Environmental Challenges 3 (2021) 100033

Contents lists available at ScienceDirect

Environmental Challenges

journal homepage: www.elsevier.com/locate/envc

A data-driven model to evaluate the medium-term effect of contingent pricing policies on residential water demand

Taís Maria Nunes Carvalho*, Francisco de Assis de Souza Filho

Dept. of Hydraulic and Environmental Engineering, Universidade Federal do Ceará, Campus do Pici, Bloco 713, Fortaleza CEP 60455-760, Brazil

ARTICLE INFO

Keywords:
Water demand
Data-driven modeling
Contingent pricing
Drought planning

ABSTRACT

The relative scarcity of water resources has encouraged cities to create mechanisms to control water demand and avoid water stress. In the decision-making process, water companies need to assess the price influence on water demand predictions to design better policies. The aim of this study is to estimate the medium-term effectiveness of the implementation of a contingent tariff and its consequences for water demand elasticity to price. A novel model that requires only secondary data is proposed, that can be useful for guiding the drought planning process. The methodology consists of a framework that provides monthly predictions of water demand at the household level, considering price, seasonality, and previous water use. The results indicated that the contingent tariff promoted a reduction of 11–17% in water demand, but at a higher cost for low income households. Also, reduction in water demand was found to be inelastic to price increase. Using google search hits as a proxy for public interest, we found that water cost has a higher influence on users' decision to save water than drought awareness.

1. Introduction

The growing water demand associated with urbanization processes has increased water stress and the risk of shortage in several regions of the world (McDonald et al., 2014). For some of them, the elevated temporal and spatial variability in water availability offer an additional challenge to water supply management (Orlowsky and Seneviratne, 2012; Pal et al., 2013; Campos et al., 2014).

In this context, water companies and policymakers have been implementing demand control measures, since increasing water supply capacity is not always possible or effective (Romano et al., 2014; Whittington and Nauges, 2020). A widely used approach is the adoption of increasing block rates (IBR), which is expected to encourage rational water consumption (Rietveld et al., 2000; Zhang et al., 2017). This kind of policy is typical of regions affected by droughts and developing countries and has complex impacts on consumer behavior (Rinaudo et al., 2012). Pricing strategies might also include tariffs that vary seasonally with temperature and/or precipitation (Pesic et al., 2012; Molinos-Senante, 2014) or adjusted with the level of water storage (Chu and Grafton, 2019), and household size (Arbués and Barberán, 2012).

A less common strategy to reduce water use under drought conditions is the implementation of penalty fees for those households with an elevated consumption (García-Rubio et al., 2015; Braga and Kelman, 2020). In Brazil, water utility companies have used this approach to deal with water crisis (Braga and Kelman, 2020). In Fortaleza, located in northeast Brazil, water pricing follows an IBR structure, and a contingent tariff, i.e., a penalty fee, was adopted three years after the beginning of a severe drought that reduced reservoir storage by about 63% (Pontes Filho et al., 2020). This tariff was influenced by the consumption quantity that exceeded a predefined threshold.

Previous studies have reported that water scarcity impacts price elasticity, but the consequences are adverse. While early research indicated that price elasticity is more significantly affected by pricing structure and season (Espey et al., 1997), recent studies show that consumers response to price change is related to different exogenous factors, such as climate (Monteiro and Roseta-Palma, 2011), income (Ma et al., 2014) and environmental attitude (Garrone et al., 2019). Dalhuisen et al. (2003) pointed out that income elasticities are relatively inelastic under IBR pricing, and that water scarcity does not seem to affect elasticity. Molinos-Senante and Donoso (2016) proposed a tariff scheme that accounts for the scarcity value of water and that can promote equity, based in a IBR structure and cross-subsidy. However, the measure might be difficult to implement due to lack of adequate water metering. Another strategy aiming equity and sustainability was presented by Ward and Pulido-Velázquez (2008), that presented a two-tiered pricing setup. Debate continues about the effectiveness of price control policies for demand control, especially on IBR schemes (Mansur and Olmstead, 2012; Zhang et al., 2017; Matikinca et al., 2020).

The research to date has extensively explored the price influence on water consumption (Arbués et al., 2004; Olmstead et al., 2007; Ward and Pulido-Velázquez, 2008) – together with other socioeconomic and/or

* Corresponding author.
E-mail address: taismarianc@gmail.com (T.M. Nunes Carvalho).

https://doi.org/10.1016/j.envc.2021.100033
Received 24 November 2020; Received in revised form 20 January 2021; Accepted 21 January 2021

climatic variables – but only a few studies are able to address it over long time horizons (Grafton et al., 2014). Most studies on water price use survey (which can be expensive and time consuming), aggregate, or household level data to assess the empirical implications of economic variables on water demand (Rujis, 2009). Although these analyses have improved the understanding of the scientific community and decision-makers, they do not allow continuous learning as new data become available.

Water companies have a huge amount of smart meter data available that could be useful to extract information on use patterns and consumer behavior (Cominola et al., 2019). In this research, we present a method that benefits from this data to support managers on how to adjust the pricing policy for a planning horizon of up to one year. The model can be coupled with reservoir/supply systems operation or water distribution models to provide further insights on supply-demand balance strategies.

This study proposes a data-driven predictive model to assess the medium-term effect of price-based water conservation policies at the household level. In addition, we calculate the elasticity of water demand reduction to price and we assess how much water price and public interest in the drought can affect consumption habits. The methodology can be used by water companies to assess price-related strategies of water conservation and does not require additional variables that could be difficult to obtain in a refined scale. An advantage of this model is that prediction can be performed at a disaggregated level, making it possible to design policies tailored to socioeconomic or even structural characteristics of the households. Although this study considers a block tariff structure, the framework can be adapted to any other price strategy, if it is applied at the household level.

2. Materials and methods

2.1. Study area

The city of Fortaleza, capital of Ceará, located in the Northeast region of Brazil, is the fifth most populated city of the country, with over 2.6 million inhabitants distributed across 314.9 km^2. The population is expected to grow to 3.1 million people in 2040 (Iplanfor, 2015). The city is part of the Metropolitan Region of Fortaleza, which comprises 19 municipalities of Ceará.

Fortaleza is supplied by the Jaguaribe-Metropolitano supply system (JMS), which consists of eight reservoirs which sum up to a storage capacity of 11,112 hm^3. JMS transfers water from the Jaguaribe basin and supplies 36 municipalities. Urban and industrial demand of Fortaleza is 6.77 m^3/s, corresponding to 56.5% of the volume released by the supply system. Past research has indicated that water demand in Fortaleza is highly heterogeneous and that socioeconomic factors play an important role on consumption habits (Nunes Carvalho et al. 2021).

2.2. Water tariff structures

During the period between 2012 and 2018, the northeast of Brazil suffered from a historic drought that significantly impacted its economy and water storage (Pontes Filho et al., 2020). The main reservoirs of Fortaleza's supply system were affected by the 2012–2018 drought, resulting in a significant reduction in water availability. To encourage domestic water conservation, which accounts for more than 80% of Fortaleza's water demand, the local water company implemented a contingent tariff.

The contingent tariff was implemented in December 2015 (Fig. 1) and defined a minimum reduction of 20% of the average consumption between October 2014 and September 2015. If a household did not meet this reduction goal, an extra charge of 110% on the exceeded volume would be added to the bill, i.e., the contingent tariff is calculated on the difference between the volume consumed and the goal. This percentage was updated to 120% in October 2016. Water price follows an increasing block tariff structure (Table 1); thus, the contingent tariff also varies with the consumption block of the household. Users with a monthly consumption of up to 10 m^3 did not have to pay the contingent tariff.

