Catalogue no. 16-001-M
ISSN 1917-9693
ISBN 978-0-660-75074-3

Environment Accounts and Statistics Analytical and Technical Paper Series

Estimation of drinking water quantity in Canada from 2005 to 2019

by Rezvan Taki and Michael Schimpf

Release date: March 27, 2025

Statistics Canada Statistique Canada

Canada

How to obtain more information

For information about this product or the wide range of services and data available from Statistics Canada, visit our website, www.statcan.gc.ca.

You can also contact us by

Email at infostats@statcan.gc.ca

Telephone, from Monday to Friday, 8:30 a.m. to 4:30 p.m., at the following numbers:

• Statistical Information Service	1-800-263-1136
• National telecommunications device for the hearing impaired	1-800-363-7629
• Fax line	1-514-283-9350

Standards of service to the public

Statistics Canada is committed to serving its clients in a prompt, reliable and courteous manner. To this end, Statistics Canada has developed standards of service that its employees observe. To obtain a copy of these service standards, please contact Statistics Canada toll-free at 1-800-263-1136. The service standards are also published on www.statcan.gc.ca under "Contact us" > "Standards of service to the public."

Note of appreciation

Canada owes the success of its statistical system to a long-standing partnership between Statistics Canada, the citizens of Canada, its businesses, governments and other institutions. Accurate and timely statistical information could not be produced without their continued co-operation and goodwill.

Published by authority of the Minister responsible for Statistics Canada

An HTML version is also available.

Cette publication est aussi disponible en français.

Estimation of drinking water quantity in Canada from 2005 to 2019

by **Rezvan Taki** and **Michael Schimpf**[1]

Abstract

The global demand for drinking water has increased substantially with population growth, urbanization and industrialization. As a result, many countries, including Canada, are facing growing demand for high-quality drinking water, placing a significant strain on water treatment plants. Water treatment facilities provide water for various purposes, including thermoelectric power generation and for domestic, commercial and industrial use. To better understand drinking water production trends, forecast future water use, and achieve sustainable development in designing and planning drinking water supply systems, it is essential to have precise, dependable and uninterrupted data on potable water volumes processed by water plants. This paper examines the limitations caused by the frequency of Statistics Canada's Biennial Drinking Water Plants Survey on water production analysis. It explores various water modelling techniques and introduces a robust water prediction model aimed at providing estimates of water production for non-surveyed years at the national level in Canada.

This research uses quarterly water data from 2005 to 2019, combined with auxiliary data, such as gross domestic product and population estimates, for modelling. This study examines different techniques, including partial least squares regression, random forest and spline regression. After conducting cross-validation analyses, the spline regression model was selected as the most effective technique and was used to provide a comprehensive evaluation of drinking water production by water treatment plants from 2005 to 2019. This approach effectively filled in data gaps that were not previously surveyed, allowing for a more accurate and complete assessment of water production. Overall, this study's findings demonstrate the potential of using a spline regression model for predicting drinking water production and its use in filling data gaps, highlighting its significance for water resource management and policy making in Canada.

Keywords: Drinking water, spline regression, random forest, prediction model.

1 Introduction

The global need for drinking water is rising because of the rapid pace of population growth, urbanization and industrialization (Smeti et al., 2009).To ensure long-term prosperity, governments and policy makers must have access to accurate, reliable and continuous data on potable water volume for effective planning of water supply and demand.

From 2005 to 2019, the number of Canadians served by public water supplies grew by more than 5 million (+18.8%). However, during the same period, the total production of potable water in Canada decreased from 5,706.2 million cubic metres in 2005 to 4,866.1 million cubic metres in 2019. The recent decrease in drinking water production could be in response to a shift toward more sustainable consumption patterns, possibly attributable to better household water conservation practices, advances in technology and other factors. However, the extent of this trend may vary by region, depending on factors such as the local economy, policies, the regional climate and population growth (Ryan & Wang, 2012).

Providing continuous time series data on the volume of water produced by water treatment plants is important to better understand water use trends and forecast future water use. Statistics Canada's Biennial Drinking Water Plants Survey provides ongoing water production estimates, but the frequency of this survey limits analysis. This paper therefore explores different water modelling techniques and proposes a robust water prediction model to supply water estimates in non-surveyed years at the national level for Canada.

1. Statistics Canada. Email: Rezvan.taki@statcan.gc.ca and Michael.schimpf@statcan.gc.ca.

Numerous studies have been conducted to examine the relationship between drinking water use and demographic variables, such as population (Schleich & Hillenbrand, 2009); meteorological variables, including precipitation and temperature (Brown et al., 2013; Guo et al., 2013; Heyn & Winsor, 2015; Worland et al., 2018); and socioeconomic predictors (Babel et al., 2007; Sanchez et al., 2020; Sun et al., 2019). A comparable study (Sanchez et al., 2020) examined two rapidly growing U.S. states—North Carolina and South Carolina—using a geographically weighted regression model to analyze the interplay between socioeconomic factors, environmental variables and landscape patterns.

