different CWS-level spatial analysis methods (areal weighting, dasymetric mapping, areal interpolation) and a county-level spatial analysis method. Negative binomial regression was used to evaluate whether these sociodemographic characteristics, and other water system variables, including size of the water system, ownership, and water source, are associated with SDWA violations. This research demonstrates that the spatial analysis method selected for an environmental justice study can affect the results and conclusions of the research. Evidence that SDWA violations were associated with race or socioeconomic status was not found; however, this study did determine that small CWSs (<200 connections) and CWSs serving rural areas are less likely to be compliant with the SDWA.

1. Introduction

1.1. Background – Safe Drinking Water Act

Environmental justice research that investigates potential disparities in unequal access to safe drinking water based on race and/or socioeconomic status (SES) has increased since 2014 as a result of the drinking water crisis in Flint, Michigan. Due to the corrosive quality of its new water source and the lack of appropriate corrosion control, the lead levels in Flint's drinking water spiked, resulting in an increase in blood lead levels of the community's children (Campbell, Greenberg, Mankikar, & Ross, 2016). Several case studies have concluded that environmental injustice played a role in authorities' slow response to evaluate and address this crisis (Butler, Scammell, & Benson, 2016; Campbell et al., 2016). These studies highlight the need to conduct environmental justice analyses of public utility water systems' compliance with the Safe Drinking Water Act (SDWA). While compliance with the SDWA does not ensure safe drinking water quality as there may be contaminants detected that are not currently regulated, it does suggest safer drinking water, and data on health-based and non-health-based violations to the SDWA is the best nationally available public data to use to predict the drinking water quality of public water systems.

Under the SDWA, enacted in 1974 (amended in 1986, reauthorized in 1996) the U.S. Environmental Protection Agency (EPA) is required to identify and develop rules related to harmful contaminants in drinking water distributed by public water systems (PWSs) (U.S. EPA, 2017a).[1] Although the SDWA was enacted to protect drinking water in the U.S., it does not guarantee Americans clean water (Balazs, 2011). The EPA

* Corresponding author.
E-mail address: lnanus@sfsu.edu (L. Nanus).

[1] Although this project looks at community water systems in its analysis, PWSs are referenced throughout the paper in situations where the cited source or study references PWSs. PWS is the umbrella term that is defined by the EPA as serving at least 25 people or having at least 15 service connections for at least 60 days a year. Community water systems are a subset group that serve the same population year-round (U.S. EPA, 2017a).

https://doi.org/10.1016/j.apgeog.2020.102264
Received 19 April 2019; Received in revised form 23 May 2020; Accepted 4 June 2020
Available online 20 June 2020

regulates certain contaminants, including nitrate, arsenic, and lead, for which it has determined a maximum contaminant level goal (MCLG) where there would be no expected health risk over a lifetime of exposure on a daily basis (U.S. EPA, 2004; 2017b). For most contaminants, the EPA then sets the maximum contaminant level (MCL), the determined enforceable limit of the contaminant allowed to be distributed by PWSs, as close to the MCLG as possible, taking treatment cost into consideration (U.S. EPA, 2004; 2017b). If there is no reliable method of detecting contaminants, or if the EPA does not consider it economically or technically feasible to set an MCL, a required treatment technique is then used by the PWS (U.S. EPA, 2004; 2017b). The EPA is required to reassess its regulations of all contaminants under the SDWA every six years; however, it has not established a regulatory standard for a new contaminant in approximately 20 years (Environmental Working Group, 2017; Fedinick, Wu, Pandithartne, & Olson, 2017).

If a PWS does not comply with an existing SDWA standard, it is cited for a violation. Not all SDWA violations indicate a contaminant was found above its MCL; a violation could mean an administrative SDWA rule was violated (U.S. EPA, 2017c). Health-based violations consist of exceedances of MCLs, exceedances of disinfectant concentration thresholds, and improper water treatment (U.S. EPA, 2017c). Non-health-based violations include failure to monitor regularly or report results on-time, failure to notify the public that there is a serious health problem with drinking water, and failure to publish annual consumer confidence reports (U.S. EPA, 2017c). Even seemingly non-health-based violations, such as failure to monitor and report, can mask an underlying health-based issue (Fedinick et al., 2017; Wallsten & Kosec, 2008). Thus, it is important to look at disparities across all SDWA violations.

Quantitative environmental justice studies have examined the relationship between sociodemographic factors and socioeconomic status (SES) of communities and nitrate and arsenic concentrations as indicators of drinking water quality. A recent nationwide environmental justice study on nitrate concentrations by community water system (CWS) found that the percentage of the population served by a CWS that is Hispanic was positively correlated with nitrate concentrations, while the percentage of low-income residents and black residents were both negatively correlated (Schaider, Swetschinski, Campbell, & Rudel, 2019). Balazs, Morello-Frosch, Hubbard, and Ray (2011) compared nitrate concentrations at points of entry into each CWS in California's San Joaquin Valley to the sociodemographic characteristics of the population potentially exposed, and found that CWSs with less than 200 service connections had greater nitrate concentrations in communities with a higher Latino population (Balazs et al., 2011). Balazs, Morello-Frosch, Hubbard, and Ray (2012) found that smaller CWSs in San Joaquin Valley serving areas with higher home ownership rates (an indication of economic security) had lower arsenic concentrations, and the inverse was true for CWSs serving more people of color (Balazs et al., 2012).

Recent studies have analyzed national trends in drinking water violations with respect to environmental justice indicators. The Environmental Justice Coalition (EJC) assessed disparities in access to safe drinking water at the county level in California between 1995 and 2000, and found that counties with the greatest number of drinking water violations had the highest percentage of people of color, people living below the poverty line, and Latinos (all calculated separately) (The Environmental Justice Coalition for Water, 2005). Switzer and Teodoro (2017) investigated the impact of race, ethnicity, and SES on compliance with the SDWA on a national scale, with a focus on health-based regulations. They concluded that the racial and ethnic composition of the community served by a PWS is a predictor of drinking water quality, and that violations to the SDWA in low-SES communities are strongly predicted by the percentage of the population that is black and Hispanic (Switzer & Teodoro, 2017). Allaire, Wu, and Lall (2018) analyzed spatial and temporal trends in Safe Drinking Water Information System (SDWIS) data by county and by CWS between 1982 and 2015 at the national level, assigning each CWS census information of the county in which it was located. They found rural CWSs were more likely to have violations, and that CWSs serving low-income communities of color were more likely to receive total coliform violations (Allaire et al., 2018). McDonald and Jones (2018) also conducted a national environmental justice study on SDWA violations at the county level (2011–2015) and found that initial violations were more prevalent in smaller systems, and that systems serving communities of color with low SES had a greater chance of initial and repeat violations.

1.2. Spatial environmental justice analysis

Analyzing the sociodemographic characteristics and SES of populations potentially exposed to environmental harms is critical to environmental justice and health equity research. Geographic Information Systems (GIS) are useful for quantitative environmental justice analysis (Holifield, Chakraborty, & Walker, 2017). Typically, geospatial environmental justice analysis identifies an environmental "bad" (e.g., a polluting facility) or an environmental "good" (e.g., a park) and evaluates the sociodemographic and socioeconomic characteristics of the community affected by the target feature, often with the use of a defined buffer of impact (Chakraborty, Maantay, & Brender, 2011). Since these discrete or continuous areas do not usually coincide spatially with available geospatial population data, researchers must identify a method that will best estimate the characteristics of a population within the exposure area (Holifield et al., 2017). Commonly used methods in the related literature include areal weighting, dasymetric mapping and areal interpolation described below:

- **Areal weighting:** This method assigns a proportion of the population to the affected area relative to the percent of the geographic unit within the discrete affected area boundary (e.g., a buffer) (Holifield et al., 2017). This method assumes equal distribution of the population within the census tracts.
- **Dasymetric mapping:** This advanced analysis technique utilizes additional ancillary data, such as land use or zoning data, to better estimate the population distribution within a given geographic unit (Mennis, 2003).
- **Areal Interpolation:** This geostatistical kriging-based method creates a continuous prediction surface from polygon data that can be reaggregated to new polygons (Krivoruchko, Gribov, & Krause, 2011).