Table 1
Water tariff in Fortaleza for each consumption category for the years of 2016 and 2017.

Monthly consumption (m^3)	2016 (BRL)	2017 (BRL)
0 to 10	2.79	3.48
11 to 15	3.61	4.51
16 to 20	3.92	4.88
21 to 50	6.71	8.36

For example, for a household that had a mean consumption of 14 m^3 between October 2014 and September 2015, the goal was to use up to 11 m^3, corresponding to a 20% reduction in the monthly consumption. If in a certain month of 2017 the water demand of this household was 13 m^3, in addition to the water tariff (13×4.51), they would have to pay the contingent tariff, which would be charged over the 2 m^3 that exceeded the consumption goal (1.2 × 2*4.51). The base price here corresponds to the second block of consumption (R$ 4.51 in 2017).

Although we consider these specific conditions in the prediction model, the methodology could be replicated under different price-associated water conservation measures..

2.3. Predictive model

The predictive model has three explanatory variables: previous water demand, monthly seasonality of water demand, and the cost of the penalty fee, i.e., the contingent tariff cost per household. The model was tested for multiple leading times, ranging from one to twelve months.

The penalty fee was calculated as the cost of the volume of water consumed in the previous month that exceeded a threshold. This threshold sets how much water should be saved and is a percentage of the average monthly water consumption of the household for a baseline period. Here, the baseline period goes from October 2014 to September 2015 and the threshold is 20%.

At each month, the predictions for the previous month are used to determine the tariff block of each household. Then, we calculate the volume of consumed water that exceeded the threshold and how much it cost for the user. For example, when calculating water consumption at n-months ahead, the predictions for the month n-1 are used to assess the water conservation measure (Fig. 2). This strategy allowed us to avoid the simultaneity issue associated with water consumption modeling under block tariff policies.

Previous studies have used different price variables in econometric models of water demand, and there is not a generally accepted approach. Many authors find it more appropriate using the marginal price, i.e., the cost of increasing the water consumption at each time step (Rinaudo et al., 2012), while others prefer the average price (Zhang et al., 2017) or both (Ma et al., 2014; Deyà-Tortella et al., 2016). Although some researchers argue that the users might be more influenced by the average price (Deyà-Tortella et al., 2016), in case of a contingent tariff policy, they might pay special attention to the additional charge expressed on the bill.

In addition to the lagged water consumption and the price component, a seasonal variable was included to account for seasonal behavior. This variable corresponded to the seasonal component extracted for each household with the Seasonal and Trend decomposition using Locally estimated scatterplot smoothing (STL) method. This approach captures different patterns of seasonal behavior and adds more information to the model than the usual approach of using 11 dummy variables for the months. We chose a machine learning regression model that has been widely used for electricity and wind prediction, Gradient boost regression. This algorithm also performs better than other linear and machine learning models in predicting residential water demand (Lee and Derrible 2020).

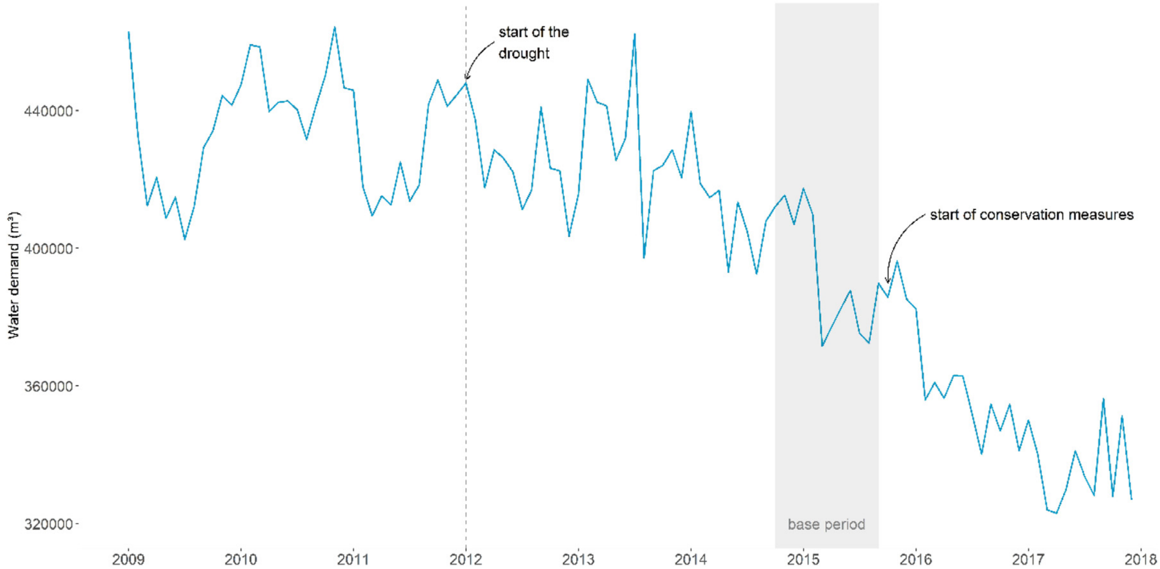

Fig. 1. Total domestic water demand (m^3) in Fortaleza from 2009 to 2017. The baseline period was used by the local water company to calculate the reduction goal for each household.

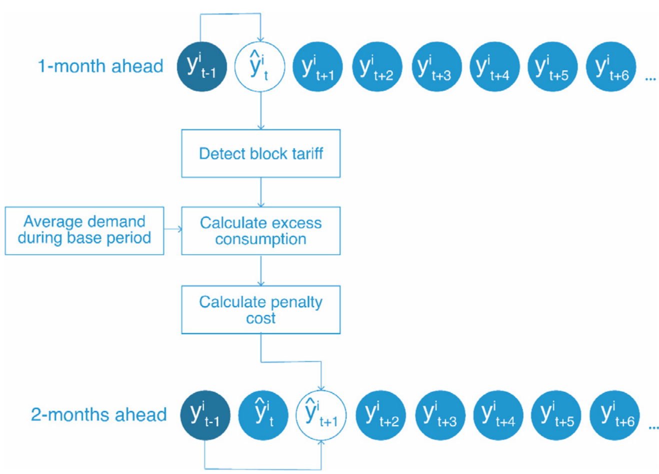

Fig. 2. The predictive model has an autoregressive component (previous month water demand) and the penalty fee as explanatory variables, in addition to the seasonality of the corresponding month. Starting from January, the water demand in December would be used to calculate the cost of the contingent tariff. For the next month, the penalty cost is calculated using the predicted water demand in January.

The predictive model can be summarized in the following steps:

(i) Select a dataset χ of monthly household water demand and set a time horizon n (in months) for the predictive model.
(ii) Extract the seasonal component s^i of each household's water demand time series y^i using the STL method.
(iii) Set a consumption reduction goal or threshold and the penalty cost policy $p(.)$. The goal might be a percentage of the average consumption over a certain period, named the baseline consumption b^i.
(iv) Split the dataset into two subsets for training and validating the model. Initialize the gradient boosting model at month $t = 1$, setting the predictive variable $\hat{y}$ to y^i_t and the predictors to s^i, $p(y^i_{t-1}, b^i)$, and y^i_{t-1}. Choose arbitrary values for the main parameters of the model i.e., the number of trees, the minimum number of observations in each node and the learning rate (usually ranges from 0.001 to 0.1).
(v) Run the model again using the predicted water demand $\hat{y}^i_t$ to calculate the penalty cost and estimate $\hat{y}$ at month $t + 1$. If the water tariff follows an IBR structure, it might be necessary use a function $f(\hat{y}^i_t)$ to set the tariff block to the household prior to calculating the penalty cost. Repeat this procedure until $t = p$.
(vi) Compute model's performance $D(\hat{y}, y^i_t)$ on the training and testing sets and compare the measures to adjust the parameters and avoid overfitting the model.