While some studies focus primarily on urban centres and megacities (Babel et al., 2007; Chu et al., 2009; Gharabaghi et al., 2019; Liu et al., 2023; Yurdusev & Firat, 2009; Zubaidi et al., 2022), others are restricted to rural areas (Keshavarzi et al., 2006; Singh & Turkiya, 2013) and limited to a single season or year (Machingambi & Manzungu, 2003; Makoni et al., 2004; Nyong & Kanaroglou, 2001).

Despite this vast body of literature, the availability of suitable datasets and the modelling of drinking water production at a national and annual scale are challenges. Although some studies have aimed to simulate drinking water consumption with population and income as a large-scale assessment, they did not consider seasonal variations in their modelling (Mitchell & Jones, 2005; Wada et al., 2011). The temporal scale of most of these studies was annual (Makoni et al., 2004) or much finer (e.g., daily or hourly resolution) (Herrera et al., 2010; Wong et al., 2010; Zhou et al., 2000), with no inclusion of quarterly variations. Furthermore, several studies have explored drinking water consumption patterns during wet and dry seasons, but the scope of their research was limited to a specific geographic region. Hence, many of the previous studies in the literature have limitations in terms of the period considered, geography analyzed or amount of available data.

To the best of the authors' knowledge, there is no common technique for calculating drinking water estimates at the national level in Canada. Hence, this paper fills a gap in the literature by outlining a methodology for modelling national-level drinking water consumption on an annual and a quarterly basis from 2005 to 2019. The outcomes of this modelling exercise are also assessed. Accurate forecasting of long-term drinking water demand is an essential tool for effective long-term planning and the expansion of water facilities to meet future needs. By providing insights into future demand trends, such forecasting can inform strategic decision-making processes and enable water authorities to allocate resources effectively in anticipation of future demand.

This paper is organized as follows: Section 2 introduces the datasets and the methodology used to simulate the volume of potable water; Section 3 presents the results of applying the proper model for modelling non-survey years; and Section 4 discusses results, conclusions and recommendations for future studies.

2 Datasets and methods

2.1 Data sources

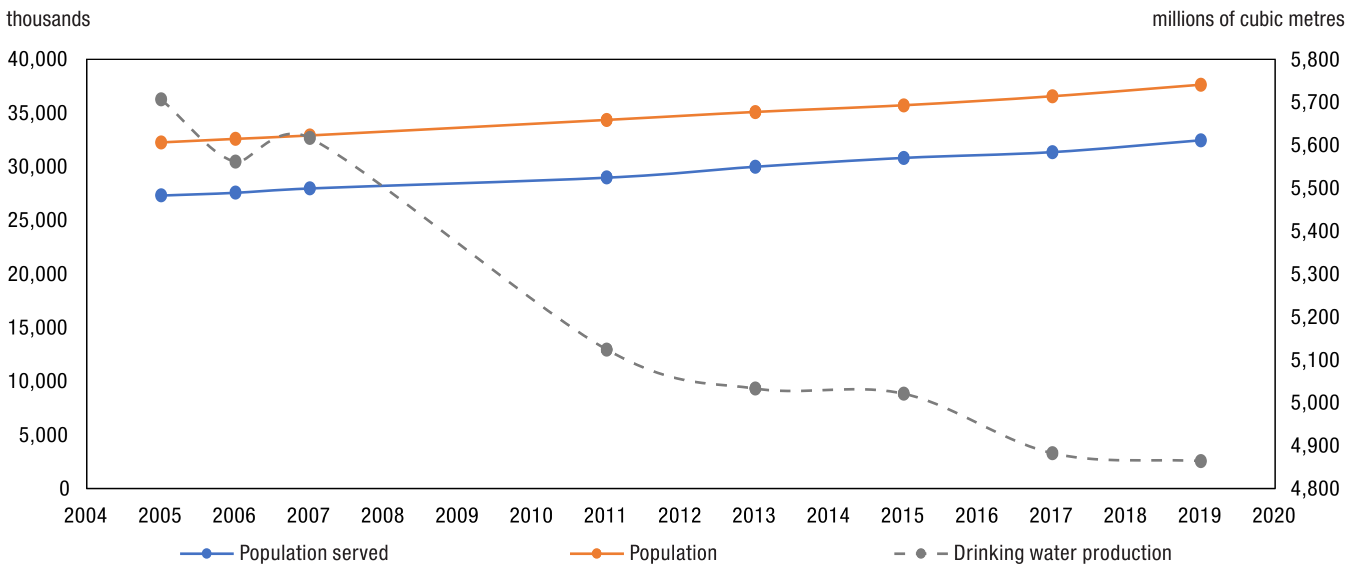
The Statistics Canada datasets used include quarterly estimates of gross domestic product (GDP) and the population as of July 1 derived from census data (Statistics Canada, 2021a), as well as biennial estimates of the population served by drinking water for 2005, 2006, 2007, 2011, 2013, 2015, 2017 and 2019 at the national level (Statistics Canada, 2021b). Based on the available GDP dataset, which only provided data on a quarterly basis at the national level, all modelling and comparisons were conducted at this specific temporal scale.

