Urmia University

Water and Environmental Challenges

2026;1(1):17-28

Journal homepage: https://jwec.urmia.ac.ir

Water & Environmental Challenges

Structural Breaks and Reversible Temporal Regimes in Urban Water Consumption: A Comparative Analysis of Salmas City, Iran

Akram Najmi Gharehgheshlaghi[a], Keivan Khalili [b*] , Hanieh Mohammadi[c]

[a] MSc. Student, Environmental Engineering, Saba University, Urmia, Iran

[b] Urmia Lake Research Institute, Urmia University, Urmia, Iran

[c] MSc in Water Resources Engineering, Urmia University, Iran

A B S T R A C T

Ignoring structural changes in urban water consumption can undermine forecasting accuracy. Since most methods assume only permanent shifts, this study examines both persistent and reversible consumption regimes in Salmas City (2004–2025) using Bai–Perron, Pettitt, SNHT, and Markov-switching models in R. Bai–Perron, run under mean- and population-based specifications, served as the primary framework, with Pettitt and SNHT providing independent validation and Markov-switching testing whether regimes could reverse over time. Water consumption rose 149%, far outpacing the 44.7% growth in population. All Bai–Perron specifications identified 2015 as the primary breakpoint, closely matching Pettitt (2014) and SNHT (2015). The population-based model also suggested a possible second breakpoint in 2021, though this wasn't confirmed by the other tests. Unlike Bai–Perron's permanent breaks, the Markov-switching model's smoothed probabilities suggest the series leaned toward a lower-consumption regime during 2011–2017 before shifting back — though with only 22 observations, this is best read as a statistically inferred pattern rather than a confirmed change, and one that depends on including population as a covariate. Overall, relying on a single method risks an incomplete picture of these dynamics.

© 2025 Urmia University

Keywords:
Bai–Perron
Markov switching,
structural break,
temporal regimes,
water consumption

1. Introduction

A combination of the increasing size of urban populations, exacerbated effects of climate change, and a decline in the availability of renewable sources of water has made sustainable water resources management one of the primary issues of the twenty-first century. It is expected that by 2050, the number of people living in cities and suffering from water scarcity would grow from 933 million in 2016 to 1.7–2.4 billion people (He et al., 2021). Together with the forecasted increase in the total number of people on Earth – over 9.7 billion by 2050 – the process will have an additional strain on urban water supply systems. In such conditions, the accurate prediction of urban water demand becomes crucial for proper planning, allocation of resources, and policy-making. The accuracy of predictions, in turn, can be achieved only with a proper understanding of past behavior of water consumption time series. There is a great number of research that has revealed that urban water consumption does not always conform to the assumptions about stationarity that classical forecasting models work with. For instance, it was proved by the analysis of water consumption dynamics in Melbourne, Australia, during more than 70 years that both consumer behavior and weather conditions contribute to the dynamics of water consumption in a nonstationary way and the ratio between the two factors varies with time (Hemati et al., 2016). Moreover, a review of existing studies on the dynamics of urban water consumption revealed that household size, economic situation, and conservation policies are the main drivers of

* Corresponding Author
E-mail Address: k.khalili@urmia.ac.ir
https://doi.org/ 10.30466/jwec.2026.57513.1015
Receive Date: 19 Aug 2026
Accepted Date: 26 Sept 2026

changes in water consumption patterns (Dias & Ghisi, 2024). Thus, structural changes in urban water consumption are usually determined by the joint effect of demographic, economic, climatic, and policy-related factors. Not taking these changes into account while constructing the models will inevitably result in significant forecasting errors as a single model for the whole period may be unable to reflect the difference in the system dynamics.

In response to this challenge, researchers in econometrics and hydroclimatology have developed various methods for detecting changes and structural breaks in time series. Some approaches search for a break point in the statistical characteristics of a series, for example, the Bai-Perron approach in econometrics, SNHT, and Pettitt approaches in hydrology, while other approaches account for the process of transition from one regime to another and then back to the initial regime that occurs in the case of Markov switching models. Although these two categories of approaches are widely used in economics and hydrology for various purposes, including runoff time series modeling (Lu & Berliner, 1999; Akintug & Rasmussen, 2005) and short-term water consumption forecasting under water scarcity conditions, such as research on Tehran metropolitan area (Sajadifar and Pakseresht, 2024), the application of both approaches in parallel to one urban water consumption time series is rather unusual.

Despite the rich methodological background of these techniques in the field of econometrics and hydroclimatology, their practical use for identifying structural breaks within urban water consumption time series rather than precipitation or runoff is scarce. Only one research study can be mentioned in which the comparison between the breakpoint technique using Bai–Perron implementation in the structure package and two other methods—binary segmentation method and PELT algorithm using the changepoint package—was performed on an artificially generated time series representing water consumption for a hotel and a hospital by Santos et al. (2021). It is noted that the breakpoint method is more accurate in detecting structural breaks in time series based on the RMSE and MAE criteria. Therefore, it is justified to use the Bai–Perron method as the main tool for detecting structural breaks in this research study.

In addition to these foreign literature works applying the Markov chain models to climatic phenomena, there is also literature available from Iran. Mahavarpour (2018) employed a first-order two-state Markov chain to determine the probability of occurrence and return period of temperature extremes, i.e., Those temperatures which were more than two times the meanin Iran. The findings revealed that these events have a higher chance of occurrence and persistence during the cold season of the year. Similarly, Zeinali and Rouhi (2023) employed this method to examine the probability of occurrence and return period of heat waves of different durations in Ardabil Province. Even though both these research studies examined temperature instead of water usage, they nevertheless are an example of using a two-state first-order Markov-chain model with a transition-probability matrix in climatic and environmental data.