The tabular version of the algorithm is described below:
Initialize: Set the variable $\hat{y}$ equal to y^i_t.

- Calculate the baseline consumption and the reduction goal
- *Decompose the water demand time series using STL and extract its seasonal component* s^i

repeat

- *Determine the tariff block of each household based on the consumption of the previous month using a function* $f(y^i_{t-1})$. *This step can be ignored if the water tariff does not follow an IBR structure.*
- *Calculate the penalty cost using a function* $p(y^i_{t-1}, b^i)$
- *Estimate a gradient boosting regression model that predicts y using* s^i, $p(y^i_{t-1}, b^i)$, *and* y^i_{t-1} *as predictors*
- *Compute model's performance using the selected measure(s)* $D(\hat{y}, y^i_t)$ **until** $t = n$
- *Adjust model's parameters based on the performance of the training and testing subsets.*

The model was validated with a classical out-of-sample evaluation and was trained for the year of 2016 and tested for the year of 2017. Fig. 3 provides a general outline of the predictive model and the performed analysis.

2.3.1. Seasonality extraction

The water demand time series was decomposed into trend, seasonal and remainder components using the STL method (Cleveland and Cleveland 1990). This procedure was used to extract the seasonality of water consumption for each household. STL consists in sequential applications of the local regression model and provides an additive decomposition of the original signal (D) into three components:

$$D(t) = S(t) + T(t) + R(t) \tag{1}$$

where S, T and R are the seasonal, trend and remainder components, respectively. The algorithm work as follows:

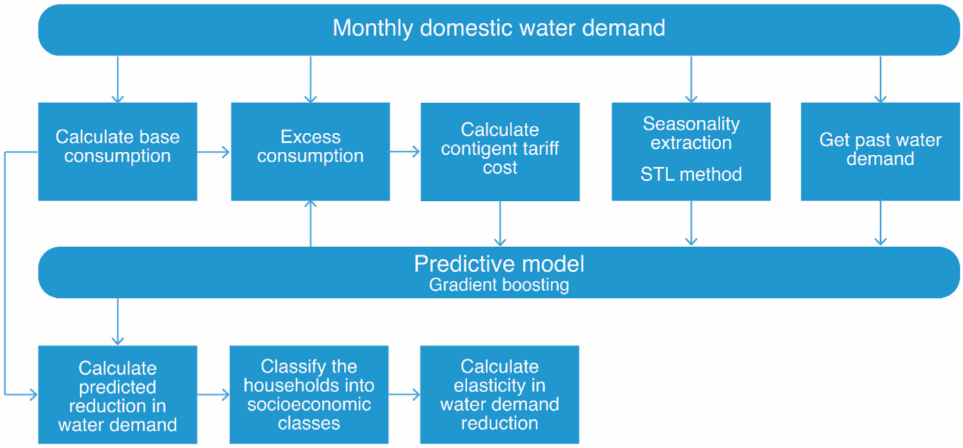

Fig. 3. Predictive model outline. The contingent tariff cost is recalculated as new predictions become available.

The local regression smoothing estimates a function g(x) for the independent variable at any value of x rather than for the measurements x_i of the dependent variable. To calculate the regression curve g, an initial value for the parameter q is chosen; q values of x_i that are closest to x are selected and weighted on their distance from x. For $q \leq n$, where n is the number of observations in the data set, the neighborhood weight for x_i is calculated as follows:

$$v_i(x) = W\left(\frac{|x_i - x|}{\lambda_q(x)}\right) \tag{2}$$

where $v_i(x)$ is the neighborhood weight for x_i, $\lambda_q(x)$ is the distance between x and the most distant x_i. For $q > n$, $\lambda_q(x)$ is multiplied by q/n. W(u) is the tricube function, expressed as:

$$W(u) = \{ \begin{matrix} \left(1-u^3\right)^3 \; if \; 0 \leq u < 1 \\ 0 \; if \; u \geq 1 \end{matrix} \tag{3}$$

Next, a polynomial of degree d is fit to the weighted data at (x_i, y_i). The value of d can be 0 (constant), 1 (locally linear) or 2 (locally quadratic). In this paper, $d = 1$. The fitted function corresponds to g(x). It is possible to add a robustness weight ρ_i for each pair (x_i, y_i) by multiplying it by v_i.

STL consists of two nested loops (Cleveland and Cleveland 1990). In the outer loop, robustness weights are calculated for each time point. Initially, the trend and remainder component are set to 0 and ρ_i is set to 1. In the next loops, the remainder component is found by removing the trend and seasonal components calculated in the inner loop from the original series. The robustness weight is then calculated as follows:

$$\rho_i = B(|R|/h) \tag{4}$$

$$h = 6 * median(|R|) \tag{5}$$

where B is the bi-square weight function, given as:

$$B(u) = \{ \begin{matrix} \left(1-u^2\right)^2 \; if \; 0 \leq u < 1 \\ 0 \; if \; u > 1 \end{matrix} \tag{6}$$

The outer loop is repeated n_o times; if one does not wish to add robustness into STL, n_o should be set to 0. In this paper, $n_o = 15$. The inner loop follows these steps: (i) Detrend the original signal; (ii) Estimate a smoothing function using Loess for each cycle-subseries, where q is the cycle periodicity (e.g. for a monthly time series, q is set to 12) and d is equal to 1; (iii) Apply a low pass filter to the smoothed cycle-subseries, which consists in sequential applications of a moving average; (iv) Detrend the smoothed cycle-subseries; (v) Remove the seasonality from the series; (vi) Smooth the deseasonalized series using Loess. The STL decomposition can be easily performed using the stl function from base R.

2.3.2. Gradient boosting

Gradient Boosting (GBM; Friedman 1999) is a learning method that converts weak learners, usually regression trees, into strong learners by combining them sequentially. The idea behind the method is that new weak learners can learn from the residuals of the output from the previous model; this ensemble technique is called bagging. For regression tasks, we want to find the function that best fits the data points in a set containing input variables x and a corresponding output variable y. To do this, the algorithm minimizes a loss function between y and the predicted values, in our case, the Mean Squared Error.

$$L(y, \hat{y}) = \frac{1}{N}\sum_{i=1}^{N}\left(y_i - \hat{y}_i\right)^2 \tag{7}$$

The gradient boosting method consists in a combination of weak learners that are added together. The individual models f_m are added one after the other to improve model performance.

$$\hat{y}_i = \sum_{m=1}^{M} f_m(x) \tag{8}$$

The weak learners, in this case, regression trees, are fitted on the residuals of the previous model. The general representation of GBM is expressed as follows:

$$F_m(x) = F_{m-1}(x) + \nu f_m(x), \tag{9}$$

meaning that the model f_m does not change the previously fitted model $F_{(m-1)}$. The term ν is a regularization parameter or the learning rate, which determines the number of iterations. Small values of the learning rate ($\nu < 0.1$) reduce the chances of overfitting.

Gradient boosting applies a functional gradient descent method to minimize the loss function, where each new weak model is equivalent to the negative gradient of the MSE. The negative gradient is given as:

$$-g_m(x_i) = -\left[\frac{\partial L(y_i, F(x_i))}{\partial F(x_i)}\right]_{F(x)=F_{m-1}(x)} \tag{10}$$

The algorithm stops when the loss reaches a threshold, or the maximum number of trees is built. An important element to consider when fitting machine learning models (or any predictive model) is the bias-variance tradeoff and the chance of overfitting the model. If the algorithm misses important connections between the predictors and the response variable, the model will have a high bias, i.e., an elevated difference between predictions and the observed data. However, if during the model fits too perfectly to the training data, resulting in a high variance, it will not generalize well (overfit).