Drinking water data were sourced from the Biennial Drinking Water Plants Survey conducted by Statistics Canada (Statistics Canada, 2021b). This survey has collected biennial national and provincial information on the production, quality and associated costs of drinking water since 2005. The survey includes drinking water facilities that provide potable water for various uses, such as residential, commercial, industrial, institutional and other non-residential purposes. The survey focuses on water plants that serve 300 people or more, and sources of water include groundwater, surface water and groundwater under the direct influence of surface water (GUDI) (Statistics Canada, 2011).

According to the survey, surface water accounts for slightly more than 88% of public supply production, with the remaining water provided by groundwater and GUDI. From 2011 to 2019, the average daily water use across all sectors, including residential, industrial, losses and wholesale, decreased from 485 litres per person per day to 411 litres per person per day. However, the residential sector's proportion of drinking water use increased from 43% in 2011 to 51% in 2019 (Chart 1).

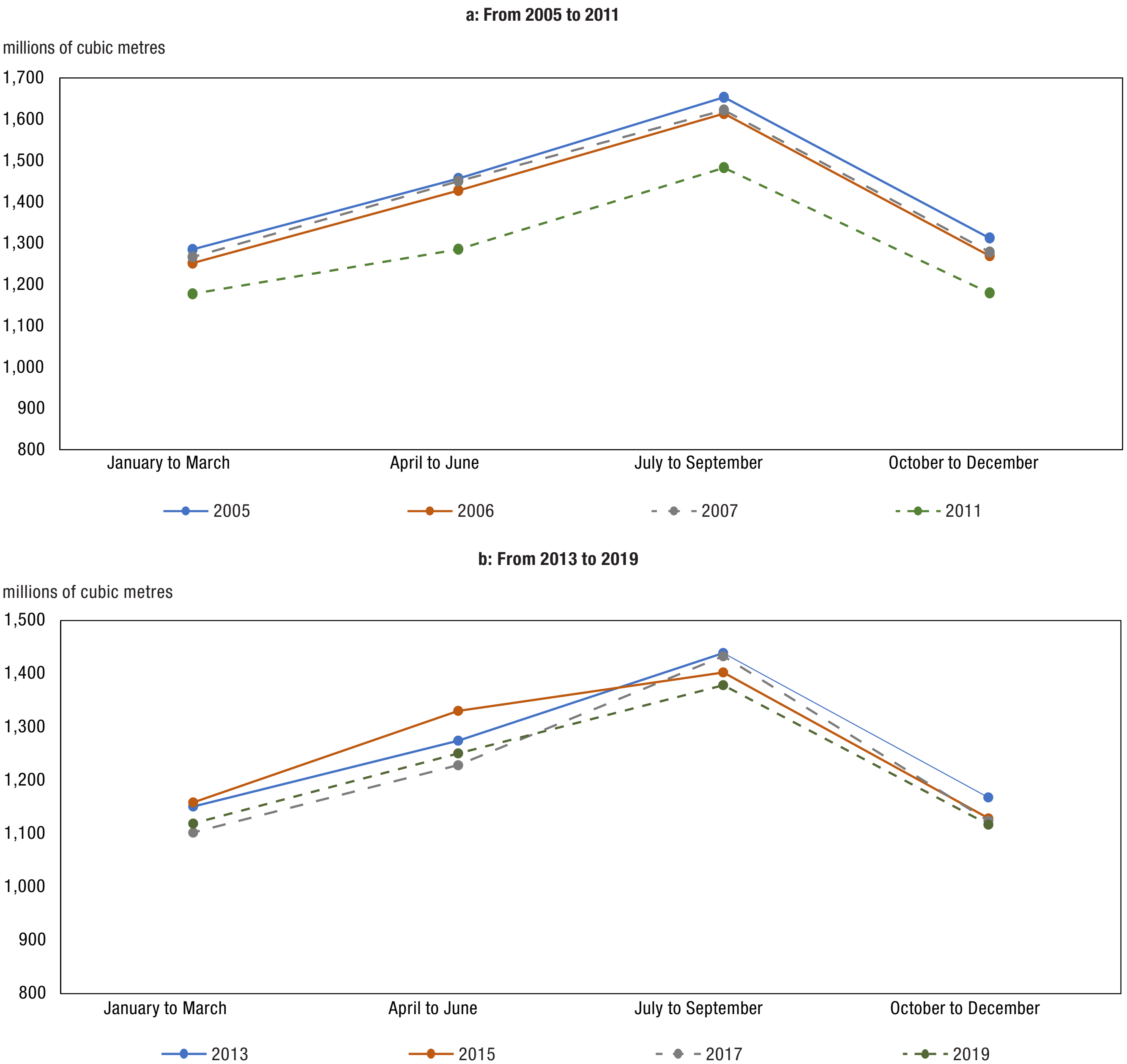
Drinking water use in Canada follows a seasonal trend, with the highest consumption occurring from July to September—because of factors such as warmer weather, increased outdoor activities, vacationing, gardening, higher water loss through evaporation and individual hydration habits—and the lowest consumption occurring from October to March because of colder weather. Chart 2 shows the average monthly drinking water production in Canada from 2005 to 2019.

