ENVIRONMENTAL RESEARCH
COMMUNICATIONS

INCLUSIVE PUBLISHING TRUSTED SCIENCE

Part of the ENVIRONMENTAL RESEARCH SERIES

PURPOSE-LED PUBLISHING™

PAPER • OPEN ACCESS

Statistical case study analysis of the quantity of water resources in Azerbaijan: Absheron Peninsula

To cite this article: Elshan Karimov *et al* 2026 *Environ. Res. Commun.* **8** 075031

IOP Publishing Environ. Res. Commun. 8 (2026) 075031 https://doi.org/10.1088/2515-7620/ae8f15

Environmental Research Communications

OPEN ACCESS

RECEIVED
22 February 2026

REVISED
13 July 2026

ACCEPTED FOR PUBLICATION
22 July 2026

PUBLISHED
29 July 2026

PAPER

Statistical case study analysis of the quantity of water resources in Azerbaijan: Absheron Peninsula

Elshan Karimov[1], Mustafa Sundu[2,*], Mertol Tüfekci[3,4,*], İnci Pir[5], Sevgi Güneş Durak[6], Güler Türkoğlu Demirkol[1] and Neşe Tüfekci[1]

[1] Faculty of Engineering, Department of Environmental Engineering, Istanbul University-Cerrahpaşa, Avcilar, Istanbul 34320, Türkiye
[2] Department of Management Information Systems İstanbul, İstinye University, Faculty of Economics, Administrative and Social Sciences, Istanbul, Türkiye
[3] Centre for Engineering Research, University of Hertfordshire, College Lane Campus, Hatfield AL10 9AB, United Kingdom
[4] School of Physics, Engineering and Computer Science, University of Hertfordshire, College Lane Campus, Hatfield AL10 9AB, United Kingdom
[5] Faculty of Mechanical Engineering, Istanbul Technical University, Gumussuyu, Istanbul 34437, Türkiye
[6] Department of Environmental Engineering, Faculty of Engineering-Architecture, Nevsehir Haci Bektas Veli University, Nevsehir 50300, Türkiye
* Authors to whom any correspondence should be addressed.

E-mail: mustafa.sundu@istinye.edu.tr and m.tufekci@hertss.ac.uk

Keywords: Absheron Peninsula, water sources, time series models, Prophet model, statistical analysis

Abstract

This study develops long-term water supply and demand forecasting models for the Absheron Peninsula of Azerbaijan, a semi-arid region with 4.0 million inhabitants entirely dependent on external water sources. Daily production and consumption data from 2016–2024 are analysed using three complementary specifications: ARMA(2,1) on the level series as a conservative parametric baseline, ARIMA(2,1,1) as a comparison model for assessing the effect of first differencing, and Prophet with automatic trend and seasonality decomposition. A seasonal-naive baseline ($y_t = y_{t-365}$) serves as a non-parametric reference, and forecasts are validated through rolling-origin cross-validation across five hold-out years (2020–2024). Strong annual seasonality is present, with 30%–35% amplitude variation between summer peaks and winter troughs. On the 2024 hold-out, Prophet attains the strongest near-term accuracy (RMSE 0.75 m^3 s^{-1}), followed by the seasonal-naive baseline (1.07 m^3 s^{-1}); ARMA(2,1) on level (2.30 m^3 s^{-1}) outperforms ARIMA(2,1,1) (3.20 m^3 s^{-1}) among the ARIMA-family models and is therefore used as the conservative parametric baseline. Long-term projections to 2050 diverge sharply: ARMA(2,1) yields a stable annual surplus of approximately 4.9 Mm^3 yr^{-1} (≈0.87% of supply), ARIMA(2,1,1) approximately 2.8 Mm^3 yr^{-1}, and Prophet an expanding surplus reaching 14.8 Mm^3 yr^{-1} by 2050. When the documented 35%–37% transmission and distribution losses are incorporated as a stress-test scenario, all three models project structural deficits of approximately 173–283 Mm^3 yr^{-1} by 2050. The findings underscore the need for multivariate forecasting frameworks, integrated hydrological modelling, and adaptive management strategies for long-term water security.