The best scenario when developing a model is to accurately capture the relationships between the variables during training but also make good predictions during training. In machine learning models, one can

Table 2
Socioeconomic classes and number of households in each of them. The total number of household analyzed here is 37,689.

Class	Number of minimum wages	Number of households	Percentage of the total number of households
A	20 or more	53	0.14%
B	$10 < N < 20$	969	2.57%
C	$4 < N < 10$	5186	13.76%
D	$2 < N < 4$	17,554	46.58%
E	$N < 2$	13,927	36.95%

control the bias-variance tradeoff by controlling model parameters. The main parameters of GBM are the number of trees, which should not be too high to avoid overfitting; the minimum number of observations in each node, which defines how depth the tree might become; the learning rate or shrinkage, which relates to the size of the incremental steps, usually ranging from 0.01 to 0.1, and the distribution of the response variable, which in our case, was Gaussian.

In our framework, parameter tuning was performed in a trial and error manner, i.e., we defined arbitrary values for them, compared model performance for the train and test datasets and chose those parameters that resulted in comparable performances for both and could not be improved anymore. The number of trees was set to 300, the learning rate to 0.1, and the number of observations per node to 10. All analyses were performed using R programming language. The gradient boosting model was implemented with the package gbm (Greenwell et al., 2019).

2.3.3. Performance assessment

Model performance was evaluated for the entire prediction horizon, i.e., for twelve months of the testing period. Two measures were used: Root Mean Squared Error (RMSE) and R squared (R^2).

$$RMSE_j = \sqrt{\sum_{i=1}^{n} \frac{\left(\hat{y}_{i,j} - y_{i,j}\right)^2}{n}} \tag{11}$$

$$R_j^2 = \frac{\sum_{i=1}^{n} \left(y_{i,j} - \hat{y}_{i,j}\right)^2}{\sum_{i=1}^{n} \left(y_{i,j} - \bar{y}_j\right)^2} \tag{12}$$

where $y_{i,j}$ is the observed water demand in household i at month j, $\hat{y}_{i,j}$ is the predicted water demand in household i at month j, $\bar{y}_{i,j}$ is the mean observed water demand at month j, and n is the number of households.

2.4. Elasticity of water demand reduction to price

Different scenarios of price increase were considered, based on the tariff for the previous year (2015 for the training and 2016 for the validation period): no increase, 5, 10, 15 and 25%. To calculate the elasticity of water demand reduction to price, we used the predictions for the year of 2016 obtained with the model. The reduction is related to the average consumption during the baseline period (October 2014 to September 2015).

$$E = \frac{\Delta R / R}{\Delta P / P} \tag{13}$$

where R is the monthly average reduction in water demand and P is the average water block tariff.

Water demand elasticity was assessed for different socioeconomic classes, as users' response to water conservation policies tend be heterogeneous. These classes were based on the criteria used by the Brazilian Institute of Geography and Statistics (IBGE), which is based on per capita family income. IBGE uses the minimum wage to classify the families in five classes (Table 2). The monthly per capita income of a household is divided by the minimum wage to find the correspondent socioeconomic class (i.e., Income = N*minimum wage). We also compared the predicted monthly reduction with the actual reduction aggregated water demand.

2.5. Public interest and media coverage

In this study, water demand reduction is associated with the implementation of a price control measure, which was expected to change public behavior. However, demand control policies may include other strategies, such as promotional events, water conservation education programs and mass media advertising campaigns (Sharma and Vairavamoorthy 2009). In Fortaleza, the water company created an app to encourage users to report leaks and frauds and promoted educational campaigns in schools, public buildings, and social media.

Google Trends data has been proven a useful tool for characterizing public response to certain matters and has been successfully applied to analyze private consumption (Vosen and Schmidt 2011) and to assess drought awareness (Quesnel and Ajami 2017; Kam et al., 2019). The idea here was to use the frequency of Google searches for the key words "contingent tariff" and "drought" to address people's interest on these matters and their awareness about the implementation of the tariff.

We acknowledge that mass media plays an important role on social systems (Luhmann 2000), hence media coverage on the contingent tariff might have influenced public response. For reference, we plotted the number of articles related to the contingent tariff published between 2012 and 2017, which were collected from the websites of the three main local newspapers (Tribuna do Ceará, OPovo and Diário do Nordeste). These sources have a strong online presence and usually share the news on social media such as Instagram and Twitter. Data was collected with web scraping using Python and the BeautifulSoup 4 library.

To assess the marginal response and the relative influence of public interest in drought and the contingent tariff on water demand, a regression analysis was performed using both as explanatory variables (Fig. 4). Water demand was predicted as a function of water demand in the previous month, public interest, and the contingent

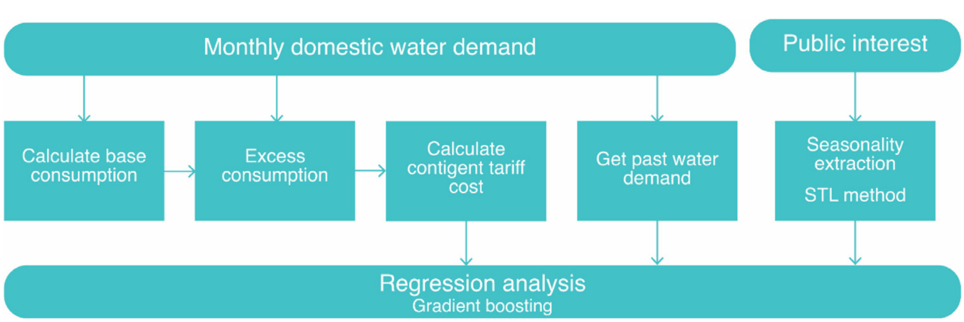

Fig. 4. Regression analysis outline.

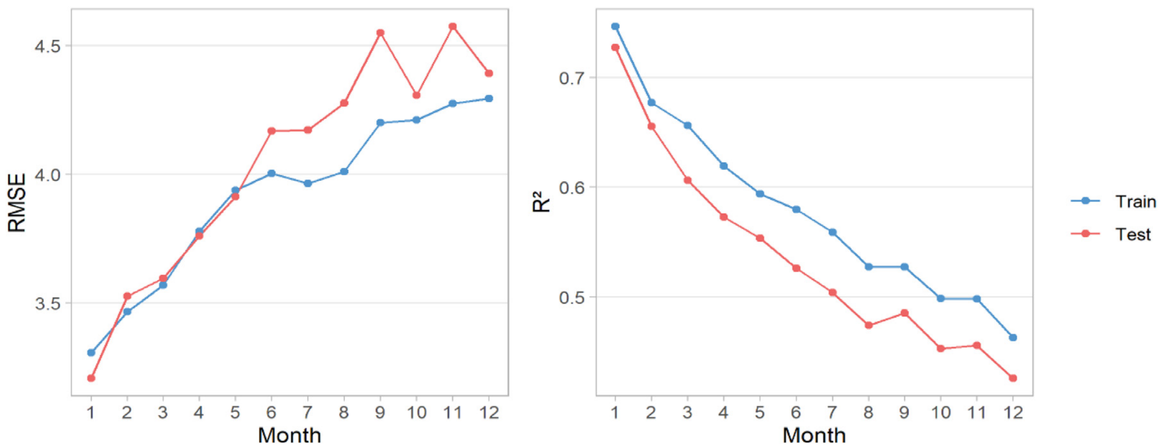

Fig. 5. Model performance.