Chart 1
Drinking water production, population and population served by drinking water plants from 2005 to 2019, obtained from survey data, Statistics Canada (2021a)



Source: Statistics Canada. Table 38-10-0092-01 Potable water volumes processed by drinking water plants, by source water type (x 1,000,000). Table 38-10-0093-01 Population served by drinking water plants. Table 17-10-0005-01 Population estimates on July 1, by age and gender.

Chart 2
Seasonal pattern for average monthly drinking water produced from 2005 to 2019, Statistics Canada (2021a)



Note: To ensure accessibility for all users, the data is presented in two graphs.
Source: Statistics Canada. Table 38-10-0272-01 Potable water volumes processed by drinking water plants, by month (x 1,000,000).

2.2 Methodology applied for drinking water production

The research undertaken for this paper determined that regression models are a methodology for examining the relationship between a dependent variable, y, and a set of auxiliary variables, x, in predicting drinking water use patterns by incorporating multiple explanatory factors. These factors include demographic variables, such as population size, and socioeconomic predictors, such as GDP.

Sophisticated forecasting methods using time series models, including multiple linear regression (MLR), random forest (RF), spline regression (SR) and partial least squares regression (PLS), were investigated. This section outlines four distinct methods used for modelling drinking water use, with the aim of predicting the annual amount of drinking water produced by water plants. The R programming language was used for both model fitting and model validation.

2.2.1 Multiple linear regression

An MLR model was used for comparing the forecasts made by the prediction models. This model regresses quarterly drinking water use on quarterly predictors of population and GDP, along with their interactions. The Akaike information criterion was used to conduct model selection and eliminate insignificant interactions (Friedman et al., 2001). The model is defined by the following equation:

$$y_i = \alpha_0 + \sum_{j=1}^{j} \alpha_j x_{ij} + \in_i \qquad (1)$$

where y_i is drinking water production calculated based on x_{ij}, which is the j^{th} predictor for the i^{th} quarter; α is the vector of regression coefficients $(\alpha = \alpha_1,\ \alpha_2,..)$; and α_0 and $\in$ are a constant and error term, respectively.

2.2.2 Random forest

RF is a non-parametric supervised learning algorithm that includes multiple decision trees (Breiman, 2001). This approach is constructed using a fixed number of trees (T), each of which is trained on a different bootstrap subsample of the training data.

To prevent any one tree or variable from dominating the model, a random subset of m covariates is used for splitting at each stage of tree growth. This selection process ensures that the ensemble is not overpowered by dominant trees and that a wider range of possible trees is explored. The optimal values for T and m are tuning parameters that are determined based on a balance between the computational costs of fitting additional trees and the resulting increase in accuracy. To generate predictions with an RF model, new data are fed into each decision tree, and the results from all terminal nodes are averaged to produce a prediction for each new observation, as shown below:

$$\hat{Y}(x) = \sum_{tree=1}^{T} f_{tree}(x) / T \qquad (2)$$

The R package randomforest (Liaw & Wiener, 2002) was used for the modelling.

2.2.3 Partial least squares regression

The PLS regression technique is a standard constructed predictive model used when highly collinear explanatory variables exist (Quenouille, 1949). In this technique, the relationship between a matrix of predictor variables (X) and the response variable (Y) is explained by latent variables or X-scores (T). The X-scores can explain the maximum amount of variability in both X and Y. The equations are as follows:

$$X = TP' + \varepsilon \qquad (3)$$

In this equation, T and P' are the score matrix and loading matrix, and $\in$ is the matrix of the X-residuals. T can also be calculated by the transformed PLS weights matrix as below:

$$T = XW^* \qquad (4)$$

Finally, the response (Y) is computed by the Y-weight matrix (C*) and the related residuals (F).

$$Y = TC^{*} + F \tag{5}$$

2.2.4 Spline regression

Cubic SR is a nonparametric piecewise polynomial regression that fits separate low-degree polynomials over different periods of the drinking water use curve. The curve is divided into k+1 segments by k breakpoints, also known as knots.

The study employed the leave-one-out cross-validation method to evaluate the performance of predictive models. This approach involves leaving out one observation from the dataset and computing the error estimate of the remaining data points. The model was trained using all data points, except a datum for one year, and subsequently validated with that year's data point using a cross-validation scheme. This process was repeated for all possible data combinations.