Previous research on drinking water disparities did not include complete datasets of the digital boundaries of the water systems studied so these standard methods of estimating populations within a given buffer could not be applied. Cory and Rahman (2009) averaged the contaminant levels of arsenic for every PWS within a given zip code, and then compared that average concentration of arsenic to the demographic characteristics of the zip code. Balazs et al. (2011) compared two techniques of estimating the demographics of the population served by each CWS for their analysis of nitrate contamination in drinking water in San Joaquin Valley, California. The first method used areal weighting with the use of digitized CWS boundaries, the second included averaging the demographic characteristics of every census block that contained a CWS source (well field, surface water intake, and treatment plants). The second approach was found to be adequate and utilized for their analysis. Given the range of methods currently used and important implications for CWSs, there is a critical need to determine whether the spatial analysis method selected for an environmental justice study on safe drinking water can affect the results and conclusions of the research.

This study is an environmental justice analysis of water system compliance with the SDWA and focuses on the following research questions:

1) Are there socioeconomic disparities in CWS compliance with the SDWA? Are there more violations (total and health-based) in low-

income communities, communities with a higher proportion of people of color, and/or rural communities?
2) How do the results differ depending on the spatial analysis method used to estimate the demographic characteristics of the population served by the CWS?
3) Are there additional characteristics of the CWS such as size of the water system, public versus private ownership, and water source parameters that influence compliance with the SDWA?

This is the first time multiple spatial analysis methods have been used through the entirety of an environmental justice study on SDWA compliance in order to assess how the analysis method affects the results.

1.3. Environmental justice parameters

Percent below poverty line will be used as the proxy variable for SES in this study. The percentage of non-Hispanic whites within the population is the second demographic variable, one used frequently in environmental justice studies (Balazs et al., 2011; Cory & Rahman, 2009). The type of CWS ownership (i.e., private vs. public) may have an effect on SDWA compliance and has been included in several previous studies (Allaire et al., 2018; Balazs et al., 2012, 2011; Konisky & Teodoro, 2016; Wallsten & Kosec, 2008). Publicly owned public water systems have been found to have more violations than privately owned public water systems (Konisky & Teodoro, 2016). However, one study that focused primarily on the effects of ownership on the number of SDWA violations found ownership type did not affect compliance (Wallsten & Kosec, 2008). Smaller systems generally have less technical managerial and financial capacity (TMF) for proper regulation and enforcement of the SDWA, and thus system size is a key variable (National Research Council, 1997). As noted above, Balazs et al. (2011) found evidence of environmental injustice in small systems with less than 200 connections but not in larger systems. Other studies have included size of the water system as well (Konisky & Teodoro, 2016). SDWA violations or drinking water contamination have also been found to be higher in rural areas compared to urban areas (Allaire et al., 2018).

Water source parameters, both groundwater vs. surface water and purchased vs. unpurchased water, are also included variables in drinking water quality studies (Allaire et al., 2018; Balazs et al., 2011, 2012; Switzer & Teodoro, 2017). Other covariates related to the characteristics of the CWS will also be factored into the statistical model based on the literature and are described in greater detail in the methods section.

2. Study area

Pennsylvania was selected as the study area as it is one of a few states that have publicly available PWS or CWS boundary data. It is the sixth most populous state in the U.S. (12.8 million people) and ranks ninth in population density (286 people/square mile). The two largest cities are Philadelphia (1.6 million people) and Pittsburgh (305,000 people) (Cedar Lake Ventures Inc., 2018).

Pennsylvania has a higher percentage white population than the rest of the U.S., with 77.7% white non-Hispanic, approximately 15% higher than the U.S. as a whole. The largest minority populations are black (11.0%) and Hispanic (6.1%), both smaller than their respective proportions in the U.S. Rural Pennsylvania is heavily white non-Hispanic (Cedar Lake Ventures Inc., 2018). The mean income in Pennsylvania is similar to the mean in the U.S., at $54.9 thousand annually (Cedar Lake Ventures Inc., 2018). The percentage of households on food stamps (13.0%) ranks Pennsylvania 26th in the U.S. in this poverty metric (Cedar Lake Ventures Inc., 2018).

3. Methods

3.1. Description of data

The source of each dataset is identified below, followed by a description of any data processing that was conducted, including the criteria used for selection or categorization.

The Pennsylvania Department of Environmental Protection (PA DEP) has produced a near-complete GIS shapefile of the active state CWS boundaries (Pennsylvania Department of Environmental Protection, 2017).[2] The dataset contains 1853 CWS boundaries, which represent over 90% of the states' CWS boundaries. Review of the dataset determined that 20 of the water systems were either not CWSs or were inactive. These were removed and 1833 active CWS boundaries remained. This PWS data was used to identify the area served by each CWS.

SDWA violation data was queried and downloaded from the online SDWIS database (U.S. Environmental Protection Agency, 2018) for January 1, 2012 to December 31, 2016, a five-year period, to assess the operation of the CWSs. Following the methods of Allaire et al. (2018), the time period in which a violation occurred was defined by the compliance period begin date. It is assumed that a violation did not occur within the time period of interest if a CWS did not have a violation entry in the SDWIS database for the years selected. Since the PA DEP geospatial public water supply boundary data only includes active CWS data, "CWS" was selected as the PWS Type and "Active" as the activity status. These data were processed to produce a dataset that includes the number of total violations and number of health-based violations within the time period of interest. The data were then joined to the PA DEP CWS boundary shapefile, and all CWSs that did not have any corresponding SDWIS violation data were identified as having zero violations.

The population served by each CWS was estimated based on census tract-level data using a 2017 geodatabase containing census tract and county boundaries from the U.S. Census Bureau (U.S. Census Bureau, 2017). Census tracts were selected as the unit of analysis as data associated with census tracts reflect a larger number of survey respondents and a smaller margin of error than census block groups, and thus provide a more reliable estimate of the population (Ogneva-Himmelberger & Huang, 2015).

Population and demographic characteristics of the Pennsylvania census tracts and counties from the U.S. Census Bureau were also included in the analysis (U.S. Census Bureau, 2016a). Total tract population, percent population below the poverty level and percent people of color (including the Hispanic population) were compiled from the 2011–2015 American Community Survey (ACS) 5-Year Estimates. The census tract data contained 3218 tracts, and the county data contained 67 counties (U.S. Census Bureau, 2017). A threshold of >0.24 km^2 was used for census tract size to allow for application of the areal interpolation method, resulting in 3166 census tracts.

Delineated urban areas in Pennsylvania were also obtained from the U.S. Census Bureau (U.S. Census Bureau, 2016b). For the CWS-level analysis, a CWS was classified as urban when 50% or more of its service area was located within an urban area; otherwise it was classified as rural. The percentage of the CWS within an urban area was determined to be the best available method to estimate whether the water system was primarily urban or rural. 754 of the CWSs were classified as urban and 1079 were classified as rural.

Alternatively, counties were classified as rural or urban based on the percentage of the population living within a rural or urban area in a given county according the U.S. Census. The classification method was

[2] Although the data is titled "Public Water Supplier's (PWS) Service Areas," the metadata states that all non-transient noncommunity water systems and transient noncommunity water systems are excluded (Pennsylvania Department of Environmental Protection, 2017), leaving only community water systems.

different for the county-level analysis compared to the CWS-level analysis because the county rural/urban demographics are readily available from the U.S. Census Bureau (U.S. Census Bureau, 2010). If 50% or more of the county's population was defined as urban, then the county was classified as urban. Of the 67 counties, 37 were classified as urban and 30 were classified as rural.