1. Introduction

Water-resource management is increasingly shaped by climate change, population growth, and industrial expansion. Climate change alters hydrological stability (Intergovernmental Panel on Climate Change (IPCC) 2023) through drought, depletion, contamination, and uneven spatial and temporal distribution of water resources (Schewe *et al* 2014) and through linked surface-water and groundwater responses (Taylor *et al* 2013). Severe water scarcity already affects large populations for part of each year (Mekonnen and Hoekstra 2016), and risks are greatest in regions already operating near scarcity thresholds (Vörösmarty *et al* 2010).

Table 1. Volumetric discharge rates of drinking water sources serving the Absheron Peninsula.

	Water Resource	Water Resource Class	Flow rate of the water source ($m^3 s^{-1}$)	Maximum Flow rate of the water source ($m^3 s^{-1}$)
1	Baku I-Şollar Wells	Underground water resources	1	1.27
2	Baku II-Haçmaz Well	Underground water resources	1.5	2.73
3	Oğuz-Gebele-Baku Pipeline	Underground water resources	2.5–3.5	5
4	Kura River	Surface Water Resources	1–3	6.2
5	Ceyranbatan Reservoir	Surface Water Resources	10–12	15
	Total		16–21	30.1

Observed global warming is already altering climatic baselines (Lüthi *et al* 2008, Tierney *et al* 2020, Calvin *et al* 2023, Intergovernmental Panel on Climate Change (IPCC) 2023), with measurable effects on the hydrological cycle (Huntington 2006, Trenberth *et al* 2014).

Even under Paris-aligned mitigation pathways, warming-related water risks remain substantial (Intergovernmental Panel on Climate Change (IPCC), 2022, United Nations Environment Programme 2022). Exposure to water stress rises sharply between 1.5 °C and 2 °C of warming (Schleussner *et al* 2016, Döll *et al* 2018).

Consequently, the conservation and strategic management of water resources are of critical importance to Azerbaijan, mirroring the global urgency but with unique regional characteristics. Azerbaijan, situated in the South Caucasus region at the crossroads of Eastern Europe and Western Asia, faces particular challenges related to its semi-arid climate, transboundary water dependencies, and socioeconomic development pressures (Ahmadov 2020). The country's water resources are highly vulnerable to climate variability, with approximately 70% of surface water originating from transboundary rivers, including the Kura and Aras (Huseynov *et al* 2025). This dependence on external water sources compounds the challenges posed by climate change, as upstream activities and climate impacts in neighbouring countries directly affect Azerbaijan's water availability (Klaphake and Kramer 2011). Recent studies have documented concerning trends in Azerbaijan's hydrological regime, including decreasing river discharge in major basins, increasing frequency of low-flow periods, and shifts in the timing of snowmelt-driven peak flows (Zarghami *et al* 2011). These changes have critical implications for agriculture, which accounts for approximately 60%–70% of total water consumption in the country (Ahmadov 2020). Furthermore, the Caspian Sea, which borders Azerbaijan and is a critical water resource for the nation, has experienced significant instability influenced by both climate change and human-driven factors (Chen *et al* 2017). Projections suggest continued instability in Caspian Sea levels, with potential implications for coastal ecosystems, infrastructure, and water quality (Kosarev 2005).

The Absheron Peninsula is the region in Azerbaijan where water scarcity challenges are most widely observed (Namazov *et al* 2024). Located on the western coast of the Caspian Sea, the peninsula encompasses 2110 km^2 and supports a population of approximately 3.6–4.6 million (Ismaylov *et al* 2015). This demographic concentration accounts for more than 30% of Azerbaijan's total population and approximately 20% of the total population within the broader Caspian region (Pasha *et al* 2023).