2.3 Statistical analysis

The best model was selected from among the above methods based on the lowest sum of squared estimate of errors (SSE) and mean absolute percentage error (MAPE):

$$MAPE = \frac{1}{n}\sum_{t=0}^{n}\left|\frac{A_t - F_t}{A_t}\right| \tag{6}$$

$$SSE = \sum_{t=0}^{n}\left(A_t - F_t\right)^2 \tag{7}$$

where

n = number of sample size

A_t = surveyed water production

F_t = predicted water production

The explained techniques were applied and compared in the following section.

3 Results and discussion

Table 1 in this study details the assessment of eight cross-validation runs for four models (RF, MLR, PLS and SR) conducted at the national level and on a quarterly time basis. The analysis of predictive accuracy across all statistical methods indicates that the SR model consistently exhibits the lowest total SSE and an average MAPE of approximately 1% over the validation years. Consequently, this study suggests that SR emerges as a fitter modelling technique for drinking water data, excelling because of its proficiency in capturing non-linear relationships and effectively handling intricate patterns within drinking water data, a conclusion that aligns with the findings of (Rinaudo, 2015).

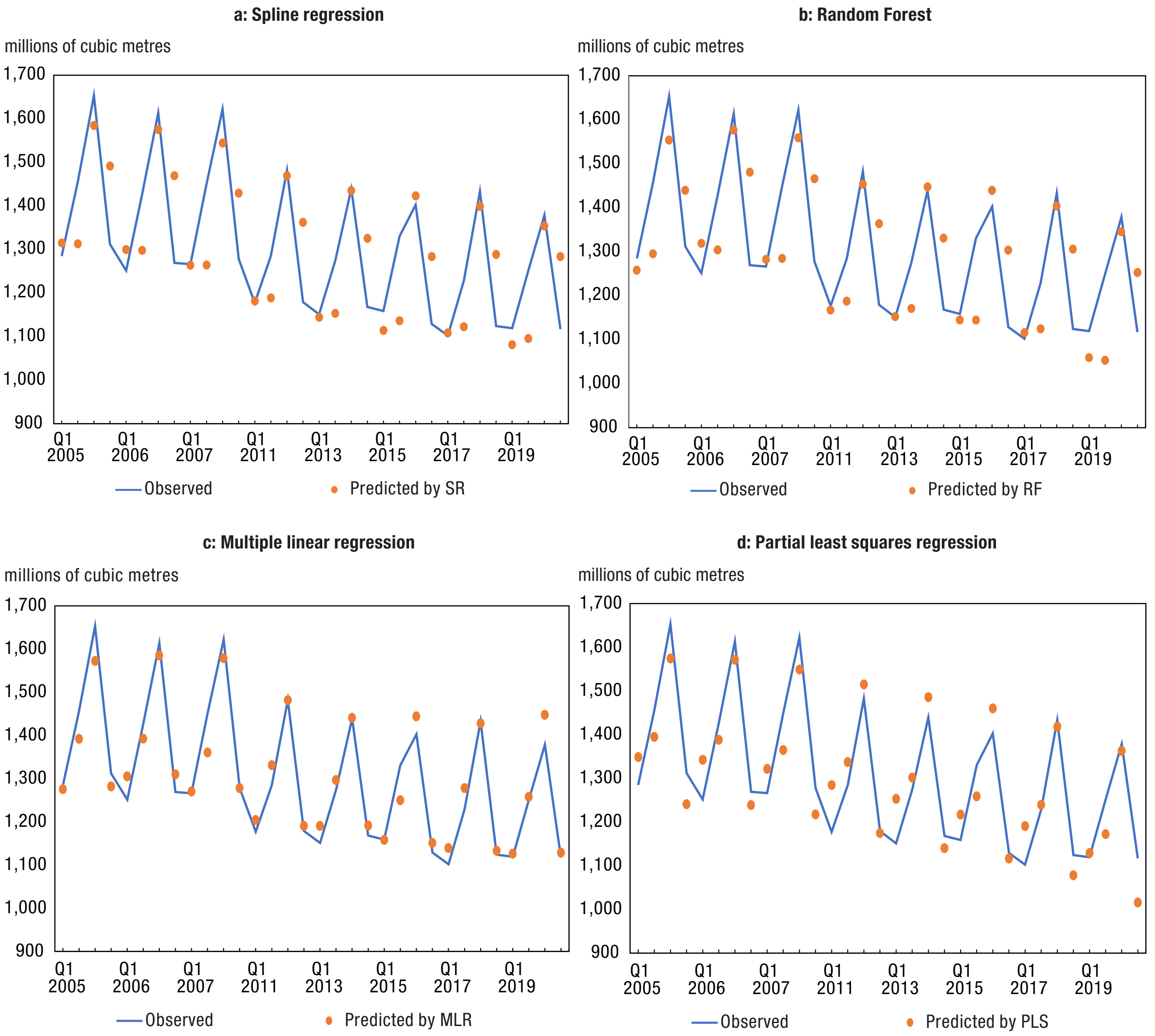
As a result of this strong performance, the performance of the SR model was investigated at the annual scale, demonstrating a decrease in performance of about 28% compared with the quarterly scale. The lower performance could be attributable to the lack of sufficient data available at the annual scale to fit an SR model accurately.