Land cover data obtained from the U.S. Geological Survey helped estimate the population served by each water system. A shapefile identifying residential land was created by selecting the land classified in the 2011 National Land Cover Database (NLCD) (U.S. Geological Survey, 2011) as developed at low, medium, and high intensity, which consists of areas with 20%–100% impervious surface land cover according to the NLCD metadata. These three NLCD classes are all defined as residential in the NLCD; they include single-family housing units, and areas where people live and work at a high density, such as apartments, row houses, and commercial and industrial areas. Some studies have used the same classification techniques to identify residential land cover (Ogneva-Himmelberger & Huang, 2015).

3.2. Spatial environmental justice analysis

Four different spatial analysis methods were selected to estimate the

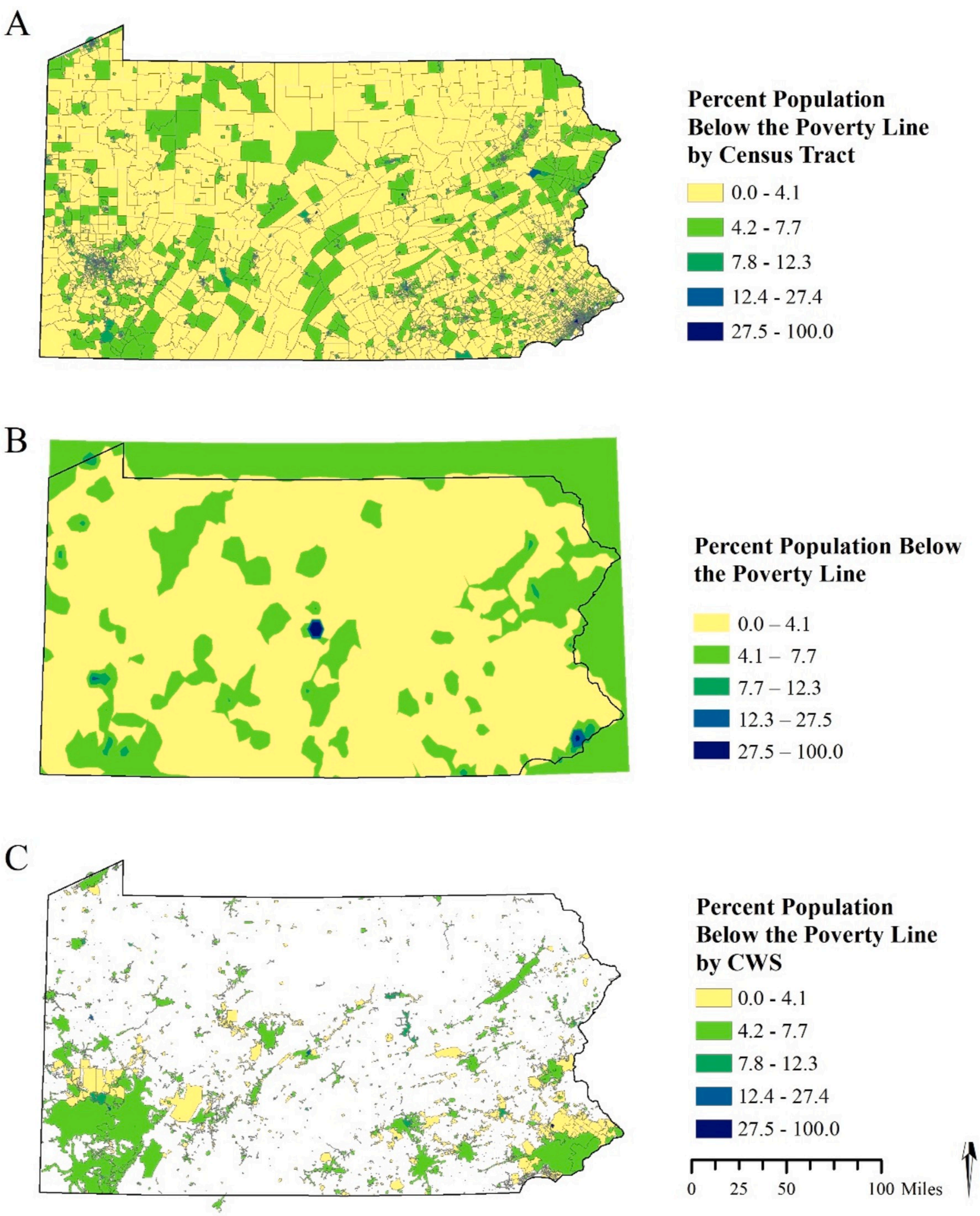

Fig. 1. Visual display of the steps of areal interpolation: A) percent below the poverty line by census tract; B) continuous prediction surface of the percent below the poverty line; and C) estimates of percent below the poverty line reaggregated to the CWS boundaries.

demographic and socioeconomic characteristics of the population served by each CWS. These estimates were then compared to the total violations and health-based violations of each CWS. The first three methods were areal weighting, dasymetric mapping, and areal interpolation (the “CWS-level analyses”). The last method used the demographic characteristics of the county in which the CWS is located (the “county-level analysis”). The city or cities in which the CWS is located were not used to estimate demographics of the community served because like counties, the city boundary does not necessarily match up with the water system boundary and Pennsylvania’s PWS are missing a “city served” designation in the SDWIS database.

Areal Weighting: An areal weighting model was developed in GIS (ArcMap 10.5.1) to estimate the population served by a given CWS based on the proportion of the intersecting census tracts that lie within the CWS boundary. Several of the smaller CWSs had estimates of zero people served (n = 82), and these were excluded from the analysis.

Dasymetric Mapping: The areal weighting model was modified such that the population served by each CWS was estimated based on the proportion of the intersecting census tracts’ residential land within the CWS boundary, based on NLCD data. Dasymetric mapping is the most detailed and provides a more realistic distribution of the population (Holifield et al., 2017); it is used as the standard for comparison in this paper. CWSs where the model estimated zero people served were excluded from the analysis (n = 84).

Areal Interpolation: The population of a given characteristic within each CWS was estimated using the binomial areal kriging model, which is designed to be used for rate data. Using the visual variography tools available in the Geostatistical Wizard (ArcMap 10.5.1), a kriging interpolation model was fit to a plot of covariance versus distance. A Stable model was used for the percent below the poverty line data and a Spherical model was used for the percent people of color data. The model parameters, such as lattice spacing, lag size, and number of lags, were adjusted with the goal of achieving a standardized root mean square as close to 1.0 as possible. The result was a continuous prediction surface of the population characteristics, including percentage of the population below the poverty line. This data was then reaggregated to the CWS boundaries. Fig. 1 shows a visual display of the steps of this method.

County-Level Analysis: The last method assigned the percent people of color and the percent below the poverty line of the surrounding county to each CWS. Although some CWSs serve more than one county, the primary one was determined based on the county listed in the SDWIS database. The rural or urban classification of the county was also assigned to the CWS.

Since CWSs had to be excluded from the areal weighting and dasymetric models, these same CWSs were excluded from all four analyses to make the results comparable. A total of 1694 CWSs were assessed.

3.3. Statistical analysis

A negative binomial regression model was used to determine which characteristics of the water system and population served by the water system best predicted the number of violations. A univariate linear regression analysis was first conducted to evaluate each variable across the spatial analysis methods. Since the dependent variable is a count of SDWA violations, a Poisson model was determined to be more appropriate than a linear regression model (Switzer & Teodoro, 2017; Wallsten & Kosec, 2008). However, a Poisson model requires the variance and the mean of the dependent variable to be equal, and the variance of the SDWA violations is much greater than the mean of the violations, a situation referred to as overdispersion (NCSS Documentation, 2018; Switzer & Teodoro, 2017; Wallsten & Kosec, 2008). Thus, negative binomial regression, a model similar to Poisson which allows for overdispersion, was used (NCSS Documentation, 2018; Switzer & Teodoro, 2017; Wallsten & Kosec, 2008). Negative binomial regression was applied to all four spatial analysis results for both health-based violations and total violations.