Regime-based methodologies similar to HMMs have also been applied in hydrology as a way of modeling the time-variant states of the system. For example, Bracken et al. (2014) used Hidden Markov Models coupled with climate indices to model several decades of runoff and showed that the technique is capable of reproducing statistical properties and nonstationary spectra of the observed time series. In a study with a practical approach, Turner & Galelli (2016) illustrated how the use of the hidden river-flow regimens in reservoir water supply operations could lead to improved operating policies. The method has been further developed for use in multi-reservoir systems by Espanmanesh & Tilmant (2022). Using Hidden Markov Models to analyze hydrological regimes in multiple Australian catchments, Peterson et al. (2021) found out that some of the catchments do not return to their initial regimes after the end of drought, which shows a lack of regime recovery.

Apart from Markov models, another class of Threshold Autoregressive (TAR) models, which was proposed by Chan and Tong (1986) and Tong (1978) in the first place, can serve as an alternative tool for modeling reversible regimes. Contrary to the previous model type, the transition between regimes in TAR models depends not on the hidden Markov chain but rather on the value of the observable threshold variable exceeding some predefined threshold. The comparison of the forecasting accuracy of regime-based nonlinear models (STAR, TAR, and autoregressive Markov-switching models) in the context of Colorado River streamflow prediction revealed the superiority of the regime-based models in comparison with traditional linear models in terms of seasonal dynamics and flow magnitude (Amiri, 2015). In the study which compares these two classes of models directly, Clements and Krolzig (1998) have tested Markov-switching and TAR models in forecasting U.S. gross domestic product and discovered statistically significant differences in the forecasting abilities of these two groups depending on the data properties. One area that remains relatively uncharted territory is whether a break that has been identified using a technique for identifying permanent breaks should be considered as a definitive irreversible break or is rather an

element of regime-based dynamics that can be uncovered only using a different methodology.

A second gap filled in this research project relates to the scale of the data. Most of the aforementioned papers have been dealing with relatively long time series lasting over several decades or high frequency data such as daily or monthly series. However, in many small and medium-size cities, especially in developing countries, the only data available is usually an annual series with few data points, generally less than 30 years. It is precisely in such conditions that the validity of any single method of inference becomes questionable, whereas the probability of obtaining unstable results is very high, especially when it comes to relatively complicated models such as Markov-switching models.

To fill these two gaps, the proposed approach in the current study will be the comparison, with the application of Bai-Perron test as the basic technique used to detect structural break points, with the independent usage of Pettitt and SNHT tests to evaluate the detected change points and a Markov switching model to analyze the existence of reversible temporal regimes. The study case is the water consumption time series of Salmas City located in the western watershed of Lake Urmia from 2004 to 2025, with 22 observations. In this region, there has been a severe reduction in the amount of available water resources during the past few decades, making it suitable to conduct this research.

In such circumstances, this study aims to investigate if these three independent change point tests identify a main structural break point or not, and if regime analysis identifies any temporal regime.

2. Material and Method

2.1. Study Area

Salmas City is one of the cities that belong to West Azerbaijan Province of northwestern Iran, which is located within the Lake Urmia catchment (Figure 1). Similar to other sub-basins within the Lake Urmia catchment, there has been an observed decline in surface water and groundwater resources in this catchment due to a number of factors, both climatic and anthropogenic in nature. These include decrease in precipitation, increase in evaporation rates, increase in agricultural lands, extraction of water from aquifers, and water structures on rivers discharging water to the lake (Alipour, 2006). Such circumstances have imposed an additional burden on the water resources of cities, including the city of Salmas, calling for accurate estimation of water consumption in urban areas.



Figure 1. Geographical location of the study area

2.2. Study Data

The data set utilized in this research is the annual time series data of urban water consumption and population in the city of Salmas in the years 2004–2025 (22 years of data). These data have been collected from the Regional Water Company of West Azerbaijan Province and the Statistical Center of Iran. Due to the limited sample size, some special considerations have been made during the application of the structural break methods and temporal regime approach.

Prior to conducting the structural break tests, the z-score standardization was applied to the water consumption time series:

$$z_t = (y_t - \bar{y}) / s_y \quad (1)$$

Where this is the actual value of the variable in year t, ȳ is the series mean, and s_y is its standard deviation. Unlike nonlinear transformations, such as logarithmic transformation, such a standardization would have no bearing on the positions of the breaks determined through the Bai-Perron algorithm since the latter algorithm tries to minimize the residual sum of squares in each regime and any linear transformation of the series leaves the relative structure of the data intact.

The population series has not been standardized since the dependent variable needed to be scaled in order to allow the estimation of the Markov-switching models.

The population variable has been kept in its natural scale in the merged models ("water ~ Population") since population was an explanatory covariate instead of a dependent variable, i.e., It was used as a covariate, not a breakpoint testing variable in this research. Since population was used as a covariate only, it was kept in its natural scale for the interpretation of the regression coefficient.

2.3. Materials and Methods

A total of four different but complementary methods have been used to detect and confirm structural breaks and time regimes in the water consumption series of the city of Salmas, including: The Bai-Perron approach (Bai & Perron, 2003); the Pettitt test (Pettitt, 1979) and SNHT (Standard Normal Homogeneity Test) (Alexandersson, 1986); and the Markov-switching model, which seeks to test the existence of reversible time regimes. All the tests were carried out using the R programming language version 4.5 through the "strucchange", "trend", and "MSwM" packages.

2.3.1. Bai–Person Multiple Structural Break Test

A linear regression model with m structural breaks, and consequently m+1 regimes, can be expressed as:

$$y_t = x_t'\beta_j + u_t, \;\; T_{j-1} < t \le T_j, \;\; j = 1, \ldots, m+1 \qquad (2)$$

where y_t is the dependent variable (standardized water consumption during period t), x_t is the vector of explanatory variables, which was considered under two separate specifications in this study: (1) a mean model that does not have an explanatory variable and (2) a population model with population as the explanatory variable; βj is the vector of regression coefficients for the j-th regime; and T1,…, Tm are the unknown breakpoints to be estimated (Bai-Perron, 2003).