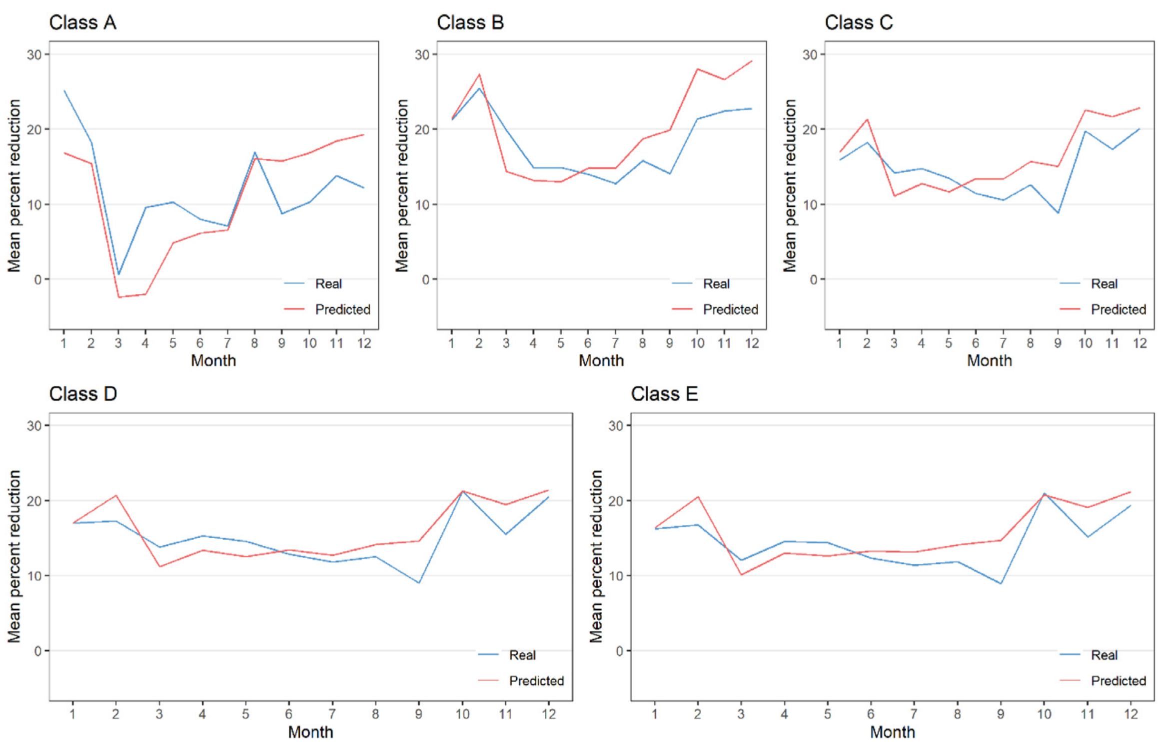

Fig. 6. Real and predicted monthly reduction in aggregated water demand for the year of 2017 for each socioeconomic class.

tariff cost for the previous month. Google search hits between 2012 and 2017 for the term "contingent tariff" and "drought" by users located in Fortaleza were used as a proxy for public interest in water scarcity, from which the trend component was extracted using the STL method.

The GBM algorithm was used to perform the regression. For this analysis, we used data from 2012 (beginning of the drought) to 2017. Note that here we fit the model using only observed data, i.e., the contingent cost is not iteratively calculated, since our intention was not to build a forecast but rather to assess the importance of the explanatory variables. For the same reason, seasonal water demand was not added as a predictor. The dataset was randomly split into 80% train and 20% test. After obtaining the regression model, we extracted the marginal response of each variable using partial dependence plots and their relative influence. The relative influence is measured with the reduction of squared error associated with each variable, i.e., how much worse the model's performance would be without that variable.

2.5.1. Partial dependence plot

The partial dependence plot (PDP) represents the marginal effect of independent variables on the response of a machine learning model (Friedman 1999). The partial dependence of the response on a variable x_l is represented by:

$$\hat{f}_{x_l}(x_l) = E_{x_s}\left[\hat{f}(x_l, x_s)\right] = \int \hat{f}(x_l, x_s) P(x_s) dx_s \tag{14}$$

Where x_l is the independent variable analyzed in the partial dependence plot, x_s is the subset of the other input variables of the regression model $\hat{f}$ and $P(x_s)$ is the marginal probability density of x_s. The function shows the effect of the variable x_l on the dependent variable by marginalizing over the other explanatory variables.

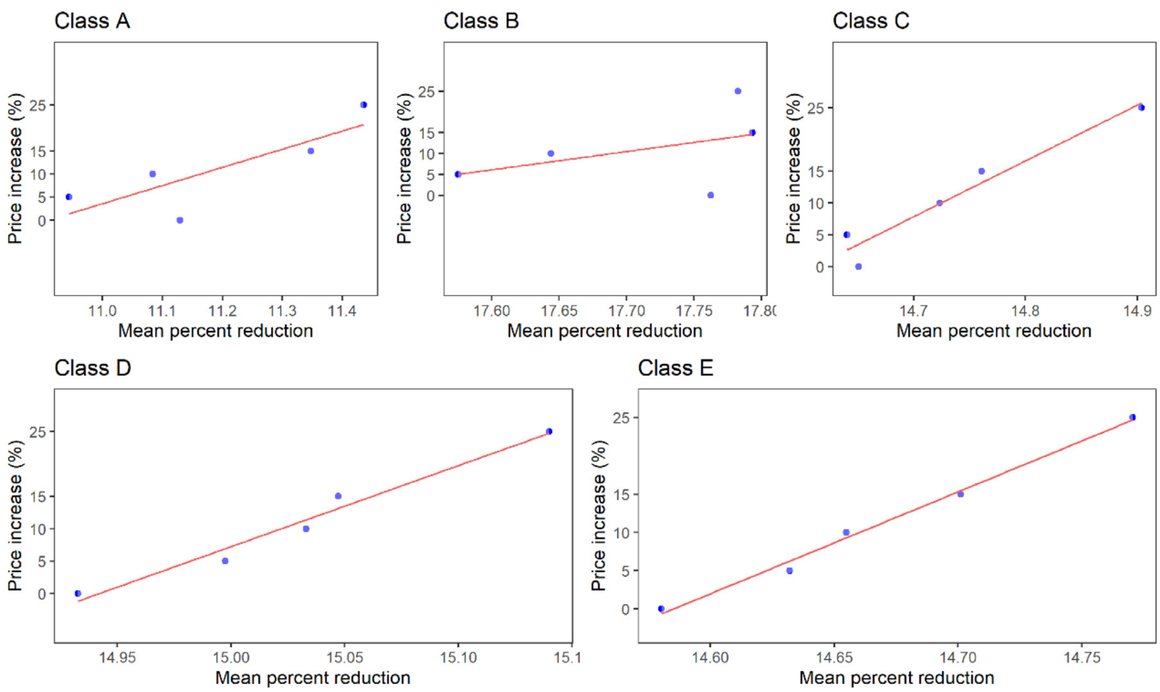

Fig. 7. Elasticity of water demand reduction to price for each socioeconomic class.

Table 3
Reduction in water demand elasticity to price increase and characteristics of the socioeconomic classes.