Characteristics of water systems are often included in environmental justice studies on drinking water quality as they are considered potentially confounding variables, but also can provide insight into the operations of CWSs. The covariates used in the model were selected based on the related literature. For this study the following variables were included in the regression model: 1) percent below the poverty line; 2) percent non-Hispanic white; 3) rural/urban; 4) private/public; 5) purchased/not purchased; 6) size of the system (simplified to include systems with less than 200 and more than 200 connections); and 7) groundwater/surface water source. The multicollinearity was tested using the variance inflation factor (VIF), which confirmed that none of the selected variables were strongly correlated (VIF less than 5), and all could be included in the regression model.

4. Results

4.1. Spatial analysis

The spatial distribution of the number of total and health-based SDWA violations by CWS for Pennsylvania is shown in the Supplementary Material, Fig. S1 and the summary statistics of this data are shown in Table 1. There is no clear pattern in the spatial distribution of the total SDWA violations by CWS or health-based SDWA violations (Fig. S1).

4.1.1. CWS-level analyses

The summary statistics of the relevant SES and sociodemographic data by census tract in Pennsylvania are shown in the Supplementary Material, Table S1, and their spatial distribution is shown in the Supplementary Material, Fig. S2. There is no clear pattern in the spatial distribution of the population below the poverty line by census tract (Fig. S2A). However, census tracts with a higher percentage people of color appear to be located in the larger urban centers, specifically Philadelphia and Pittsburgh (Fig. S2B).

The results of the three CWS-level spatial analysis methods for assessing percent below the poverty line by CWS are shown in Fig. 2A–C, and the results of these three methods for assessing percent people of color by CWS are shown in Fig. 3A–C. At the state scale, the spatial distribution of the percent below the poverty line by CWS does not appear to vary by spatial analysis method (Fig. 2A–C, Table 3). However, the percent people of color by CWS estimated using areal interpolation (Fig. 3C) appears to vary spatially compared to the results of the areal weighting and dasymetric mapping methods (Fig. 3A and B, respectively). The CWSs in the Philadelphia area in Fig. 3A and B have a higher percentage people of color than that shown in Fig. 3C. There is also slight variation in the CWSs surrounding Pittsburgh. The mean values of the percentage below the poverty line and people of color are relatively similar across CWS-level spatial analysis methods (Table 2).

4.1.2. County-level analysis

The summary statistics of the SES and sociodemographic data by county in Pennsylvania are shown in Supplementary Material, Table S1, and their spatial distribution is shown in Supplementary Material, Fig. S3. There is not a clear spatial pattern in the percent below the poverty line by county, but Fig. S3B shows the highest percentages of people of color reside in the counties surrounding Philadelphia.

Table 1
Violations per CWS in Pennsylvania (2012–2016).

Statistics	All	Health-based
Min	0	0
Max	804	41
Mean	24.4	0.75
Median	8	0
Standard Deviation (SD)	55.2	2.1

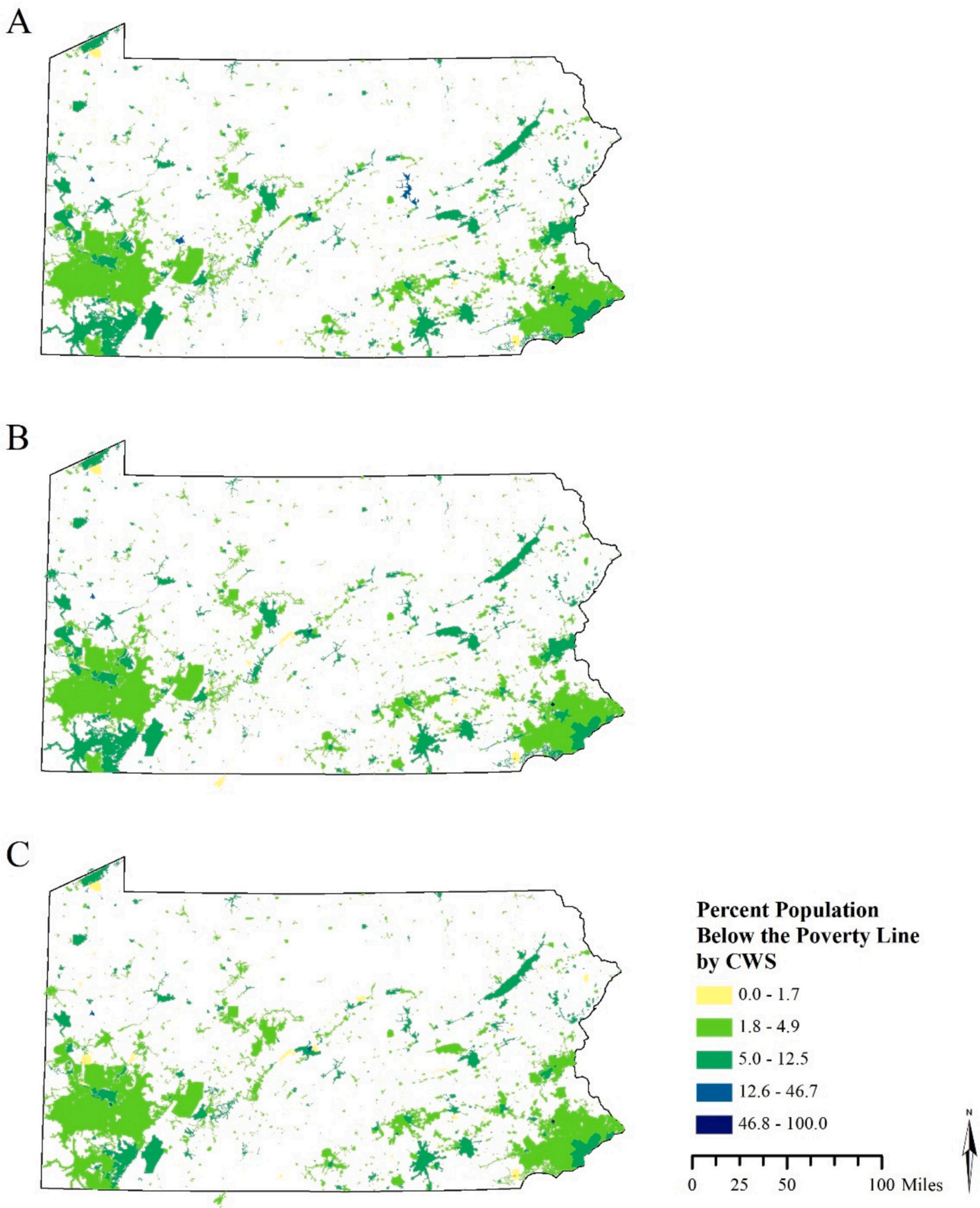

Fig. 2. Percent below the poverty line by CWS calculated using different spatial analysis methods: A) areal weighting; B) dasymetric mapping; and C) areal interpolation.

4.1.3. Summary statistics by CWSs

The summary statistics for sociodemographic and SES variables by CWS and the categorical variables by CWS are presented in Tables 2 and 3, respectively. The summary statistics of the continuous variables (percent below the poverty line and percent people of color) are relatively similar across the three CWS-level analyses but there are stark differences between these three analyses and the county-level analysis (e.g., percent below the poverty line maximum).

Scatter plot matrices of the percentage of people below the poverty line and the percentage people of color as estimated by the four spatial analysis methods are shown in Supplementary Materials, Figs. S4 and S5 respectively, along with the Pearson correlation coefficient between each set of estimates. Similar to the summary statistics, the county-level estimates are the least correlated with the estimates of the other three spatial analysis methods. The estimates from areal weighting and dasymetric mapping are the most correlated.