Systematic assessment of Azerbaijan's freshwater resources started during the 1970s and 1980s (Aliyev and Veliyeva 2024). Evaluations conducted during this period quantified the country's total freshwater reserves at 35.6 billion m^3, comprising 9 billion m^3 of groundwater, 6 billion m^3 of indigenous river runoff, and 20.6 billion m^3 of transboundary surface waters (Pasha *et al* 2023). Despite accommodating a significant proportion of the national population (with Baku and its surrounding areas representing approximately 25% of the population), the region accounts for only 6% of national water consumption, totalling approximately 744 million m^3 annually (General Country Assessment—Azerbaijan: General Water Security Assessment n.d.). Potable water for the Absheron Peninsula is sourced from five primary origins, with their respective flow rates presented in table 1.

In response to these challenges, Azerbaijan must prioritise implementing integrated water resources management approaches that account for climate uncertainty, enhance water-use efficiency, and promote adaptive capacity (Ahmadov 2020). This includes investing in modern irrigation technologies, improving water monitoring and forecasting systems, strengthening institutional frameworks for transboundary cooperation, and developing climate-resilient water infrastructure (World Bank 2024). Additionally,

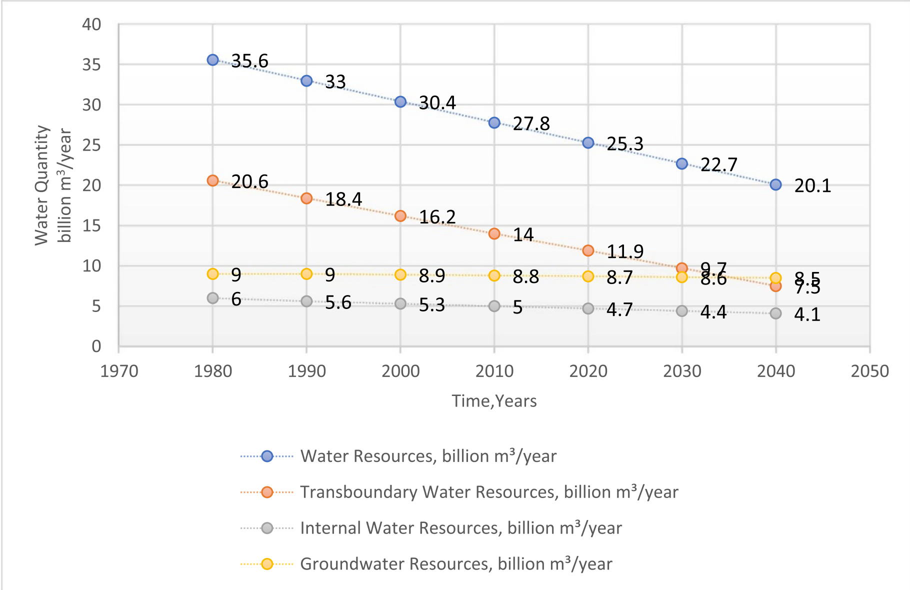


Figure 1. Estimated indicators of water resources in the Republic of Azerbaijan.

nature-based solutions such as wetland restoration, reforestation in critical watersheds, and soil conservation measures can provide cost-effective means of enhancing water security while simultaneously contributing to climate change mitigation and biodiversity conservation objectives (Ramírez-Agudelo *et al* 2020).

Climate change is increasingly affecting water resources and requires coordinated management at multiple scales. For Azerbaijan, located in a climate-sensitive region, strategic adaptation is necessary for sustainable development. As illustrated in figure 1, the country's freshwater resources are in significant decline. This trend is driven by climate change, pollution, and upstream activities by neighbouring countries.

Trans-boundary surface water resources decline by 42.2%, falling from 20.6 billion m^3 to 11.9 billion m^3. Similarly, local surface water resources decrease by 21.7% (from 6 billion m^3 to 4.7 billion m^3), and groundwater reserves decrease by 3.3% (from 9 billion m^3 to 8.7 billion m^3). Consequently, total water resources diminished by 28.9%, dropping from 35.6 billion m^3 to 25.3 billion m^3.