Breakpoints are estimated jointly using the RSS minimization technique across all regimes:

$$(\hat{T}1, \ldots, \hat{T}m) = \operatorname{argmin}_{(T1,\ldots,Tm)} \; \Sigma_{j=1}^{m+1} \; \Sigma_{t=Tj-1+1}^{Tj} (y_t - x_t'\beta_j)^2 \qquad (3)$$

The number of breaks that minimize the Bayesian Information Criterion (BIC) is selected endogenously. The use of BIC avoids the problem of overfitting, which is the consequence of adding more parameters to the model, while RSS does not provide a proper criterion for selecting the number of breaks since it will always be decreasing in the number of parameters.

As there is a relatively small sample (n=22), a minimum length of 4 observations is needed to have enough data for estimating a break point in a segment. Moreover, both mean and population-based models were estimated for two scenarios – the first one imposed an upper limit of 3 breaks based on the preliminary analysis, and the second one had no upper limit, and the number of breaks was determined only by the h parameter. This scenario was chosen as a sensitivity check. The procedure was conducted using the "breakpoints ()" function in the "strucchange" package.

2.3.2. Complementary Validation Tests: SNHT and Pettitt

Two classical homogeneity tests were used to independently confirm the breakpoints obtained using the Bai-Perron approach. The non-parametric Pettitt test uses rank statistics and checks whether the observations come from the same distribution under the null hypothesis against the alternative of a change-point in one step. The test statistic is calculated as:

$$X_y = |2\Sigma_{i=1}^{y} r_i - y(n+1)|, \quad y = 1, 2, \ldots, n \qquad (4)$$

where the estimated change point corresponds to the time at which

$$X_k = \max(X_y), \text{ for } 1 \le y \le n \qquad (5)$$

Attains its maximum absolute value.

Where *r* denotes the rank of the i-th observation.

Since the conventional Pettitt test picks up only one breakpoint per analysis, this test was performed sequentially on the subseries obtained after detecting the first breakpoint, to look for any second breakpoint. The Standard Normal Homogeneity Test (SNHT) was carried out as another test to detect inhomogeneity in the mean of the series. Both tests were conducted using the "pettitt.test()" and "snh.test()" functions from the "trend" package.

2.3.3.Temporal-Regime Analysis Using the Markov-Switching Model

In contrast with traditional structural break approaches, where change is assumed to be deterministic, irreversible, and instantaneous in nature, the Markov-switching approach relaxes these restrictions by introducing an alternative view of urban water use, where it is viewed as a stochastic process capable of switching between different regimes, and not once but several times:

$$y_t = \mu_{St} + \gamma \cdot \text{Population}_t + \varepsilon_t, \;\; \varepsilon_t \sim N(0, \sigma^2) \qquad (6)$$

where $S_t \in \{1, 2\}$ is a latent state variable following a first-order Markov chain with transition probabilities (Hamilton, 2018):

$$p_{ij} = P(S_t = j \mid S_{t-1} = i) \qquad (7)$$

Given the limited sample size, the model specification was simplified so that only the intercept (μ_{St} was allowed to switch between the two regimes, while the population coefficient (γ) and error variance (σ^2) were assumed to be common and constant across both regimes. This specification was adopted because preliminary tests showed that allowing both the mean and variance to switch with only 22 observations repeatedly resulted in singularity during the computation of the Hessian matrix.

Variable population was chosen as the covariate, which reflects the main demographic factor affecting water consumption and allows for easier numerical convergence of the algorithm. Coefficient on this covariate was not considered in the interpretation, as the aim of the model was to capture temporal regimes and not a causal relationship between population and consumption. To ensure that the model converged to a correct rather than to a degenerate solution, where the probability of a regime is constant throughout the whole series, several fittings were performed, each with a different set of initial values. The model with the maximum log-likelihood was used in the end. Two regime specification (k=2) was chosen instead of three or more regimes, as, given n=22, fitting with three regimes yielded uninterpretable results (near-singular Hessian matrices and regimes with only one or two observations supporting them), suggesting overparameterization of the model, and k=2 was the most complex specification that could converge correctly under different initializations. The test of the selected two-regime switching model against a one-regime (non-switching) model through log-likelihood, AIC, and BIC is provided in the Appendix. The test procedure was carried out using the "msmFit()" command from the "MSwM" R package (Goldfeld and Quandt, 1973; Hamilton, 1989, 1994; Perlin, 2007).

3. Results and Discussion

3.1. Preliminary Descriptive Findings

A summary of descriptive statistics for the variables analyzed in this research is provided in Table 1. In particular, the population of Salmas City rose from 74,546 to 107,888 individuals, which amounts to a 44.7% increase. In the meantime, urban annual water demand rose from 4.96 Mm³ to 12.34 Mm³, thus reflecting a growth of about 149%—three times more than the population growth (Table 1). The aforementioned disproportionate increase is linked to rapid changes in the consumption time series as per the structural break analysis in Sections 3.2 and 3.3.

Table 1. Descriptive statistics of water consumption and population in Salmas City during the 2004–2025 study period (22 annual observations)

Variable	Mean	Standard deviation	Minimum	Maximum
Water consumption (Mm³)	8.32	2.7	4.96	12.34
Population (persons)	91783	10703	74546	107888

3.2. Bai–Perron Test

In the mean-based model ("water ~ 1"), the BIC value fell from 67.59 when there was no break (m=0) to the minimum of 17.38 when there was one break (m=1), before rising to 17.83 and 18.44 when m=2 and m=3, respectively. In the unrestricted specification, this upward trend went on up to m=4 (BIC=23.92) (see Figure 2). In both the three-break constrained and the unrestricted models, this trend clearly indicated 2015 as the main breakpoint, with a 95% confidence interval ranging from 2014 to 2016. According to the regression results of the two-regime model, the mean standardized water consumption changed from –0.857 to 1.028 between the two regimes, which means an increase of 1.885 standard units (std error = 0.122, $p < 0.001$). The model accounted for 92.3% of the variability in the data ($R2 = 0.923$, $F = 239.6$, $df = (1, 20)$, $p = 1.35\times10$).