Socioeconomic class	A	B	C	D	E
Elasticity of water demand reduction to price	0.515	0.212	0.426	0.314	0.295
Number of households	53	969	5186	17,554	13,927
Percentage of the average per capita income related to the water tariff (%)	0.46	0.89	2.00	3.77	22.97
Average daily per capita consumption (L/hab/day) for the baseline period	102.25	123.36	105.19	96.78	94.96
Average daily per capita consumption (L/hab/day) after the restriction measures	90.06	104.99	92.37	84.60	83.69

3. Data

Monthly water demand data for the period between 2009 and 2017 from 45,141 households were provided by the Water and Wastewater Company of Ceará (CAGECE). This analysis focused on households with consumption up to 50 m^3/month. Households with monthly water consumption inferior to 1 m^3 per month or the ones in which the total water consumption between 2009 and 2017 was less than 5 m^3 were excluded from the dataset. The data cleaning process reduced the dataset to 37,689 observations.

Socioeconomic data from the 2010 Census were used to classify the households. Average per capita income is available at the census tract level, territorial units containing a maximum number of households that allow a survey to be carried out by a single person (IBGE). Fortaleza is divided into 3043 census tracts, and 2586 of them are attended by CAGECE's water supply.

4. Results and discussion

Model performance was evaluated for each month of the testing period (Fig. 5). The model presented reliable predictions in terms of RMSE and R^2 for a short-term horizon (1 to 6 months ahead), and satisfactory results for a medium-term horizon (7 to 12 months ahead). The autoregressive component was the most important, i.e., removing it from the model would mean a significant increase in the loss function. This suggests that water demand is strongly dependent on past use.

A comparison between the predicted and observed mean percent reduction in residential water demand shows that the model provided accurate predictions (Fig. 6). For this analysis, households were grouped according to their socioeconomic class, to assess variation in model performance and mean percent reduction in water demand. Classes D and E presented a rather regular behavior during the year, with an average reduction of 14.73% and 13.99%, respectively. Households in class B had the largest reduction in water demand: 17.58% over the year. Class A, with the smallest reduction (11.22% on average), presented a peak in January but almost no change in March.

The reduction in water demand was revealed inelastic to tariff variation (Fig. 7). These results suggest that the contingent tariff itself would be enough to encourage a reduction in water consumption in all socioeconomic classes. However, the policy has adverse effects on each type of consumer. While the water tariff represents less than 1% of the average per capita income of classes A and B, it is about 23% of the income of class E, which represents 37% of the households (Table 3). The lower income classes had the lowest per capita consumptions during the baseline period, but still managed to reduce their demand after the implementation of the contingent tariff. Except for households in class B, none of the classes would reach the 20% reduction goal. Class B also had the highest average daily per capita consumption (Table 3) during the baseline period.

These findings agree with other studies that also found water demand is inelastic to price variation (Rinaudo et al., 2012; Deyà-Tortella et al., 2016). Also, Zhang et al. (2017) showed that increasing block policies are not effective to encourage a reduc-

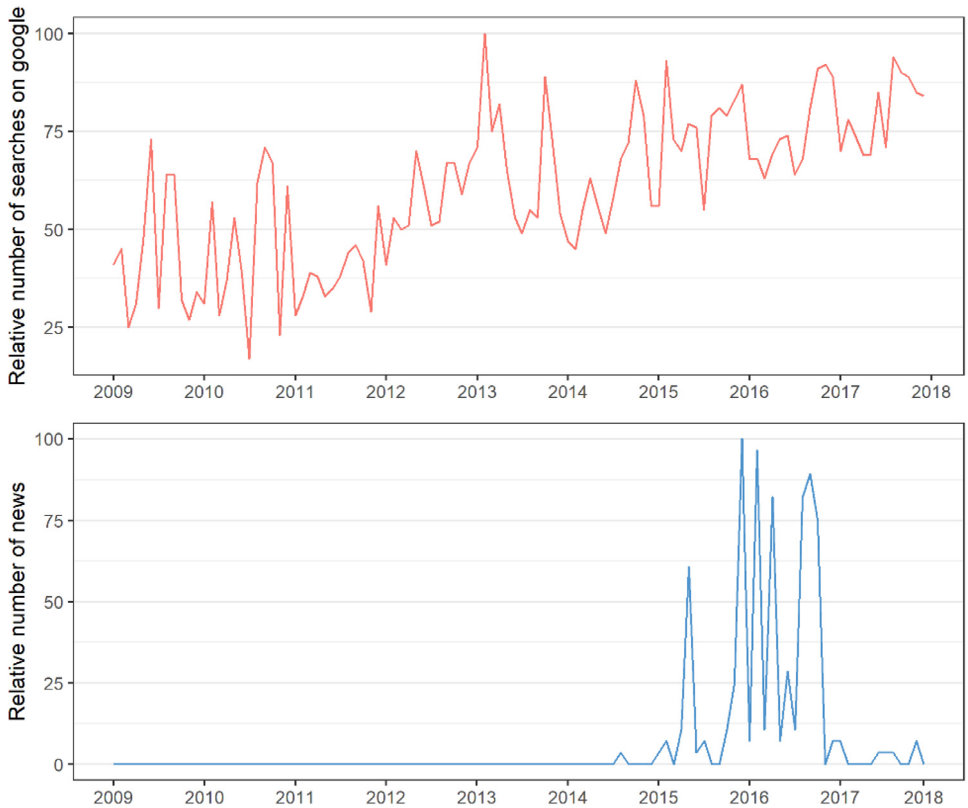

Fig. 8. Public interest and media coverage on the contingent tariff policy.

tion in water consumption. Ma et al. (2014) indicated that the highest income group is not sensitive to price changes, while residents from the lower income group respond to marginal price and might even compare the tariff for different blocks to optimize their benefit. De Maria André and Carvalho (2014) found similar values of water demand elasticity to price in Fortaleza using survey collected data. The advantage here is that we used only secondary data to calculate elasticity for different socioeconomic classes.

Overall, the results indicate that the restriction policy might be unfair with the lower income classes, for which the tariff represents a significant percentage of their income and still enforced a reduction in its already low daily per capita demand. As stated by Bernoulli (1954), benefit perception depends on the individual perception of cost. Hence, a small increase in water cost has a more significant effect on the economic value attributed to water for lower income classes.

In a scenario where the customers must pay an additional charge for their excess consumption, price increase does not seem to affect consumer behavior. This result can be explained by the fact that the customers might be at the kink point of the block rate schedule or their willingness to pay for water rises under drought conditions, since it represents only a small percentage of their income. The first is the most reasonable explanation for classes D and E, while the second is consistent with higher income classes. Another aspect to be considered is the reservation capacity of households (water tanks or cisterns, private borehole drilling), which is higher for wealthy customers (Grande et al., 2016), who might be able to maintain their standards and still reduce the water volume from public supply.

It is important to bear in mind that the consumers are not necessarily aware of the pricing policy structure. Although the contingent tariff is clearly expressed on the water bill, increasing block tariff scheme is not detailed for households.

A clear increasing trend in public interest is observed after 2012 (when the drought started), while the number of news related to the restriction measure peaked in 2016 (Fig. 8). While this could imply that the public was well informed about pricing policy, the finding cannot be extrapolated to all customers, since not all households have access to internet.

Table 4
Relative importance of the explanatory variables of the regression model between water demand, past water demand, public interest, and contingent tariff.

Class	A	B	C	D	E
Past water demand	85.87	95.78	98.31	98.67	98.55
Contingent tariff cost	10.17	3.90	1.59	1.29	1.42
Seasonal public interest	3.96	0.32	0.10	0.03	0.02
R^2	0.69	0.69	0.75	0.74	0.73
RMSE	4.37	4.08	3.26	3.01	3.05

A regression analysis between water demand, public interest, the contingent tariff, and past water demand was performed for each socioeconomic class (Table 4). The relative importance values imply that an increase in the cost associated with the contingent tariff has a higher influence on consumer behavior than information on drought. Also, it seems that residents with higher income have a more significant response to both the contingent tariff and information on drought compared to residents in classes with lower income.