Significant reductions are particularly observable in the Kura and Aras rivers. Between 2016 and 2020, the average annual inflow to the Mingachevir reservoir fell from 344 $m^3\ s^{-1}$ to 192 $m^3\ s^{-1}$ (a 44.2% decrease), while the inflow to the Mil-Mugan Hydroelectric Power Plant dropped from 134 $m^3\ s^{-1}$ to 99 $m^3\ s^{-1}$ (a 26.1% decrease). As a direct result of this decline, hydro-ecological crises occurred between 2020 and 2022; the Kura River failed to reach the Caspian Sea, and conversely, seawater intrusion was observed moving upstream along the Kura riverbed as far as the Salyan region. If these declining trends continue, water reserves are projected to decrease by an additional 21.2% by 2040 compared to 2022 levels, falling to 20.1 billion m^3. Under this scenario, transboundary surface water resources are expected to decline by 39% to 7.5 billion m^3, and local surface water resources by 14.6% to 4.1 billion m^3. This yields a combined surface water potential of 11.6 billion m^3; however, after accounting for ecological flow requirements, only 9.3 billion m^3 is forecast to be available for withdrawal.

The Absheron Peninsula is a semi-arid region characterised by limited local water resources and low annual precipitation (approximately 200–250 mm). The peninsula lacks perennial rivers, and a significant portion of its groundwater is saline or polluted due to geological factors, restricting its use for drinking or irrigation. Consequently, the water supply in Absheron is entirely dependent on external sources. Currently, the daily water demand of approximately 2 million m^3 for Baku and surrounding settlements is met by pumping water from neighbouring regions and river basins. Nearly 70% of this supply is provided by the Samur-Absheron canal system originating from the north of the country, with the remainder supplied by the Kura River and the Oguz-Gabala-Baku pipeline. Even local reservoirs, such as Lake Ceyranbatan, are primarily fed by a distant basin.

The impact of climate change on the Absheron Peninsula is critical given this heavy dependence on external resources. Flows in the Samur River are declining due to climate change and increased use in neighbouring regions. Research indicates that withdrawals from the Samur River may decrease in the future based on climate scenarios, potentially leading to water deficits in Absheron (Karthe *et al* 2015, Ahadov *et al* 2025). Similarly, the volume of water pumped from the Kura River to Baku is directly affected by droughts in the Kura basin; reduced precipitation in Armenia and Georgia due to climate factors is expected to reduce inflows via the Kura River further.

This study addresses these gaps by developing and systematically comparing complementary forecasting specifications for the Absheron Peninsula's water system. Using daily water production and consumption data from 2016–2024, comprising 3286 observations per series, this study evaluates an ARMA(2,1) specification fitted to the level series as the conservative ARIMA-family baseline. ARIMA(2,1,1) is included as a comparison model to assess the effect of first differencing, Prophet models are developed with automatic trend and seasonality decomposition, and a seasonal-naive baseline, $y_t = y_{t-365}$, is included as a non-parametric reference. Models are selected through Box-Jenkins methodology, formal stationarity testing, held-out validation against the seasonal-naive baseline, and rolling-origin cross-validation. All specifications generate 26 year daily forecasts from 2025 through 2050, providing long-range projections at policy-relevant timescales. The comparative framework explicitly quantifies forecast divergence arising from different structural assumptions about trend persistence (ARMA(2,1) mean reversion versus Prophet trend extrapolation) and uncertainty propagation mechanisms (bounded ARMA prediction intervals versus unbounded ARIMA bands versus Prophet posterior bands). This multi-specification approach yields not a single point prediction but a range of plausible futures bounded by contrasting modelling philosophies, thereby characterising the uncertainty associated with multi-decade water system projections. Analysis focuses on three dimensions: (1) seasonal pattern identification and operational planning implications, (2) long-term supply-demand balance projections under alternative trend assumptions, and (3) forecast uncertainty quantification through prediction intervals and model divergence. Results inform strategic planning for a region representing a relevant case: water-scarce, externally dependent, climate-vulnerable, and undergoing rapid urbanisation.