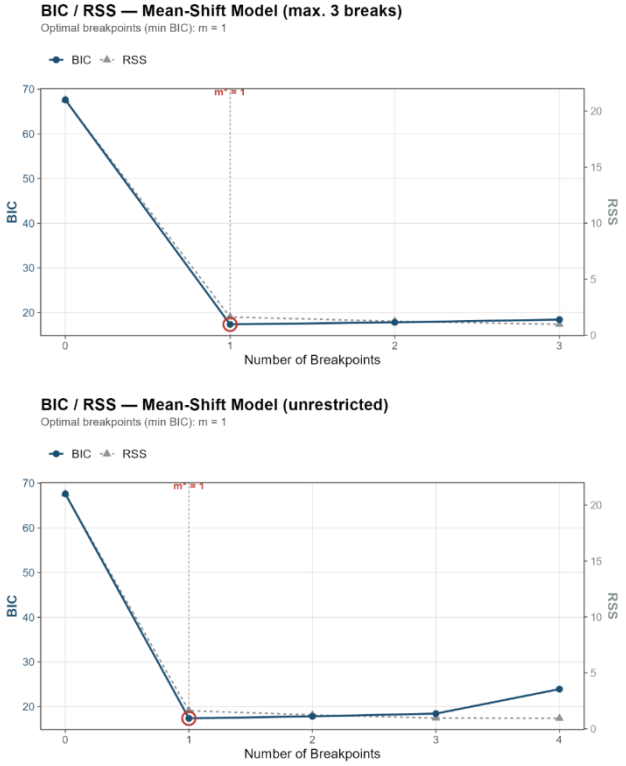


Fig. 2. BIC and RSS plots for the mean-based "water ~ 1" model under constrained and unrestricted specifications

The population-based model ("water ~ Population") had a decrease in BIC from 31.26 (m=0) to its lowest value of 7.96 at m=2, in contrast to 15.48 at m=1 and 13.31 at m=3. The results for this model are identical to those from the unconstrained and constrained versions (Figure 3). Two breakpoints have been found in 2015 and 2021. The confidence interval for the first breakpoint was 2014-2016 and for the second one 2019-2022. The three regimes regression model shows that the population parameter was statistically significant in the baseline regime ($\beta = 2.77\times10^{\wedge}$, p=0.007), and interactions of population parameter with the second regime (9.35×10-5, p=0.006) and third regime (-1.6×10-4, p=0.023) were statistically significant. The three regimes regression model accounted for 97.5% of the variance of the series (R2=0.975, adjusted R2=0.96, F=125.5, degrees of freedom = (5, 16), $p=2.99\times10^{-12}$).

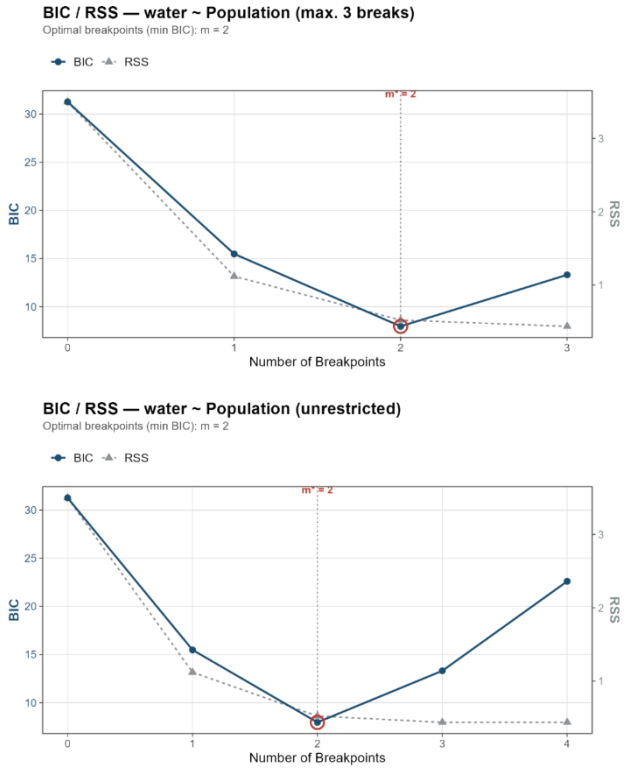


Fig. 3. BIC and RSS plots for the population-based "water ~ Population" model under constrained and unrestricted specifications

All four models, therefore, agreed that 2015 was the main year of the breakpoint, hence demonstrating the robustness of this result. Besides, the population model had an additional breakpoint in the year 2021, which was not detected using the mean model.

3.3. Results of the Complementary Tests (Pettitt and SNHT)

The Pettitt test applied to the entire series identified a change point in 2014, with a test statistic of U* = 121 and a p-value of 0.000748, indicating statistical significance at a confidence level exceeding 99.9%. This change point occurred only one year earlier than the breakpoint identified by the Bai–Perron test and falls entirely within its 95% confidence interval (2014–2016).

Sequential application of the Pettitt test to the two resulting sub-periods did not identify any statistically significant additional change points. For the first sub-period (2004–2014), a potential change point was identified around 2007–2008 with a p-value of 0.122, which was not statistically significant. For the second sub-period (2015–2025), a potential change point was identified around 2018–2019 with a p-value of 0.185, which was likewise not significant. Thus, the Pettitt test did not provide evidence supporting a statistically significant second breakpoint in 2021, reinforcing that this breakpoint should be treated as a candidate, model-dependent result rather than a confirmed structural change.

The SNHT showed a break point occurring in 2015, with a T-value of T=19.382 and a p-value less than 2.2×10-16, the same as that identified by the Bai-Perron test.

In general, the three distinct tests (Bai-Perron, Pettitt, and SNHT) found the 2014-2015 period as the main break point, while the second break point in 2021 was found solely using the population-oriented Bai-Perron method and needs to be verified independently for it to be considered an actual structural change.