4.2. Statistical analysis

The results of unadjusted univariate regressions for each parameter for total and health-based SDWA violations are shown in Supplementary Materials, Tables S2 and S3, respectively. Results show the rural variable

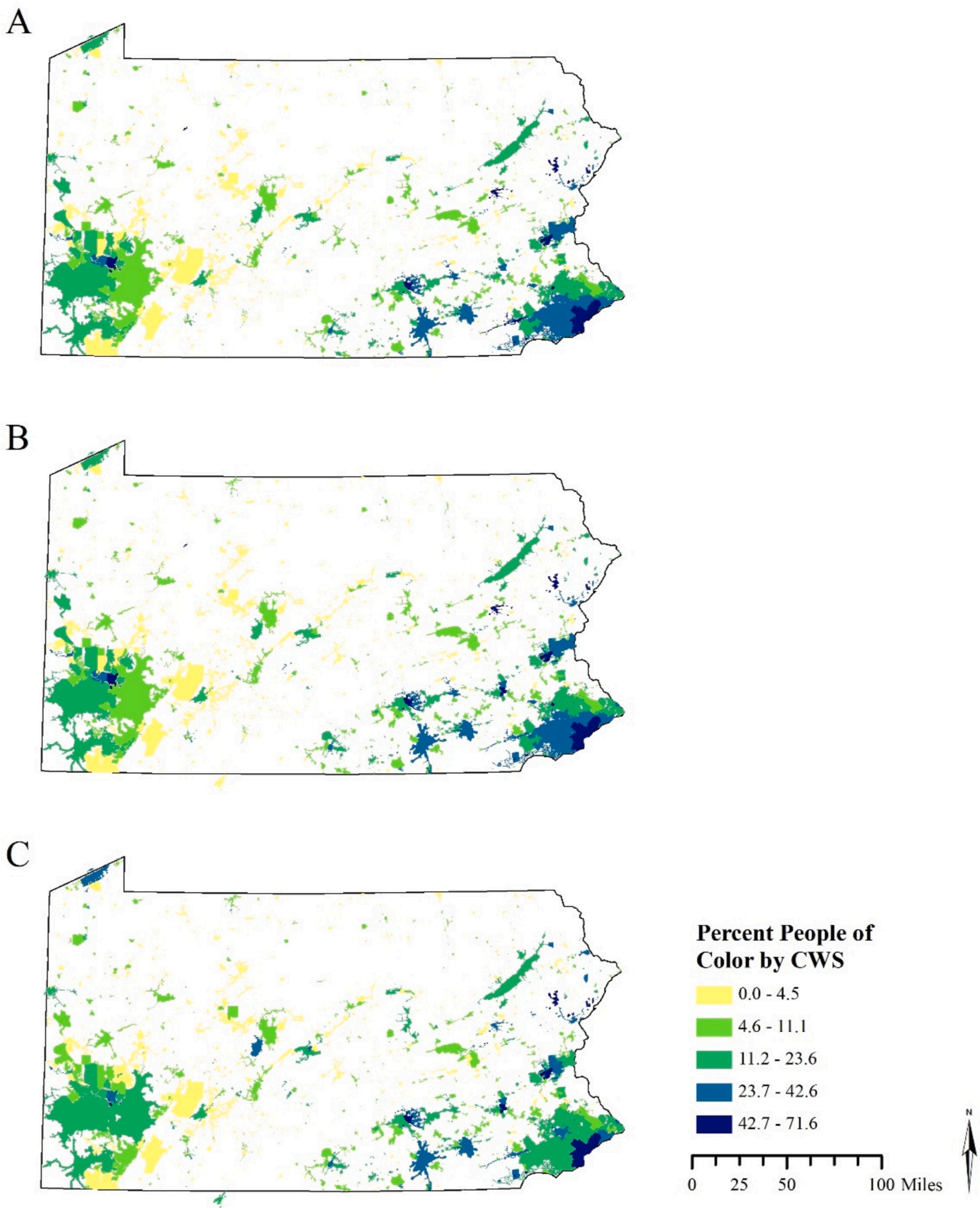

Fig. 3. Percent people of color by CWS calculated using different spatial analysis methods: A) areal weighting; B) dasymetric mapping; and C) areal interpolation. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

is significantly related to total SDWA violations ($p < 0.05$) in the CWS-level and County-level methods (Table S2). However, for percent below poverty line and percent people of color, differences were observed depending on selected spatial analysis method (Table S2). The CWS characteristic variable coefficients (e.g., size and water source) are identical across all spatial analysis methods since the data are the same and only one variable is being assessed. Purchased water, small size of CWS, and groundwater source are all significantly related to total SDWA violations (Table S2). Groundwater source was the only variable that was significantly related to health-based SDWA violations ($p < 0.001$) (Table S3).

In the areal weighting and dasymetric mapping negative binomial regression models, an increase in the percent of the population below the poverty line resulted in a decrease in the number of total SDWA violations ($p < 0.05$), while no significant effect was found in the other total violation models (Table 4). The percent below the poverty line did not have a significant effect on the number of health-based violations in the CWS-level models but did have a significant negative effect in the county-level model ($p < 0.05$) (Table 5). The percent people of color had no significant effect in any of the models (Tables 4 and 5).

The effects of the CWS characteristic variables only varied slightly by model. Small CWSs had more total violations in all the models (p <

Table 2
Summary statistics by CWS for continuous variables.

Statistics	Method			
	Areal Weighting	Dasymetric Mapping	Areal Interpolation	County-level
Percent Below the Poverty Line				
Min	0.00	0.00	0.00	6.00
Max	100	100	99.3	26.4
Mean	2.75	3.65	3.87	12.4
SD	4.57	4.24	3.46	3.17
Percent People of Color				
Min	0.00	0.00	0.00	2.40
Max	100	71.6	76.3	64.2
Mean	7.10	8.00	8.82	12.9
SD	10.7	10.1	10.3	8.27

Table 3
Count of categorical variables.

Variable	Yes	No
Rural	960	733
Public	723	970
Purchased	207	1486
Small size	951	742
Groundwater Source	1295	398

0.001) but did not have a significant relationship with health-based violations. While the groundwater variable was not significant in the models determining total violations, all models showed a negative relationship between CWSs with groundwater as a source and health-based violations ($p < 0.001$). All models found that CWSs with purchased water were more compliant in terms of both total and health-based SDWA violations ($p < 0.001$). The owner of the CWS (public vs. private) was not significant in any model ($p > 0.05$).

Rural CWSs were found to be less compliant across all models, except for the county-level model predicting health-based SDWA violations (Tables 4 and 5). Fig. 4 shows the mean total and health-based SDWA violations in the CWS-level analyses and the county-level analysis in rural and urban systems (only CWSs included in the regression analyses are factored into the figure). The county-level method estimated a slightly higher average of total SDWA violations (~33 violations/system compared to ~29 violations/system). The difference between rural and urban CWSs is more pronounced for the total SDWA violations compared to the health-based violations (Fig. 4). On average, a rural CWS experienced approximately 10–13 more total SDWA violations in the time period of interest than an urban CWS.

5. Discussion

5.1. Regression model comparison

The results of the SES and sociodemographic data coefficients show that the spatial analysis method selected can directly affect the conclusions of an environmental justice study at the state scale. The results of the regression models with regard to the percent below the poverty line were different for both sets of models predicting total and health-based SDWA violations. In the total SDWA violations analysis, only the areal weighting and dasymetric mapping methods found percent below the poverty line to have a significant effect ($p < 0.05$), while the county-level model found this variable had a significant negative effect on health-based SDWA violations ($p < 0.05$). These findings accord with recent environmental justice research that has also demonstrated that the spatial analysis method can have a direct impact on the outcome of a study (Maantay & Maroko, 2009; Ogneva-Himmelberger & Huang, 2015).

Areal weighting and dasymetric mapping had consistent regression results across all spatial analysis methods in terms of significance and direction. While dasymetric mapping is considered a more robust approach since it relies on more detailed datasets, this shows that at the state scale, utilizing more fine scaled data necessary for dasymetric mapping may be unnecessary for environmental justice studies.