This work contributes to water resources forecasting literature in several ways. First, it provides one of the first comparative daily time-series forecasting analyses for the Absheron Peninsula water system, complementing earlier qualitative assessments and resource-management overviews of Azerbaijan's water sector by quantifying multi-decadal supply–demand projections at daily resolution (Ahmadov 2020, Pasha *et al* 2023, World Bank 2024). Second, the explicit comparative framework that treats model divergence as information rather than error provides a practical basis for comparing long-range forecasts amid substantial uncertainty. Third, by quantifying the approximately 2.8–14.8 $Mm^3\ yr^{-1}$ range of annual surplus projections arising from the ARIMA(2,1,1), ARMA(2,1) on level ($\approx$4.9 $Mm^3\ yr^{-1}$), and Prophet specifications, the analysis demonstrates that modelling choices can matter as much as parametric uncertainty—a finding with direct policy relevance for robust decision-making. Fourth, the study's focus on external-dependent systems addresses an underrepresented context in water forecasting literature, which predominantly examines self-sufficient or partially self-sufficient regions. Finally, the integration of daily-to-annual aggregation with uncertainty interval analysis shows how daily statistical forecasts can be converted into annual planning metrics for water-resource assessment.

2. Materials and methods

2.1. Study area

The Absheron Peninsula is situated in the eastern region of the Republic of Azerbaijan along the western coast of the Caspian Sea. The peninsula encompasses an area of 2110 km^2 and supports a population of approximately 4.0 million, representing more than 30% of Azerbaijan's total population and approximately 20% of the total population within the broader Caspian region. This demographic concentration makes the peninsula a critical region for national water resource management.

The region is characterised by a semi-arid climate with low precipitation, with annual rainfall ranging between 200 and 250 mm. The peninsula lacks permanent rivers, and a substantial proportion of groundwater reserves are saline or contaminated by geological structures, rendering them unsuitable for potable or agricultural use. Consequently, the water supply for the Absheron Peninsula, including the capital city Baku and surrounding areas, is entirely dependent on external water sources.

The current water supply to the region is provided primarily through two major external sources: water pumped from the Kura River and water transferred from the Samur River system. The daily water demand of approximately 2 million m^3 for Baku and surrounding settlements is met through this external infrastructure. This heavy dependence on transboundary water resources makes the region particularly vulnerable to upstream water use changes and climate variability, highlighting the critical importance of accurate long-term water supply-demand forecasting for strategic planning and infrastructure development.

2.2. Data collection and preprocessing

Daily water production and consumption data for the Absheron Peninsula were obtained from the operational records of the regional water utility serving the peninsula, covering the period from 1 January 2016 to 29 December 2024. The raw records consist of daily aggregated discharge from the five primary water sources serving the peninsula (table 1) and the corresponding daily consumption recorded for the Baku, Sumqayit, Absheron, and outlying districts. Two aggregate variables were constructed for the present analysis: Total Water Supply (the sum of daily discharge across all five sources) and Total Consumption (the sum of daily consumption across all four districts, encompassing residential, industrial, and commercial uses). All values are expressed as daily means in $m^3\ s^{-1}$.

Preprocessing consisted of (i) verifying the continuity of the daily index across the full 2016–2024 window, (ii) correcting a small number of evident data-entry artefacts in individual cells (a comma decimal separator and a stray trailing period) prior to numerical conversion, and (iii) confirming the absence of missing days. The cleaned dataset contains 3 286 daily observations for each series with no gaps. Total Water Supply values range from 13.6 to 23.9 $m^3\ s^{-1}$ (mean = 17.7$m^3\ s^{-1}$, SD = 2.5 $m^3\ s^{-1}$), and Total Consumption values exhibit a similar range (13.5–23.6 $m^3\ s^{-1}$; mean = 17.6 $m^3\ s^{-1}$, SD = 2.5 $m^3 s^{-1}$); the close correspondence between production and consumption magnitudes suggests effective short-term supply–demand matching.