PDPs were plotted for each regression model (Fig. 9). The results indicate that water cost has an inverse relationship with water demand for all households, while an increase in the interest in the drought has little effect on consumer habits. It is worth mentioning that class A is the only one to present a direct relationship between public interest in the drought and water demand. However, we should be careful when interpreting these results since class A has a low number of households.

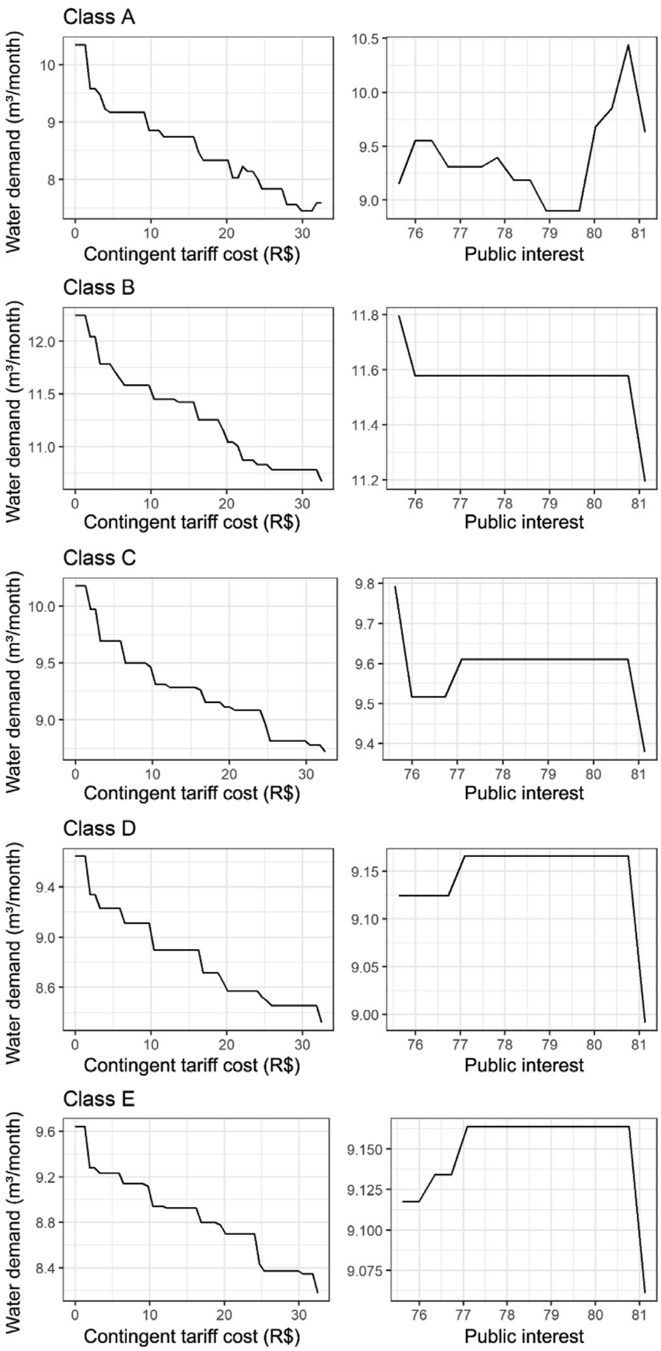

Fig. 9. Partial dependence plots for public interest and the contingent tariff cost. A regression model was built for each socioeconomic class. Public interest is dimensionless.

Conclusions

The main objective of this research was to address the influence of a contingent tariff on a predictive model of water demand in Fortaleza, Brazil. The model contained an autoregressive component and variables assessing seasonality and the cost associated with the contingent tariff. This study has found that the contingent tariff was effective and resulted in a 11–17% reduction in residential water demand. Also, reduction in consumption was inelastic to price increase in all socioeconomic classes.

The evidence from this study suggests that a price policy that associates IBT with a contingent tariff could be unfair to lower income households, for which the tariff represents a large percentage of household income. Hence, although the strategy warrants a high revenue for the water company (that can be allocated to water security projects), its equity is questionable. Managers should be careful when implementing pricing policies to ensure the affordability of water services to all consumers.

The findings of this study imply that price-related water demand control policies are effective, while drought awareness is less likely to encourage consumers to save water. The increase in public interest in the drought does not necessarily indicate that consumers are well informed about the risks associated with it. It is crucial that the users are aware of the water resources management strategies and the implications of their habits rather than having a limited perception of drought. This can only be accomplished if social dynamics aspects are considered when designing drought plans and policies.

The framework proposed here is flexible and can be useful for water companies planning to implement price-related measures to encourage water demand reduction. The predictions at the household level can be useful to design policies for different classes of consumers. The predictive model can be used to verify at what extent the changes in the price policy could influence water demand.

Declaration of Competing Interest

None.

Acknowledgments

This work was supported by the Brazilian Federal Agency for Support and Evaluation of Graduate Education [88887.123932/2015-00]; the Brazilian Council for Scientific and Technological Development [441457/2017-7]; and the Cearense Foundation for Scientific and Technological Support [Cientista-chefe program].

Data and Code Availability Statement

Data for this research was obtained from the Water and Wastewater Company of Ceará (CAGECE) and can be requested at this website: https://cearatransparente.ce.gov.br/.

References

Arbués, F., Barberán, R., Villanúa, I., 2004. Price impact on urban residential water demand: a dynamic panel data approach. Water Resour. Res. 40, W11402. doi:10.1029/2004WR003092.

Arbués, F., Barberán, R., 2012. Tariffs for urban water services in Spain: household size and equity. Int. J. Water Resour. Dev. 28, 123–140. doi:10.1080/07900627.2012.642235.

Bernoulli, D., 1954. Exposition of a new theory on the measurement of risk. Econometrica 22 (1), 23. doi:10.1142/9789814293501_0002.

Braga, B., Kelman, J., 2020. Facing the challenge of extreme climate: the case of Metropolitan Sao Paulo. Int. J. Water Resour. Dev. 36, 278–291. doi:10.1080/07900627.2019.1698412.

Campos, J.N.B., Souza Filho, F.A., Lima, H.V.C., 2014. Risks and uncertainties in reservoir yield in highly variable intermittent rivers: case of the Castanhão Reservoir in semi-arid Brazil. Hydrol. Sci. J. 59, 1184–1195. doi:10.1080/02626667.2013.836277.

Chu, L., Grafton, R.Q., 2019. Policy note: short-term pain for long-term gain: urban water pricing and the risk-adjusted user cost. Water Econ. Policy 5, 1871005. doi:10.1142/S2382624X18710054.

Cleveland, R.B., Cleveland, W.S., 1990. STL: a seasonal-trend decomposition procedure based on loess. J. Off. Stat. 6, 3–33. doi:10.4236/ojg.2014.412048.

Cominola, A., Nguyen, K., Giuliani, M., Stewart, R.A., Maier, H.R., Castelletti, A., 2019. Data mining to uncover heterogeneous water use behaviors from smart meter data. Water Resour. Res. 55, 9315–9333. doi:10.1029/2019WR024897.

Dalhuisen, J.M., Florax, R.J.G.M., de Groot, H.L.F., Nijkamp, P, 2003. Price and income elasticities of residential water demand: a meta-analysis. In: Land Econ., 79. University of Wisconsin Press, pp. 292–308. doi:10.2307/3146872.

de Maria André, D., Carvalho, J.R., 2014. Spatial determinants of urban residential water demand in Fortaleza. Braz. Water Resour. Manag. 28 (9), 2401–2414. doi:10.1007/s11269-014-0551-0.