By contrast, the PLS model exhibited the highest total SSE over all validation years from 2005 to 2019, totalling approximately 142,930 million cubic metres squared ($(MCM)^2$). This was notably higher than the SSE values for the other techniques, which were 83,408 $(MCM)^2$ for RF, 75,509 $(MCM)^2$ for MLR and 34,116 $(MCM)^2$ for SR. Furthermore, the PLS model showed a MAPE of 4%, indicating lower predictive accuracy compared with the RF, MLR and SR models, which achieved MAPE values of 2%, 1% and 1%, respectively. These findingssuggest that the PLS model may require more ancillary variables to improve its performance (Quenouille, 1949). Also, the RF model's diminished performance could be attributed to its sensitivity to the limited sample size characteristic of this type of data, as suggested by previous research (Sultana et al., 2018).

Chart 3 shows the quarterly drinking water produced by water plants and modelled by different techniques from 2005 to 2019. According to the modelled estimates, the July-to-September period exhibited the highest drinking water production values, whereas the months from October to February displayed the lowest drinking water production compared with the remaining months. However, the results obtained from the monthly scale model were not substantiated because of the non-availability of supplementary data during the study, which impeded the validation process.

It is noteworthy that weather elements and the effects of climate change have the potential to alter the drinking water demand pattern (Gober, 2010; Milly et al., 2008). This observation underscores the complexity of the factors influencing water production and consumption. Importantly, these findings should be interpreted in light of certain limitations in this study. The lack of supplementary data for the monthly scale model is one such limitation, hindering the robust validation of results. Additionally, it is essential to acknowledge that this analysis may not account for all possible variables affecting drinking water production.

Chart 3
Comparison of regression model estimates for drinking water produced by water plants from 2005 to 2019



Source: Authors' computations.

Table 1
Statistical outcomes of the modelled data using different regression models and applied cross-validation scheme, 2005 to 2019

		Calibration years									Statistical measures	
											SSE	MAPE
Regression model	Temporal scale	2005	2006	2007	2011	2013	2015	2017	2019	Validation year	$(MCM)^2$	percent
Spline	quarterly		•	•	•	•	•	•	•	2005	6.14	0
		•		•	•	•	•	•	•	2006	6,559.06	1
		•	•		•	•	•	•	•	2007	13,228.22	2
		•	•	•		•	•	•	•	2011	6,109.92	2
		•	•	•	•		•	•	•	2013	632.47	0
		•	•	•	•	•		•	•	2015	4,006.76	1
		•	•	•	•	•	•		•	2017	900.18	1
		•	•	•	•	•	•	•		2019	2,672.89	1
	annual		•	•	•	•	•	•	•	2005	595.56	0
		•		•	•	•	•	•	•	2006	6,819.17	1
		•	•		•	•	•	•	•	2007	18,453.44	2
		•	•	•		•	•	•	•	2011	11,783.17	2
		•	•	•	•		•	•	•	2013	343.46	0
		•	•	•	•	•		•	•	2015	8,562.66	2
		•	•	•	•	•	•		•	2017	606.60	1
		•	•	•	•	•	•	•		2019	6.22	0
Partial least squares	quarterly		•	•	•	•	•	•	•	2005	21,775.96	8
		•		•	•	•	•	•	•	2006	438.97	0
		•	•		•	•	•	•	•	2007	26,933.67	1
		•	•	•		•	•	•	•	2011	34,980.42	1
		•	•	•	•		•	•	•	2013	21,585.32	11
		•	•	•	•	•		•	•	2015	949.87	0
		•	•	•	•	•	•		•	2017	1,388.54	1
		•	•	•	•	•	•	•		2019	34,877.61	6
Multiple linear regression	quarterly		•	•	•	•	•	•	•	2005	25,042.88	3
		•		•	•	•	•	•	•	2006	14,462.65	2
		•	•		•	•	•	•	•	2007	525.17	0
		•	•	•		•	•	•	•	2011	2,482.49	1
		•	•	•	•		•	•	•	2013	4,832.85	1
		•	•	•	•	•		•	•	2015	179.48	0
		•	•	•	•	•	•		•	2017	3,999.74	1
		•	•	•	•	•	•	•		2019	23,984.03	3
Random forest	quarterly		•	•	•	•	•	•	•	2005	33,757.28	3
		•		•	•	•	•	•	•	2006	1,022.61	1
		•	•		•	•	•	•	•	2007	16,648.88	2
		•	•	•		•	•	•	•	2011	7,033.92	2
		•	•	•	•		•	•	•	2013	7,539.31	2
		•	•	•	•	•		•	•	2015	270.77	0
		•	•	•	•	•	•		•	2017	8,190.35	2
		•	•	•	•	•	•	•		2019	8,944.86	2

Notes: SSE: sum of squared estimate of errors; MAPE: mean absolute percentage error; PLS: partial least squares; MLR: multiple linear regression; and RF: random forest. The dot (•) for each year indicates that the respective year was applied in the modelling.
Source: Authors' computations.