The results of the regression models for both total and health-based SDWA violations were relatively similar across all four spatial analysis methods with regard to the CWS characteristics including rural vs urban, public vs private, size, and water source, due to the fact that all of the data other than the rural/urban variable came directly from the SDWIS database and do not vary across methods. The county-level analysis identified substantially fewer rural CWSs, likely because the method of categorizing a county as rural or urban was based on the proportion of people in the county living in rural areas whereas the CWS-level analysis was based on the proportion of a CWS within a U.S. Census Bureau-designated urban area. Since fewer people live in rural areas, a county had to be considerably rural to have a greater rural than urban population and thus be classified as rural. Despite this discrepancy, Fig. 4 and

Table 4
Determinants of total SDWA violations.

Variable	Method							
	Areal Weighting		Dasymetric Mapping		Areal Interpolation		County-level	
	Coeff.	SE	Coeff.	SE	Coeff.	SE	Coeff.	SE
Percent Below the Poverty Line	−0.018*	0.009	−0.018*	0.009	−0.019	0.011	−0.005	0.011
Percent People of Color	0.002	0.004	0.005	0.004	0.003	0.004	−0.009	0.005
Rural[a]	0.261***	0.074	0.285***	0.074	0.281***	0.074	0.359***	0.089
Public[b]	0.166	0.091	0.163	0.090	0.146	0.090	0.026	0.091
Purchased[c]	−0.947***	0.126	−0.944***	0.126	−0.942***	0.126	−0.862***	0.126
Small size[e]	0.390***	0.094	0.407***	0.093	0.410***	0.093	0.407***	0.092
Groundwater Source[d]	−0.165	0.109	−0.143	0.109	−0.157	0.108	−0.128	0.107
Constant	2.962***	0.133	2.916***	0.135	2.955***	0.135	3.141***	0.204
AIC	13,537.57		13,537.33		13,538.50		13,521.32	
Log-Likelihood	−6780.8		−6760.7		−6761.2		−6752.7	
Pearson chi2	3601.73		3526.81		3548.39		3503.2	

Coeff. = Coefficients.
SE = Standard Error.
Significance levels: *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.
[a] Reference group = urban.
[b] Reference group = private.
[c] Reference group = unpurchased.
[d] Reference group = large size.
[e] Reference group = surface water source.

Table 5
Determinants of health-based SDWA violations.

Variable	Method							
	Areal Weighting		Dasymetric Mapping		Areal Interpolation		County-level	
	Coeff.	SE	Coeff.	SE	Coeff.	SE	Coeff.	SE
Percent Below the Poverty Line	−0.016	0.015	−0.016	0.015	−0.010	0.017	−0.046*	0.018
Percent People of Color	0.007	0.006	0.009	0.006	0.007	0.006	0.006	0.008
Rural[a]	0.301*	0.121	0.311**	0.120	0.303*	0.120	0.277	0.145
Public[b]	0.012	0.146	0.019	0.146	−0.004	0.146	0.071	0.147
Purchased[c]	−0.809***	0.205	−0.803***	0.205	−0.802***	0.205	−0.836***	0.205
Small size[e]	0.192	0.153	0.209	0.151	0.202	0.151	0.274	0.150
Groundwater Source[d]	−1.077***	0.172	−1.060***	0.172	−1.071***	0.172	−1.127***	0.171
Constant	0.257	0.212	0.210	0.214	0.234	0.213	0.806*	0.327
AIC	3635.18		3634.37		3635.85		3633.22	
Log-Likelihood	−1809.6		−1809.2		−1809.9		−1808.6	
Pearson chi2	2259.45		2244.87		2230.4		2199.46	

Coeff. = Coefficients.
SE = Standard Error.
Significance levels: *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.
[a] Reference group = urban.
[b] Reference group = private.
[c] Reference group = unpurchased.
[d] Reference group = large size.
[e] Reference group = surface water source.

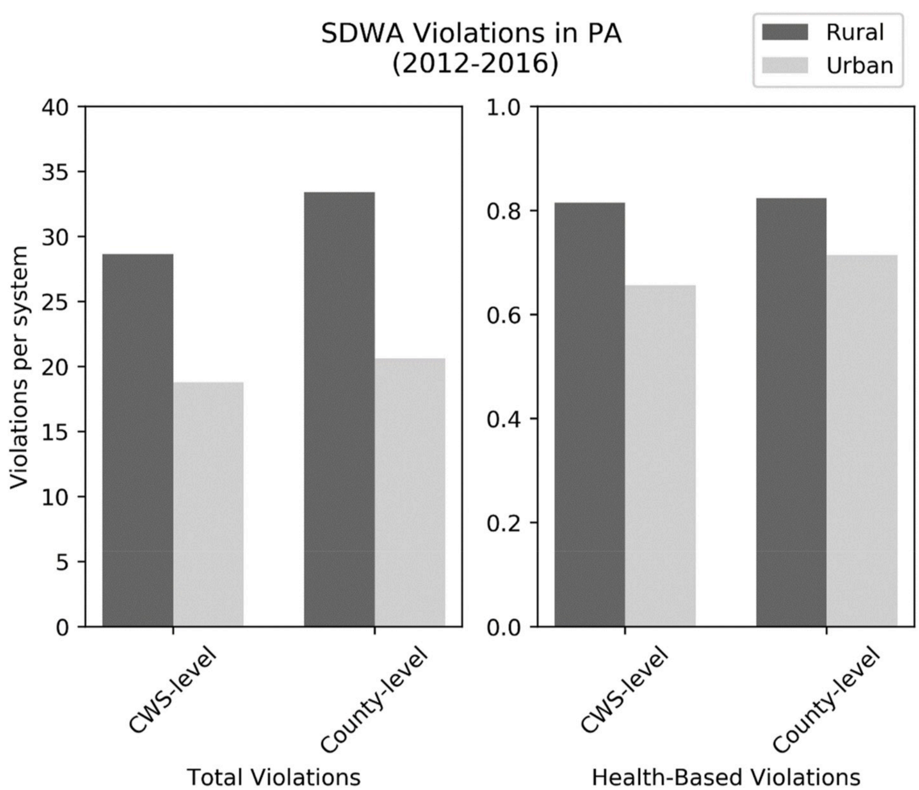

Fig. 4. Mean total and health-based SDWA violations in rural and urban CWSs as estimated by the CWS-level analyses and the county-level analysis.

the regression results (Tables 4 and 5) demonstrate that rural CWSs tend to have more total violations than urban CWSs. However, the relationship was not significant in the county-level model predicting health-based violations, possibly due to the different method of estimating rural CWSs.

The differences between the results of the county-level analysis compared to the CWS-level analyses were not surprising. CWSs may cover more than one county, and large portions of a county's area, and thus some of its population, may be factored into analysis of that CWS even when far outside the boundaries of a system. This situation is not unique to Pennsylvania and should be evaluated for other states across the U.S. as CWSs can cover more than one county, and many counties have areas not covered by CWSs.

5.2. *Environmental justice analysis*

This study's finding that poorer CWSs had fewer total SDWA violations across two of the models, and fewer health-based SDWA violations in one of the models is surprising, but not unusual. A national study that investigated violations by every CWS between 1997 and 2003 found that wealthier counties had more SDWA violations. That national study used income as a proxy for wealth (Wallsten & Kosec, 2008). Shaider et al. (2019) conducted a national environmental justice study on nitrate in drinking water, and found lower income residents associated with lower nitrate concentrations. A state-specific study found per capita income and average house value had no statistically significant effect on Arizona's implementation of a 2002 arsenic standard (Cory & Rahman, 2009).

This study showed no relationship between the percentage of people of color and total or health-based SDWA violations. Previous studies also did not find that a greater percentage of people of color necessarily resulted in more violations or in an increased concentrations of contaminants. In the study of Arizona's implementation of the new arsenic standard, Cory and Rahman (2009) did not find evidence that communities of color experienced inequitable implementation of the new standard. While Balazs et al. (2012) found that communities of color had greater odds of having an MCL violation, they did not find that these CWSs had higher arsenic concentrations. Balazs et al. (2011) also did not find a statistically significant relationship between race/ethnicity and nitrate levels in CWSs with more than 200 connections in San Joaquin Valley, California, although they did within communities of less than 200 connections. In sum, this study did not find conclusive evidence of environmental injustice in CWS violations by race or SES in Pennsylvania, as was the case in several related studies at the national and state scale.