Deyà-Tortella, B., Garcia, C., Nilsson, W., Tirado, D., 2016. The effect of the water tariff structures on the water consumption in Mallorcan hotels. Water Resour. Res. 52 (8), 6386–6403. doi:10.1002/2016WR018621.

Espey, M., Espey, J., Shaw, W.D., 1997. Price elasticity of residential demand for water: a meta-analysis. Water Resour. Res. 33 (6), 1369–1374. doi:10.1029/97WR00571.

Friedman, J.H., 1999. Greedy Function Approximation: a Gradient Boosting Machine Function estimation: numerical optimization in function space. The Annals of Statistics 1–10. doi:10.2307/2699986.
García-Rubio, M., Ruiz-Villaverde, A., González-Gómez, F., 2015. Urban water tariffs in Spain: what needs to be done? Water (Basel) 7, 1456–1479. doi:10.3390/w7041456.
Garrone, P., Grilli, L., Marzano, R., 2019. Price elasticity of water demand considering scarcity and attitudes. Util. Policy 59, 100927. doi:10.1016/j.jup.2019.100927.
Grafton, R.Q., Chu, L., Kompas, T., Ward, M., 2014. Volumetric water pricing, social surplus and supply augmentation. Water Resour. Econ. 6, 74–87. doi:10.1016/j.wre.2014.07.001.
Grande, M.H.Del, De Oliveira Galvão, C., Miranda, L.I.B.De, Sobrinho, L.D.G., 2016. The perception of users about the impacts of water rationing on their household routines1. Ambient. Soc. 19 (1), 163–182. doi:10.1590/1809-4422ASOC150155R1V1912016.
Greenwell, B., Boehmke, B., Cunningham, J., 2019. gbm: Generalized Boosted Regression Models. R Package Version 2.1.5.
Iplanfor, 2015. Fortaleza 2040. Prefeitura Municipal de Fortaleza, Brazil.
Lee, D., Derrible, S., 2020. Predicting residential water demand with machine-based statistical learning. J. Water Resour. Plan. Manag. 146. doi:10.1061/(ASCE)WR.1943-5452.0001119.
Luhmann, N., 2000. The Reality of the Mass Media. Stanford University Press, Stanford.
Mansur, E.T., Olmstead, S.M., 2012. The value of scarce water: measuring the inefficiency of municipal regulations. J. Urban Econ. 71 (3), 332–346. doi:10.1016/j.jue.2011.11.003.
Matikinca, P., Ziervogel, G., Enqvist, J.P., 2020. Drought response impacts on household water use practices in Cape Town. South Africa. Water Policy 22 (3), 483–500. doi:10.2166/wp.2020.169.
Monteiro, H., Roseta-Palma, C., 2011. Pricing for scarcity? an efficiency analysis of increasing block tariffs. Water Resour. Res. 47 (6). doi:10.1029/2010WR009200.
Molinos-Senante, M., 2014. Water rate to manage residential water demand with seasonality: peak-load pricing and increasing block rates approach. Water Policy 16, 930–944. doi:10.2166/wp.2014.180.
Molinos-Senante, M., Donoso, G., 2016a. Water scarcity and affordability in urban water pricing: a case study of Chile. Util. Policy. 43, 107–116. doi:10.1016/j.jup.2016.04.014.
Nunes Carvalho, T.M., de Souza Filho, F., de, A., Porto, V.C., 2021. Urban water demand modeling using machine learning techniques: case study of Fortaleza. Brazil. J. Water Resour. Plan. Manag. 147, 05020026. doi:10.1061/(asce)wr.1943-5452.0001310.
Olmstead, S.M., Michael Hanemann, W., Stavins, R.N., 2007. Water demand under alternative price structures. J. Environ. Econ. Manag. 54 (2), 181–198. doi:10.1016/j.jeem.2007.03.002.
Pesic, R., Jovanovic, M., Jovanovic, J., 2012. Seasonal water pricing using meteorological data: case study of Belgrade. J. Clean. Prod. 37, 299–303. doi:10.1016/j.jclepro.2012.07.031.
Pontes Filho, J.D., Souza Filho, F.de A., Martins, E.S.P.R., Studart, T.M.de C., 2020. Copula-based multivariate frequency analysis of the 2012–2018. Drought Northeast Brazil. Water. 12, 834. doi:10.3390/w12030834.
Ruijs, A., 2009. Welfare and distribution effects of water pricing policies. Environ. Resour. Econ. 43, 161–182. doi:10.1007/s10640-008-9228-6.
Sharma, S.K., Vairavamoorthy, K., 2009. Urban water demand management: prospects and challenges for the developing countries. Water Environ. J. 23, 210–218. doi:10.1111/j.1747-6593.2008.00134.x.
Kam, J., Stowers, K., Kim, S., 2019. Monitoring of drought awareness from Google trends: a case study of the 2011–17 California drought. Weather, Clim. Soc. 11 (2), 419–429. doi:10.1175/WCAS-D-18-0085.1.
Ma, X., Zhang, S., Mu, Q., 2014. How do residents respond to price under increasing block tariffs? evidence from experiments in urban residential water demand in Beijing. Water Resour. Manag. 28 (14), 4895–4909. doi:10.1007/s11269-014-0561-y.
McDonald, R.I., Weber, K., Padowski, J., Flörke, M., Schneider, C., Green, P.A., et al., 2014. Water on an urban planet: urbanization and the reach of urban water infrastructure. Glob. Environ. Change 27 (1), 96–105. doi:10.1016/j.gloenvcha.2014.04.022.
Orlowsky, B., Seneviratne, S.I., 2012. Global changes in extreme events: regional and seasonal dimension. Clim. Change 110, 669–696. doi:10.1007/s10584-011-0122-9.
Pal, I., Anderson, B.T., Salvucci, G.D., Gianotti, D.J., 2013. Shifting seasonality and increasing frequency of precipitation in wet and dry seasons across the U.S. Geophys. Res. Lett. 40, 4030–4035. doi:10.1002/grl.50760.
Quesnel, K.J., Ajami, N.K., 2017. Changes in water consumption linked to heavy news media coverage of extreme climatic events. Sci. Adv. 3 (10), e1700784. doi:10.1126/sciadv.1700784.
Rinaudo, J.D., Neverre, N., Montginoul, M., 2012. Simulating the impact of pricing policies on residential water demand: a southern france case study. Water Resour. Manag. 26 (7), 2057–2068. doi:10.1007/s11269-012-9998-z.
Rietveld, P., Rouwendal, J., Zwart, B., 2000. Block rate pricing of water in Indonesia: an analysis of welfare effects. Bull. Indones. Econ. Stud. 36, 73–92. doi:10.1080/00074910012331338983.
Romano, G., Salvati, N., Guerrini, A., 2014. Factors affecting water utility companies' decision to promote the reduction of household water consumption. Water Resour. Manag. 28 (15), 5491–5505. doi:10.1007/s11269-014-0818-5.
Ward, F.A., Pulido-Velázquez, M., 2008. Efficiency, equity, and sustainability in a water quantity-quality optimization model in the Rio Grande basin. Ecol. Econ. 66, 23–37. doi:10.1016/j.ecolecon.2007.08.018.
Whittington, D., Nauges, C., 2020. An assessment of the widespread use of increasing block tariffs in the municipal water supply sector. Oxford Research Encyclopedia of Global Public Health. Oxford University Press doi:10.1093/acrefore/9780190632366.013.243.
Zhang, B., Fang, K.H., Baerenklau, K.A., 2017. Have Chinese water pricing reforms reduced urban residential water demand? Water Resour. Res. 53 (6), 5057–5069. doi:10.1002/2017WR020463.