4 Conclusion

This study employed four techniques to estimate national-level drinking water use in Canada on a quarterly time scale. The SR model was recommended to quantify the continuous time series of drinking water use. Specifically, this approach can be used to fill data gaps from 2005 to 2019 when the Biennial Drinking Water Plants Survey did not collect drinking water data. Although the SR approach was the best method overall, its performance in estimating annual drinking water data was inferior compared with its performance in estimating data at a quarterly scale.

The limitations of the current analysis, such as the absence of supplementary data for the monthly scale model and the potential oversight of various variables, underscore the need for future research to build upon and refine the understanding gained through this analysis. Future investigations should prioritize addressing these limitations and strive to provide a more comprehensive understanding of the dynamics influencing water supply. By overcoming these challenges, future research endeavours will play a pivotal role in advancing the accuracy and applicability of models, ultimately contributing to a more refined and sophisticated estimation of factors influencing drinking water production.

Acknowledgments: We sincerely thank Beni Ngabo Nsengiyaremye and Martin Hamel for their invaluable contributions as consultants. Special thanks to Jennie Wang, Jessica Andrews, Terence Nelligan, Jenny Watt and Avani Babooram for their editing and feedback, which greatly refined this paper. We deeply appreciate their collaboration and support.

References

Babel, M. S., Gupta, A. D., & Pradhan, P. (2007). A multivariate econometric approach for domestic water demand modeling: an application to Kathmandu, Nepal. *Water Resources Management, 21*, 573–589.

Breiman, L. (2001). Random forests. *Machine Learning, 45*, 5–32.

Brown, T. C., Foti, R., & Ramirez, J. A. (2013). Projected freshwater withdrawals in the United States under a changing climate. Water Resources Research*, 49*(3), 1259–1276.

Chu, J., Wang, C., Chen, J., & Wang, H. (2009). Agent-based residential water use behavior simulation and policy implications: A case-study in Beijing City. *Water Resources Management, 23*, 3267–3295.

Friedman, J. I., Hastie, T., & Tibshirani, R. (2001). *The elements of statistical learning: data mining, inference, and prediction* (1st ed.). Springer. 10.1007/978-0-387-21606-5.

Gharabaghi, S., Stahl, E., & Bonakdari, H. (2019). Integrated nonlinear daily water demand forecast model (case study: City of Guelph, Canada). *Journal of Hydrology, 579*, 124182. http://www. doi:10.1016/j.jhydrol.2019.124182

Gober, P. (2010). Desert urbanization and the challenges of water sustainability. *Current Opinion in Environmental Sustainability, 2*(3), 144–150.

Guo, B., Chen, Y., Shen, Y., Li, W., & Wu, C. (2013). Spatially explicit estimation of domestic water use in the arid region of northwestern China: 1985–2009. *Hydrological Sciences Journal, 58*(1), 162–176. http://www .doi:10.1080/02626667.2012.745081

Herrera, M., Torgo, L., Izquierdo, J., & Pérez-García, R. (2010). Predictive models for forecasting hourly urban water demand. *Journal of Hydrology (Amsterdam), 387*(1), 141–150. http://www.doi:10.1016/j.jhydrol.2010.04.005

Heyn, K., & Winsor, W. (2015). Climate Risks to Water Utility Built Assets and Infrastructure. *Portland: City of Portland Water Bureau*.

Keshavarzi, A. R., Sharifzadeh, M., Haghighi, A. K., Amin, S., Keshtkar, S., & Bamdad, A. (2006). Rural domestic water consumption behavior: A case study in Ramjerd area, Fars province, IR Iran. *Water Research, 40*(6), 1173–1178.

Liaw, A., & Wiener, M. (2002). Classification and regression by randomForest. *R News, 2*(3), 18–22.

Liu, G., Savic, D., & Fu, G. (2023). *Short-term water demand forecasting using data-centric machine learning approaches.* IWA Publishing, 895–911. http://www.doi:10.2166/hydro.2023.163

Machingambi, M., & Manzungu, E. (2003). An evaluation of rural communities' water use patterns and preparedness to manage domestic water sources in Zimbabwe. *Physics and Chemistry of the Earth, Parts A/B/C, 28*(20-27), 1039–1046.

Makoni, F. S., Manase, G., & Ndamba, J. (2004). Patterns of domestic water use in rural areas of Zimbabwe, gender roles and realities. *Physics and Chemistry of the Earth, Parts A/B/C, 29*(15-18), 1291–1294.