3.4. Results of the Markov-Switching Model

The two regimes Markov-switching model converged stably (AIC=8.59, BIC=21.13, LogLik=-1.29). The estimated intercept was -7.723 for Regime 1 (Standard error = 0.330, p<2×10-16) and -8.533 for Regime 2 (Standard error = 0.330, p<2×10-16). Upon control of the impact of population, Regime 2 demonstrated significantly lower consumption than Regime 1 by 0.81 standardized units. The population coefficient turned out statistically nonsignificant (p=0.9998), implying that the population did not play an essential role in explaining the within-regime variability of the model. This needs to be stressed explicitly: the same population variable turns out statistically significant in the Bai-Perron model (p=0.007, Section 3.2) but statistically nonsignificant here, which demonstrates that statistical significance is model specification-dependent and thus cannot be taken as an indicator of practical importance of a variable on its own; the per capita analysis in Section 3.8,

rather than any p-value separately, is the best tool to understand the practical role of population in consumption increase.

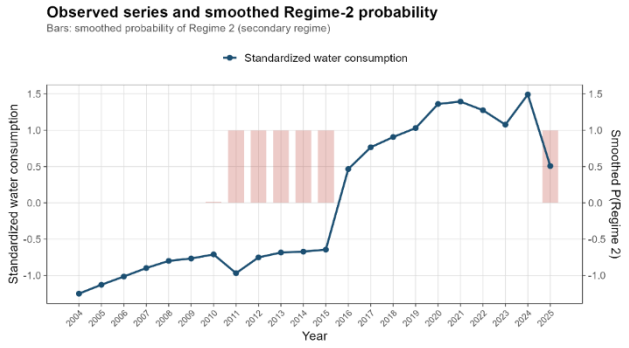


Fig. 4. Smoothed regime probabilities obtained from the Markov-switching model

The persistence probability matrix suggests that Regime 1 was highly persistent with a persistence probability of 93.7% and an average persistence duration of about 16 years, while Regime 2 was relatively more transient with a persistence probability of 66.7% and an average persistence duration of about 3 years. The 200 simulations from the parametric bootstrap procedure (out of which 164 converged) suggest that there is considerable uncertainty associated with the point estimates of persistence probabilities at n=22 – the 95% confidence interval for the persistent regime's persistence probability was [0.559, 0.984] (average duration 2.3-63.8 years) and for the transient regime's persistence probability was effectively [~0, 0.904] (average duration 1-10.4 years). In other words, the interval estimates are very broad and the point estimates cannot be interpreted as being precise parameter values. The formal statistical test comparing this model with a single-regime (no switching) alternative strongly favors the latter one: the non-switching model (standard linear regression) has AIC=27.99, BIC=31.26, and log-likelihood=-10.99 vs. AIC=8.59, BIC=21.13, and log-likelihood=-1.29 of the two regimes model (AIC and BIC criteria favor the latter even with extra parameters). Repeating the estimation with 20 random starting points yielded the same value of the log-likelihood each time (up to six decimals), suggesting the existence of one consistent optimum and not dependence on initial conditions. The standard deviation of the smoothed regime probabilities for Regime 1 was 0.4484, which means a significant variability of the regime probabilities across time and the validity of the differentiation into two regimes. Inspection of the yearly probabilities of the two regimes (see Table 2 and Figure 4) shows that the process mostly belonged to Regime 1 while the probability of Regime 2 showed a sharp increase from 2011 to 2017 (probability of Regime 2 above 0.98 in each year from 2011 to 2015), declined until 2024, and increased drastically again in 2025 (smoothed probability of Regime 2 = 1.000), the last year of the series. Since, in a Markov switching model, we are determining probabilities of latent regimes and not seeing an actual change that is taking place physically or institutionally, this phenomenon is termed in this paper as a temporary probability change inferred by statistical means instead of a verified structural change, and especially the recurrence in 2025 since it is the latest date in the series.

Table 2. Annual smoothed regime probabilities from the Markov-switching model, Salmas City, 2004-2025

Year	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025
P_Regim	1	1	1	1	1	1	0.983	0.001	0	0	0	0	1	1	1	1	1	1	1	0.997	1	0
P_Regim	0	0	0	0	0	0	0.017	0.999	1	1	1	1	0	0	0	0	0	0	0	0.003	0	1

3.5. Summary and Methodological Comparison of the Results

The findings of these six tests that were performed in this study are provided in Table 3: four Bai-Perron formulations (which cannot be considered statistically independent from each other since they are all based on the same model), together with the Pettitt and SNHT tests. The comparison of the findings shows the existence of a remarkable similarity and discrepancy trend, as detailed below.

Table 3. Breakpoints identified by the six analyses (four Bai-Perron specifications, Pettitt, and SNHT)

Method	Breakpoint
Bai–Perron ("water ~ 1", constrained to 3 breaks)	2015
Bai–Perron ("water ~ 1", unrestricted)	2015
Bai–Perron ("water ~ Population", constrained to 3 breaks)	2015. 2021
Bai–Perron ("water ~ Population", unrestricted)	2015. 2021
Petite	2014
SNHT	2015

Indeed, the result shows that all the three independent methods, namely Bai-Perron, Pettitt, and SNHT, yielded almost the same 2014-2015 period as the first major breakpoint. With the consideration of the essential difference in the theoretical basis of these three methods,

this convergence may serve as additional proof to support the significance of this breakpoint.

Nevertheless, the secondary candidate breakpoint in 2021 has been recognized only via population-based Bai-Perron test and should be considered with caution. A likely cause of this inconsistency could be explained with the control for population, which increases the power of discriminatory ability of the method, while classical homogeneity tests applied directly to the consumption series cannot provide this control.

Most importantly, there is a key difference between the Bai-Perron and the Markov-switching analysis. While Bai-Perron treats the 2015 breakpoint as a structural break (that is, a permanently occurring break under the assumptions of Bai-Perron, where no reversion to a previous state is possible), the Markov-switching model offers a more nuanced view: the time series was more prone to Regime 2 up to 2017 than in 2011 before the probability of being part of Regime 1 increased again. Both interpretations uncover different sides of a single complex process: a real structural break could have taken place in 2015, but there might also have been a statistical change in the probability of being part of Regime 1 around the new probability.