Different water sources are susceptible to different types of contaminants and problems. Systems that rely on groundwater have fewer regulatory requirements compared to systems that utilize surface water (Allaire et al., 2018). This study found groundwater supply had a negative correlation with health-based violations, which is consistent with some recent research on SDWA health-based violations on the national scale (Switzer & Teodoro, 2017). However, groundwater supply has also been found to positively predict SDWA violations (Konisky & Teodoro, 2016). CWSs relying on purchased water were found to be more compliant across the board. The association of purchased water and fewer SDWA violations is supported by other studies' findings (Allaire et al., 2018; Switzer & Teodoro, 2017).

A significant finding of this study is that rural CWSs are likely to have more SDWA violations, and small CWSs are also likely to have more total violations. This is not surprising, as rural communities have a smaller customer base to generate the necessary revenue for proper treatment technology and maintenance and to hire experienced utility managers (National Research Council, 1997). These findings are consistent with Allaire et al. (2018), who found that rural areas had more violations than urban areas, and concluded that small, rural CWSs relying on surface water sources had the highest predicted probability of an SDWA violation. Wallsten and Kosec (2008) found that small privately-owned PWSs have fewer MCL health violations and more non-health-based violations indicating that the health-based violations of small systems may be masked by the lack of monitoring and reporting.

The distribution of population within Pennsylvania is unusual, which may explain the lack of a finding of environmental injustice. Rural areas in Pennsylvania tend to have a higher percentage of white individuals compared to states such as California (Cedar Lake Ventures Inc., 2018), where studies have found higher rates of violations in communities of color (Balazs et al., 2012; The Environmental Justice Coalition for Water, 2005). This may reflect the distribution of diversity in California, where many communities of color are rural (Cedar Lake Ventures Inc., 2018). The PA DEP defines its own Environmental Justice areas as census tracts with more than 30 percent people of color and over 20% of the population "in poverty" (Pennsylvania Department of Environmental Protection, 2019) These areas tend to be clustered in urban areas such as Philadelphia and Pittsburgh, rather than in rural areas, which have a much higher rate of violations (Allaire et al., 2018). Overall, Pennsylvania also has a high percentage of non-Hispanic white residents compared to other states (Cedar Lake Ventures Inc., 2018), and the chance of missing significant trends increases. Conducting the study in a single state limits the ability to analyze some of the results outside of the context of Pennsylvania.

There are additional potential limitations associated with this analysis. Compliance with the SDWA does not equate with safe drinking water quality, so this study cannot be considered an environmental justice analysis of access to safe drinking water. This is true for the many reasons discussed in the introduction, including that the EPA is slow to regulate new contaminants (Fedinick et al., 2017). The undereporting of violations within the SDWIS database could also have an impact on the count of total and health-based violations (Balazs, 2011; U.S. EPA's Office of Enforcement and Compliance Assurance, 2013). A 2011 report from the U.S. Government Accountability Office found that of 14 states audited in 2009, 26% of health-based violations and 84% of the monitoring violations were not reported accurately or at all (U.S. Government Accountability Office, 2011). Another potential limitation is that this analysis does not include areas served by private wells, which are more prevalent in rural areas and do not undergo routine monitoring for SDWA violations. This study also excluded over 100 CWSs as a result of the limitations of the different spatial analysis methods, most of which were small and rural, which could affect the results. There could be other confounders that are not included in this analysis.

Environmental justice analysis will differ based on the place and circumstance under analysis, as well as the scale and method of analysis. Baden, Noonan, and Turaga (2007) found that environmental justice research decisions regarding scale and scope can affect the findings. The scale (CWS) and scope (state-level) of this study could have had an impact on the results and conclusions. For example, in California race and rurality often coincide, reinforcing environmental justice concerns. In contrast, Pennsylvania's rural areas are overwhelmingly Caucasian while populations of color are highly urbanized. This may serve to mask or counteract potential environmental justice concerns associated with rural living. Thus, in other regions, rural communities of color may be at even greater risk.

5.3. Future research

Two primary findings of this study can help inform future research. First, the results of environmental justice studies can vary depending on the spatial analysis methods used to estimate the sociodemographic characteristics of communities of interest. This has important implications. When the CWS boundary data exists, it is best to use the most comprehensive method, which in this case would be dasymetric mapping. However, as so few states produce CWS boundary data, the county-level analysis method is sometimes the only option. For future national-scale county-level studies, it would be prudent to include a more robust CWS-level spatial analysis, such as dasymetric mapping, in a state with available CWS boundary GIS data for comparison.

A second, significant finding is that rural CWSs tend to experience more SDWA violations, and small systems experience more total SDWA violations. Balazs and Ray's (2014) Drinking Water Disparities Framework identifies actors and systems that perpetuate social inequalities in drinking water and addresses the challenges faced by small rural water systems. They argue that funding at the regional level is needed to support TMF capacity and help CWSs develop engineering and financial strategies for infrastructure improvements and SDWA compliance (Balazs & Lubell, 2014). Future studies should focus on how regional agencies or organizations can best and most efficiently support CWSs, and policy makers need to take seriously the need for external funding, particularly in small rural communities and/or low-income communities of color. Statewide studies could be beneficial in identifying potential disparities and the areas most in need of external support. This could assist in providing the most efficient allocation of funding to the areas in most need. Case studies of regions that have successfully reduced their violations could provide a beneficial blueprint for similar CWSs.

6. Conclusion

This study demonstrates that the spatial analysis method used in an environmental justice study can significantly affect the outcome and interpretations and offers new insights. For both total and health-based SDWA violations, the results varied regarding the impact of SES and sociodemographic variables across the four spatial analysis methods. Dasymetric mapping is the most robust method and best reflects the population's distribution (Holifield et al., 2017); it was thus used as the basis of comparison for other methods. While the county-level model was comparable to the dasymetric mapping model across many variables, it differed in the key environmental justice-related categories. Areal weighting was found to have results consistent with dasymetric mapping in terms of significance and coefficient direction, indicating it could be a decent alternative spatial analysis method if the more detailed data required for the dasymetric approach is not available. It is important to be aware of the potential effects of the spatial analysis method applied; community-level variations may be masked if county-level data is used.

This study did not find conclusive evidence that SDWA violations were associated with race or SES. The research did, however, identify a disparity in SDWA compliance in rural vs urban and small vs large communities. Rural CWSs in Pennsylvania are likely to have more total and health-based SDWA violations than their urban counterparts, a discrepancy which was also found to be true for health-based violations on a national scale (Allaire et al., 2018). It is possible that the fact that poorer communities of color are centered in urban areas in Pennsylvania is the reason that evidence of environmental injustice was not found.

Disparities in many states go beyond the rural vs urban divide, however, as other studies have found evidence of environmental injustice by race and SES with regard to drinking water quality (Balazs et al., 2011, 2012). Qualitative assessments of the recent events in Flint, Michigan also demonstrate the inequalities faced by poorer communities of color in drinking water quality, compliance and enforcement (Butler et al., 2016). Prior studies analyzed in conjunction with this one show that disparities in drinking water quality and compliance need to be addressed through research and increased funding at the local, state and national level.

Declaration of competing interest

None.

CRediT authorship contribution statement

Zoe Statman-Weil: Conceptualization, Methodology, Software, Validation, Formal analysis, Writing - original draft. **Leora Nanus:** Supervision, Conceptualization, Writing - original draft, Writing - review & editing. **Nancy Wilkinson:** Writing - review & editing.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.apgeog.2020.102264.

References

Allaire, M., Wu, H., & Lall, U. (2018). National trends in drinking water quality violations. *Proceedings of the National Academy of Sciences, 115*(9), 2078–2083.

Baden, B. M., Noonan, D. S., & Turaga, R. M. R. (2007). Scales of justice: Is there a geographic bias in environmental equity analysis? *Journal of Environmental Planning and Management, 50*(2), 163–185.