Milly, P. C., Betancourt, J., Falkenmark, M., Hirsch, R. M., Kundzewicz, Z. W., Lettenmaier, D. P., & Stouffer, R. J. (2008). Stationarity is dead: Whither water management? *Science, 319*(5863), 573–574.

Mitchell, T. D., & Jones, P. D. (2005). An improved method of constructing a database of monthly climate observations and associated high-resolution grids. *International Journal of Climatology: A Journal of the Royal Meteorological Society, 25*(6), 693–712.

Nyong, A. O., & Kanaroglou, P. S. (2001). A survey of household domestic water-use patterns in rural semi-arid Nigeria. *Journal of Arid Environments, 49*(2), 387–400.

Quenouille, M. (1949). Approximate tests of correlation in time series. Journal of the Royal Statistical Society. Series B (Methodological), 11(1), 68–84 (1956). https://doi.org/10.1111/j.2517-6161.1949.tb00023.x

Rinaudo, J. (2015). Long-Term Water Demand Forecasting. Springer Netherlands. Understanding and managing urban water in transition, 239–268. https://doi:10.1007/978-94-017-9801-3_11

Ryan, S., & Wang, J. (2012). Residential water metering and pricing structures for the District of Mission, https://dspace.library.uvic.ca/bitstream/handle/1828/4287/Ryan_Sean_MPA_2012.pdf?sequence=1

Sanchez, G. M., Terando, A., Smith, J. W., García, A. M., Wagner, C. R., & Meentemeyer, R. K. (2020). Forecasting water demand across a rapidly urbanizing region. *Science of the Total Environment, 730*, 139050.

Schleich, J., & Hillenbrand, T. (2009). Determinants of residential water demand in Germany. *Ecological Economics, 68*(6), 1756–1769. https://www.doi:10.1016/j.ecolecon.2008.11.012

Singh, O., & Turkiya, S. (2013). A survey of household domestic water consumption patterns in rural semi-arid village, India. *GeoJournal, 78*, 777–790.

Smeti, E. M., Thanasoulias, N. C., Lytras, E. S., Tzoumerkas, P. C., & Golfinopoulos, S. K. (2009). Treated water quality assurance and description of distribution networks by multivariate chemometrics. *Water Research, 43*(18), 4676–4684. 10.1016/j.watres.2009.07.023

Statistics Canada. (2011). *Survey of Drinking Water Plants*. Retrieved Jan 12, 2022, from https://www23.statcan.gc.ca/imdb/p2SV.pl?Function=getSurvey&Id=108808.

Statistics Canada. (2021a). *Table: 36-10-0449-01. Gross domestic product (GDP) at basic prices, by industry, quarterly average*. Retrieved October 12, 2022, from https://www150.statcan.gc.ca/t1/tbl1/en/cv.action?pid=3610044901

Statistics Canada. (2021b). *Table: 38-10-0271-01 Potable water use by sector and average daily use*. https://www150.statcan.gc.ca/t1/tbl1/en/cv.action?pid=3610044901. Retrieved October 12, 2022.

Sultana, Z., Sieg, T., Kellermann, P., Müller, M., & Kreibich, H. (2018). Assessment of business interruption of flood-affected companies using random forests. *Water, 10*(8), 1049.

Sun, S., Fu, G., Bao, C., & Fang, C. (2019). Identifying hydro-climatic and socioeconomic forces of water scarcity through structural decomposition analysis: a case study of Beijing city. *Science of the Total Environment, 687*, 590–600.

Wada, Y., Beek, L. P. H. v., & Bierkens, M. F. P. (2011). Modelling global water stress of the recent past: on the relative importance of trends in water demand and climate variability. *Hydrology and Earth System Sciences, 15*(12), 3785–3808. https://hess.copernicus.org/articles/15/3785/2011/

Wong, L. T., Mui, K. W., & Guan, Y. (2010). Shower water heat recovery in high-rise residential buildings of Hong Kong. Applied Energy*, 87*(2), 703–709.

Worland, S. C., Steinschneider, S., & Hornberger, G. M. (2018). *Drivers of Variability in Public-Supply Water Use Across the Contiguous United States.* American Geophysical Union (AGU). 10.1002/2017wr021268

Yurdusev, M. A., & Firat, M. (2009). Adaptive neuro fuzzy inference system approach for municipal water consumption modeling: An application to Izmir, Turkey. *Journal of Hydrology, 365*(3-4), 225–234.

Zhou, S. L., McMahon, T. A., Walton, A., & Lewis, J. (2000). Forecasting daily urban water demand: a case study of Melbourne. *Journal of Hydrology, 236*(3-4), 153–164.

Zubaidi, S. L., Hashim, K., Ethaib, S., Al-Bdairi, N. S. S., Al-Bugharbee, H., & Gharghan, S. K. (2022). A novel methodology to predict monthly municipal water demand based on weather variables scenario. Journal of King Saud University. *Engineering Sciences, 34*(3), 163–169. http://www.doi:10.1016/j.jksues.2020.09.011