3.6. Domain-Based Interpretation of the Findings

Some of the potential factors behind the sudden jump in water consumption could be alterations in urban water sources, tariff systems, infrastructure development and consumer patterns. These, however, need to be treated as hypothesis rather than a proven mechanism, which can be revealed from statistical analysis, since these parameters have not been introduced in any of the model equations. Equally important, the proposed candidate breakpoint for the year 2021 is temporally consistent with the severe droughts that were observed in the Lake Urmia region between 2020-2021. Nonetheless, such an interpretation still remains speculative at this point since there is no information about the correlation between these phenomena – either in the form of hydroclimatic data (for instance, precipitation rates, reservoir inflows and groundwater levels), which could be analyzed along with the time series of water consumption. The following Figure 5 presents the disproportional trend, identified breakpoints and temporal regimes.

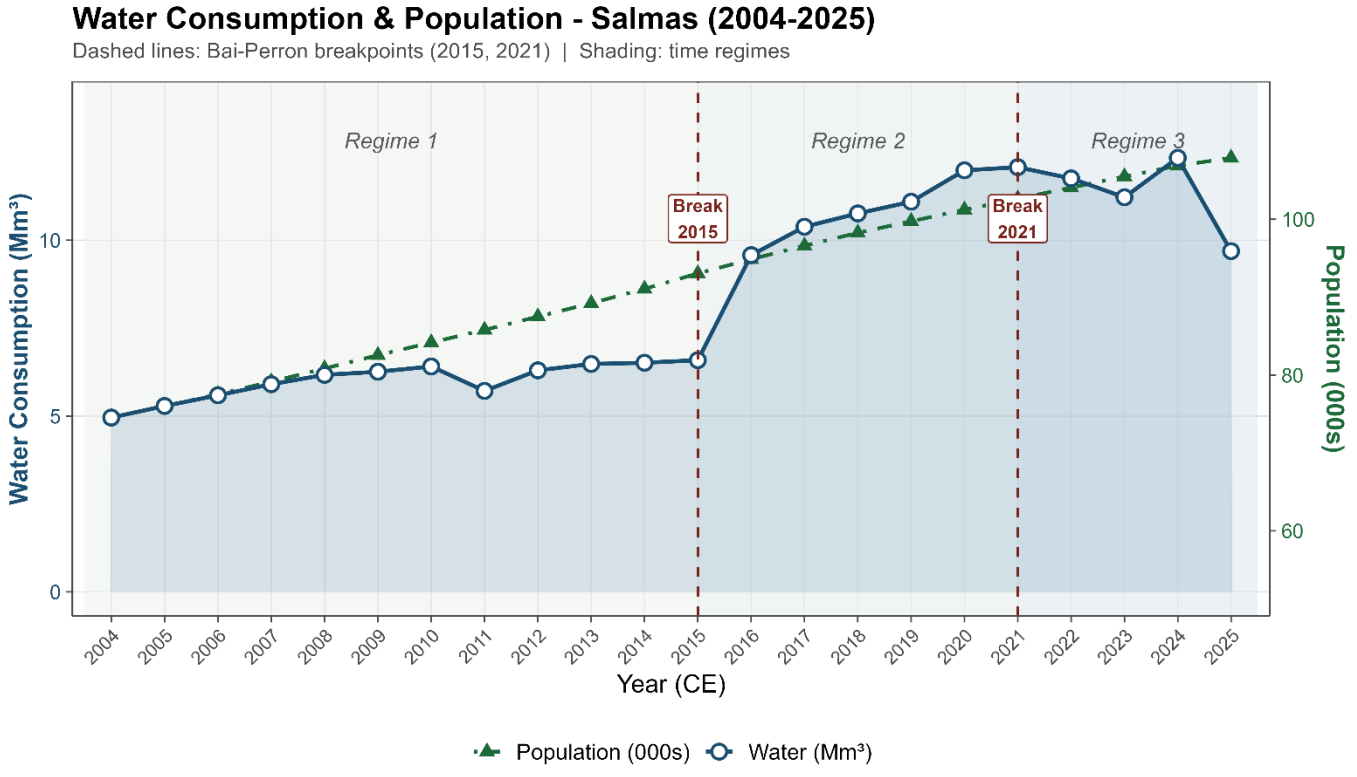


Fig 5. Temporal regimes resulting from the two identified structural breaks

3.7. Discussion

The agreement of all four variations of Bai-Perron (constrained/unconstrained × two model structures) to choose 2015 as the principal breakpoint is fully aligned with the results reported by Santos et al. (2021). By utilizing simulated data of water usage with pre-set breakpoints for a hotel and a hospital, they showed that the breakpoint analysis performed in "strucchange" package had greater accuracy in capturing the correct points of the breaks compared to two other methods – binary segmentation and PELT in "changepoint" package – based on RMSE and MAE. The above-mentioned independent empirical evidence supports the choice of Bai-Perron as the principal approach in the current paper beyond its own convergence.

The fact that Pettitt's test has identified the 2015 shift as a significant breakpoint with a very low p-value (0.000748), while failing to identify the 2021 breakpoint in the sequential analysis, has been in agreement with the findings of Mallakpour and Villarini (2016) simulation work on the statistical performance of the Pettitt test. They found out that the test power of the Pettitt test is proportional to the effect size, which is defined as the ratio between the size of the shift and the standard deviation of the series. The coefficient of the 2015 break point in the mean-based model (1.885) is indicative of a fairly large shift size, which is consistent with its detection by the Pettitt test. On the other hand, the 2021 breakpoint has only been detected in the population-based model after removing the effect of the population trend.

The transition to a transient Regime 1, indicated by the Markov-switching model, with years 2011-2017 being an intermediary period and transition back to 2018 onwards, is logically analogous to the results obtained by Peterson et al. (2021). Having studied hydrological regimes in a number of Australian catchments, these researchers found that certain catchments switch back to the initial regime after experiencing a shock, namely, drought in their case, while others undergo a permanent regime failure. The Salmas series presented an example of a system switching back to its original regime. There arises an interesting research question in this regard: is this reversible regime behavior specific to a certain type of factor that triggers the second regime shift, which could not be determined in this particular study due to a lack of contextual information, or does it have a more general character in the context of urban water consumption systems in comparison to natural hydrological systems?