Balazs, C. L. (2011). *Just water? Social disparities and drinking water quality in California's san Joaquin Valley*. Berkeley: University of California.

Balazs, C. L., & Lubell, M. (2014). Social learning in an environmental justice context: A case study of integrated regional water management. *Water Policy, 16*, 97–120.

Balazs, C. L., Morello-Frosch, R., Hubbard, A. E., & Ray, I. (2011). Social disparities in nitrate-contaminated drinking water in California's San Joaquin Valley. *Environmental Health Perspectives, 119*(9), 1272–1278.

Balazs, C. L., Morello-Frosch, R., Hubbard, A. E., & Ray, I. (2012). Environmental justice implications of arsenic contamination in California's san Joaquin Valley: A cross-sectional, cluster-design examining exposure and compliance in community drinking water systems. *Environmental Health, 11*(84), 1–12.

Balazs, C. L., & Ray, I. (2014). The drinking water disparities framework: On the origins and persistence of inequities in exposure. *American Journal of Public Health, 104*(4), 603–611.

Butler, L. J., Scammell, M. K., & Benson, E. B. (2016). The Flint, Michigan, water crisis: A case study in regulatory failure and environmental injustice. *Environmental Justice, 9*(4), 93–97.

Campbell, C., Greenberg, R., Mankikar, D., & Ross, R. D. (2016). A case study of environmental injustice: The failure in Flint. *International Journal of Environmental Research and Public Health, 13*(10), 951.

Cedar Lake Ventures Inc. (2018). The demographic stastical atlas of the United States. Retrieved from https://statisticalatlas.com.

Chakraborty, J., Maantay, J. A., & Brender, J. D. (2011). Disproportionate proximity to environmental health hazards: Methods, models, and measurement. *American Journal of Public Health, 101*(S1), 27–36.

Cory, D. C., & Rahman, T. (2009). Environmental justice and enforcement of the safe drinking water act: The Arizona arsenic experience. *Ecological Economics, 68*(6), 1825–1837.

Environmental Working Group. (2017). State of American drinking water. Retrieved from https://www.ewg.org/tapwater/state-of-american-drinking-water.php# (WZoWZyiGNPY).

Fedinick, K. P., Wu, M., Pandithartne, M., & Olson, E. (2017). Threats on tap: Widespread violations highlight need for investment in water infrastructure and protections. Retrieved from https://www.nrdc.org/resources/threats-tap-widespread-violations-water-infrastructure.

Holifield, R., Chakraborty, J., & Walker, G. (Eds.). (2017). *The routledge handbook of environmental justice* (1st ed.). New York, New York: Routledge.

Konisky, D. M., & Teodoro, M. P. (2016). *When Governments Regulate Governments, 60*(3), 559–574.

Krivoruchko, K., Gribov, A., & Krause, E. (2011). Multivariate areal interpolation for continuous and count data. *Procedia Environmental Sciences, 3*, 14–19.

Maantay, J., & Maroko, A. (2009). Mapping urban risk: Flood hazards, race, & environmental justice in New York. *Applied Geography, 29*, 111–124. https://doi.org/10.1016/j.apgeog.2008.08.002.

McDonald, Y. J., & Jones, N. E. (2018). Drinking water violations and environmental justice in the United States, 2011-2015. *American Journal of Public Health, 108*(10), 1401–1407.

Mennis, J. (2003). Generating surface models of population using dasymetric mapping. *The Professional Geographer, 55*(1), 31–42.

National Research Council. (1997). *Safe water from every tap: Improving water service to small communities*. Washington DC: National Academy Press.

NCSS Documentation. (2018). NCSS LLC. Retrieved from https://www.ncss.com/software/ncss/ncss-documentation/.

Ogneva-Himmelberger, Y., & Huang, L. (2015). Spatial distribution of unconventional gas wells and human populations in the Marcellus Shale in the United States: Vulnerability analysis. *Applied Geography, 60*, 165–174. https://doi.org/10.1016/j.apgeog.2015.03.011.

Pennsylvania Department of Environmental Protection. (2017). Public water supplier's *(PWS) service areas* [shapefile]. Retrieved from http://data-padep-1.opendata.arcgis.com/datasets/ accessed 8/20/17.

Pennsylvania Department of Environmental Protection. (2017a). Public water Supplier's PWS service areas, full metadata. Retrieved from http://www.pasda.psu.edu/uci/FullMetadataDisplay.aspx?file=PublicWaterSupply2017_07.xml.

Pennsylvania Department of Environmental Protection. (2019). PA environmental justice areas. Retrieved from https://www.dep.pa.gov/PublicParticipation/OfficeofEnvironmentalJustice/Pages/PA-Environmental-Justice-Areas.aspx.

Schaider, L. A., Swetschinski, L., Campbell, C., & Rudel, R. A. (2019). Environmental justice and drinking water quality: Are there socioeconomic disparities in nitrate levels in U.S. Drinking water? *Environmental Health, 18*(3).

Switzer, D., & Teodoro, M. P. (2017). Class, race, ethnicity, and justice in safe drinking water compliance. *Social Science Quarterly*, 1–12.

The Environmental Justice Coalition for Water. (2005). *Thirsty for justice: A people's blueprint for California water*.

U.S. Census Bureau. (2010). *Urban and rural 2010 census summary file* [population count spreadsheet]. Retrieved from https://factfinder.census.gov/faces/tableservices/jsf/pages/productview.xhtml?pid=DEC_10_SF1_P2&prodType=table accessed 2/20/2018).

U.S. Census Bureau. (2016a). *2011-2015 American community survey 5-year estimates* [population count spreadsheets]. Retrieved from https://factfinder.census.gov/faces/nav/jsf/pages/download_center.xhtml accessed 11/10/2017).

U.S. Census Bureau. (2016b). *Urban areas* [shapefile]. Retrieved from https://www.census.gov/cgi-bin/geo/shapefiles/index.php?year=2016&layergroup=Urban+Areas; accessed 2/20/2018.

U.S. Census Bureau. (2017). TIGER state level geodatabase *(Pennsylvania)* [geodatabase]. Retrieved from https://www.census.gov/geo/maps-data/data/tiger-geodatabases.html accessed 11/10/2017).

U.S. Environmental Protection Agency. (2004). *Understanding the safe drinking water Act*.

U.S. Environmental Protection Agency. (2017a). Background on drinking water standards in the safe drinking water act (SDWA). Retrieved from https://www.epa.

gov/dwstandardsregulations/background-drinking-water-standards-safe-drinking-water-act-sdwa.
U.S. Environmental Protection Agency. (2017b). How EPA regulates contaminants. Retrieved from https://www.epa.gov/dwregdev/how-epa-regulates-drinking-water-contaminants.
U.S. Environmental Protection Agency. (2017c). Safe drinking water act (SDWA) resources and FAQ. Retrieved from https://echo.epa.gov/help/sdwa-faqs#Q5.
U.S. Environmental Protection Agency. (2018). Safe drinking water information system [violation data]. Retrieved from https://ofmpub.epa.gov/apex/sfdw/f?p=108:200 accessed 10/15/2018.
U.S. Environmental Protection Agency's Office of Enforcement and Compliance Assurance. (2013). Providing safe drinking Water in America: 2013 national public water systems. Retrieved from https://www.epa.gov/sites/production/files/2015-06/documents/sdwacom2013.pdf.
U.S. Geological Survey. (2011). *National land cover database* [raster]. Retrieved from https://www.mrlc.gov/nlcd11_leg.php accessed 2/11/2018.
U.S. Government Accountability Office. (2011). Drinking water: Unreliable state data limit EPA's ability to target enforcement priorities and communicate water systems' performance. Retrieved from https://www.gao.gov/assets/320/319780.pdf.
Wallsten, S., & Kosec, K. (2008). The effects of ownership and benchmark competition: An empirical analysis of U.S. water systems. *International Journal of Industrial Organization, 26*, 186–205. https://doi.org/10.1016/j.ijindorg.2006.11.001.