On account of the concept of asymmetric regime persistence, the result of this study – 93.7% regime persistence for Regime 1 versus 66.7% for Regime 2 – is consistent with the wider theory developed by Akıntuğ and Rasmussen (2005) in analyzing the annual hydrological regimes, where the model used by them draws attention to the connection between transition probabilities and regime persistence in hydrological data. In the current study, the regime with a shorter duration (Regime 2) corresponds to the time interval that can be linked to water stress in the region (Section 3.6).

In Iran, studies by Mahavarpour (2018) and Rouhi & Zeinali (2023), despite concentrating on temperature-based phenomena and not water consumption, respectively, for Iran and Ardabil province, were able to model the temporal dependence and persistence of climatic patterns in Iran through first-order two-state Markov chains. Domestic experience in this respect provides sufficient ground to suggest that application of this approach to Iranian data, no matter what kind of variable it would be, climatic or consumption-based, is technically and conceptually feasible. One key difference in methodology is that while both cited studies were based on daily data with much larger samples over decades, the current study had to use annual data with 22 observations and, thus, was limited to a far simpler Markov-switching framework, where switching was possible only in the intercept but not variance.

Population growth cannot be the only explanation for such high levels of increase in water consumption since the latter occurred at a rate more than three times that of population growth, and fits into the general conclusion of Hemati et al. (2016). These researchers' investigation of 72-year water consumption data in Melbourne revealed that the influences of climatic and anthropogenic factors on water consumption do not stay steady but act simultaneously and non-stationary. In a similar manner, the systematic review of Dias and Ghisi (2024) has shown that the influence of multiple factors other than population has been the major driver of changes in water consumption levels. Such regional variations have already been noted by previous studies, including the study by Alzahrani and Tawfik (2025) carried out in the context of Saudi Arabian cities. This information provides additional arguments in favor of a multidimensional explanation of the trends in Salmas described in Section 3.6.

3.8. Limitations

However, there are a number of limitations associated with the research that have to be noted in connection with the interpretation of the results obtained. The first and the most important of them is that only 22 observations per year were used for the analysis. The estimation of the two-state Markov-switching model that includes regime-specific intercepts, common slope and variance and 2x2 transition matrix from such a sample is inevitably associated with great estimation uncertainty. The model converged to the same log-likelihood in 20 different runs with randomly selected initial parameters, and two regimes were clearly preferable to one regime in terms of AIC (8.59 against 27.99) and BIC (21.13 against 31.26; see section 3.4). In other words, the model is well identified using

the available sample, but the parametric bootstrap (200 runs, 164 successfully refitted) shows quite broad 95% confidence intervals for the persistence probability (0.559 – 0.984 for the persistent regime and practically 0 – 0.904 for the transient one; see section 3.4).

Secondly, since both population and total water consumption show very high (near-linear) temporal trends (correlation between Year and Population = 0.999 over the time range of this study), one part of the relationship between population and water consumption could be explained by a common trend rather than by population itself. For this reason, we calculated annual water consumption per capita (total water consumption divided by population). Annual water consumption per capita still saw a rapid rise after 2015, from an average pre-2015 value of about 72.1 m³ per capita to an average post-2015 value of around 109.0 m³ per capita (a rise of 51%), suggesting that the increase in consumption after 2015 was not due to population growth. Also related, the Markov-switching model was re-estimated without the population covariate (intercept-only, two-regime model) to see if the 2011-2017 regime pattern was affected by including population growth adjustment. The alternative model fit considerably worse (AIC=41.94, BIC=50.30, log-likelihood=-18.97, compared to AIC=8.59, BIC=21.13, log-likelihood=-1.29 for the population-adjusted model) but, more importantly, produced different regimes altogether: without population growth, the two regimes corresponded roughly to a split before and after 2015 (in line with the Bai-Perron results for mean shift) rather than to the 2011-2017 temporary regime pattern. The correlation between the smooth probabilities for Regime-1 in the two models is -0.356 (with and without population), suggesting a weak negative relationship rather than a consistent regime structure. This is a crucial caveat: the temporary and reversible regime pattern outlined in sections 3.4-3.5 is contingent upon accounting for population growth; the regime is not present in the water consumption series alone. We include it explicitly so that the interpretation of the reversible regime is seen to be contingent upon the specification.

Third, in answer to a reviewer query concerning whether the z-score standardization process could have affected the estimated break point locations, it was confirmed that this was not the case: minimization of the residual sum of squares for all possible single break points for the water consumption series on the raw scale and z-score standardized scale finds the exact same year of break point (2015–2016), and the ratio of the two residual sums of squares curves for all possible break points is constant, indicating that the z-score standardization was just a linear transformation with no effect on the break point location.

Fourth, the explanatory variables used in this analysis are limited to annual water consumption and population. Thus, the proposed mechanisms described in Section 3.6 (changes in tariff, changes in supply source, network expansion, and drought events) cannot be tested directly in this analysis and are offered as hypotheses instead of causes. Further analysis using tariff history data, supply source data, or hydroclimatic indicators (e.g., precipitation, reservoir level) will enable testing of these hypotheses.

All of the above listed issues indicate that the results of this analysis, especially the results based on the Markov-switching analysis and the proposed 2021 breakpoint, are to be considered illustrative but not conclusive.

4. Conclusions

The main objective of this study was to detect and test structural breaks and temporal regimes in the time series data for the annual water consumption in Salmas City during the period 2004–2025. For this purpose, a comparative analysis using a four-step approach involving the Bai-Perron test, Pettitt test, SNHT test, and the Markov switching model was adopted.

These results can be summed up along three key points. First, all three independent procedures for detection of the breakpoint (Bai-Perron, Pettitt, and SNHT) pointed out the high consistency of breakpoint timing in the 2014-2015 interval as the only and the most important breakpoint in the time series of Salmas's water consumption, showing how robust this finding is. Second, the population-based Bai-Perron model found a second breakpoint in 2021; however, it appeared only in the population-based specification and was not confirmed by homogeneity tests. This result proves that accounting for context might improve the discriminatory power of the methods in detecting structural breaks. Third, and this point is the most important from the methodological standpoint, the Markov-switching procedure showed that, unlike the assumption of irreversibility made in the Bai-Perron model, the smoothed probabilities of the Salmas water consumption time series show that there was a temporarily inferred shift to another regime in 2011-2017, with the probability of the primary regime increasing after 2017. This pattern cannot be detected using the classical approach to the problem of structural breaks.

Combining these results, the following methodological recommendation can be made regarding the future research of time series on urban water consumption: the use of a sole approach to analysis, no matter how efficient it is—in this case, the Bai-Perron procedure, which was proven in terms of identification of breaks in water consumption in prior research, such as Santos et al. (2021), may be insufficient to show all peculiarities of the real-world system under investigation. In parallel using several approaches based on entirely different premises—permanent break or reversible regime—it becomes possible to observe different characteristics of the structural

changes that may not be revealed by means of any of the procedures in isolation.

In practical terms, the results can assist in the water resources management of Salmas City, especially in the construction of water consumption forecasting models, accounting for both structural change found around 2015 and regime shifts found in 2011-2017.

The findings need to be considered with caution due to the small sample size (n = 22) and associated uncertainty with respect to the Markov switching model estimations (as described in section 3.8); the key 2014-2015 breakpoint detected using three different approaches can be viewed as the main result of this study, while the 2021 breakpoint and 2011-2017 regime probabilities can be regarded as assumptions for further investigation.

References

1. Akintug, B., & Rasmussen, P. F. (2005). A Markov switching model for annual hydrologic time series. *Water resources research*, *41*(9).
2. Alexandersson, H. (1986). A homogeneity test applied to precipitation data. *Journal of climatology*, *6*(6), 661-675.
3. Alipour, S. (2006). Hydrogeochemistry of seasonal variation of Urmia Salt Lake, Iran. Saline systems, 2(1), 9.
4. Alzahrani, F., & Tawfik, R. (2025). Regional heterogeneity in urban water consumption in Saudi Arabia. Water, 17(8), 1156.
5. Amiri, E. (2015). Forecasting daily river flows using nonlinear time series models. *Journal of Hydrology*, *527*, 1054-1072.
6. Bai, J., & Perron, P. (2003). Computation and analysis of multiple structural change models. *Journal of applied econometrics*, *18*(1), 1-22.
7. Bracken, C., Rajagopalan, B., & Zagona, E. (2014). A hidden M arkov model combined with climate indices for multidecadal streamflow simulation. *Water Resources Research*, *50*(10), 7836-7846.
8. Chan, K. S., & Tong, H. (1986). On estimating thresholds in autoregressive models. *Journal of time series analysis*, *7*(3), 179-190.
9. Clements, M. P., & Krolzig, H. M. (1998). A comparison of the forecast performance of Markov-switching and threshold autoregressive models of US GNP. *The Econometrics Journal*, *1*(1), 47-75.
10. Dias, T. F., & Ghisi, E. (2024). Urban water consumption: A systematic literature review. *Water*, *16*(6), 838.
11. Espanmanesh, V., & Tilmant, A. (2022). Optimizing the management of multireservoir systems under shifting flow regimes. *Water Resources Research*, *58*(6), e2021WR030582.
12. Goldfeld, S. M., & Quandt, R. E. (1973). A Markov model for switching regressions. Journal of Econometrics, 1(1), 3-15.
13. Hamilton J.D. (1989). A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle. Econometrica 57: 357-384
14. Hamilton, J.D. (1994). Time Series Analysis. Princeton University Press.
15. Hamilton, J. D. (2018). Regime switching models. In *The new palgrave dictionary of economics* (pp. 11421-11426). Palgrave Macmillan, London.
16. He, C., Liu, Z., Wu, J., Pan, X., Fang, Z., Li, J., & Bryan, B. A. (2021). Future global urban water scarcity and potential solutions. *Nature communications*, *12*(1), 4667.
17. Hemati, A., Rippy, M. A., Grant, S. B., Davis, K., & Feldman, D. (2016). Deconstructing demand: the anthropogenic and climatic drivers of urban water consumption. *Environmental Science & Technology*, *50*(23), 12557-12566.
18. Lu, Z. Q., & Berliner, L. M. (1999). Markov switching time series models with application to a daily runoff series. *Water Resources Research*, *35*(2), 523-534.
19. MAHAVARPOUR, ZAHRA. (2018). Forecasting of Heat Wave Occurrence and its Return Time in Iran Using Markov Chain Model. GEOGRAPHIC SPACE, 18(63), 71-86. SID. https://sid.ir/paper/91605/en
20. Mallakpour, I., & Villarini, G. (2016). A simulation study to examine the sensitivity of the Pettitt test to detect abrupt changes in mean. *Hydrological Sciences Journal*, *61*(2), 245-254.
21. Perlin, M. (2007). 'Estimation, Simulation and Forecasting of a Markov Switching Regression', (General case in Matlab).
22. Peterson, T. J., Saft, M., Peel, M. C., & John, A. (2021). Watersheds may not recover from drought. *Science*, *372*(6543), 745-749.
23. Pettitt, A. N. (1979). A non-parametric approach to the change-point problem. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, *28*(2), 126-135.
24. Sajadifar, S., & Pakseresht, M. (2024). A probabilistic Markov chain model for short-term water demand forecasting. *Journal of Water Management Modeling*, *32*.
25. Santos, M. C., Borges, A. I., Carneiro, D. R., & Ferreira, F. J. (2021, June). Synthetic dataset to study breaks in the consumer's water consumption patterns. In *Proceedings of the 2021 4th International Conference on Mathematics and Statistics* (pp. 59-65).
26. Tong, H. (1978). On a threshold model. *Pattern recognition and signal processing*, 575-586.

27. Turner, S. W. D., & Galelli, S. (2016). Regime-shifting streamflow processes: Implications for water supply reservoir operations. *Water Resources Research*, *52*(5), 3984-4002.
28. Zeynali, B., & Rouhi, F. (2023). Estimation of the probability of occurrence of heat waves with different periods of continuation in Ardabil province.