2.3. Statistical analysis methods

Four time series forecasting specifications were employed to generate and benchmark long-term water supply and demand projections: an ARMA(2,1) model fitted to the level series as a conservative parametric baseline, an ARIMA(2,1,1) comparison model used to assess the effect of first differencing, the Prophet forecasting framework, and a seasonal-naive baseline. These were selected for their complementary strengths—the ARIMA-family models for short-term autocorrelation structure, Prophet for flexible trend and seasonality modelling, and the seasonal-naive baseline as a non-parametric reference.

2.3.1. ARIMA-family models

ARIMA is a univariate time series modelling technique that combines three components to capture temporal dependencies and forecast future values. The general ARIMA(p,d,q) model is specified by three integer parameters: p (the order of the autoregressive (AR) component), d (the degree of differencing required to achieve stationarity), and q (the order of the moving average (MA) component).

The *AR* component of order p models the current observation as a linear combination of p previous observations. Mathematically, the AR(p) process is expressed as:

$$y_{\mathrm{t}} = \varphi_1 y_{t-1} + \varphi_2 y_{t-2} + \cdots + \varphi_p y_{t-p} + \varepsilon_t \tag{1}$$

where y_t represents the value at time t, φ_1, φ_2, .., φ^p are AR coefficients, and ε_t is white noise error.

The *integrated (I)* component of degree d addresses non-stationarity through differencing. First-order differencing ($d = 1$) is computed as:

$$\Delta y_t = y_t - y_{t-1} \tag{2}$$

which removes linear trends. Higher-order differencing may be applied for more complex non-stationary patterns, though d rarely exceeds 2 in practice.

The *MA* component of order q models the current observation as a linear combination of current and past forecast errors:

$$y_t = \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \cdots + \theta q \varepsilon_{t-q} + \varepsilon_t \tag{3}$$

where θ_1, θ_2, …, θ_q are MA coefficients.

The combined ARIMA(p,d,q) model is applied to the differenced series and can be written in compact form using the backshift operator B as:

$$\varphi(B)(1-B)^d y_t = \theta(B)\varepsilon_t \tag{4}$$

where B is the backshift operator defined by $B^j y_t = y_{t-}{}^j$ (i.e. $By_t = y_{t-1}$, $B^2 y_t = y_{t-2}$, and so on), and $\varphi(B)$ and $\theta(B)$ are polynomials of orders p and q, respectively.

Model identification and parameter selection followed the Box-Jenkins methodology. Stationarity was first assessed through visual inspection of time series plots and formal testing using autocorrelation function (ACF) and partial ACF (PACF) plots. The ACF at lag k is defined as:

$$\rho_k = \mathrm{Cov}\left(y_t, y_{t-k}\right)/\mathrm{Var}\left(y_t\right) \tag{5}$$

while the PACF measures the correlation between observations at lag k after removing the effects of intermediate lags.

For non-stationary series exhibiting slow ACF decay, first differencing ($d = 1$) was applied. The PACF plot was used to identify the AR order p (significant spikes at lags 1 through p followed by cutoff), while the ACF plot of the differenced series indicated the MA order q (significant spikes at lags 1 through q followed by cutoff).

Candidate models were compared using information criteria, specifically the Akaike information criterion (AIC) and Bayesian information criterion (BIC), defined as:

$$\mathrm{AIC} = -2\ln\left(L\right) + 2k \tag{6}$$

$$\mathrm{BIC} = -2\ln\left(L\right) + k\cdot\ln\left(n\right) \tag{7}$$

where L is the maximised likelihood, k is the number of estimated parameters, and n is the sample size. Lower AIC and BIC values indicate superior model performance, balancing goodness-of-fit against model complexity.

Model validation was performed through residual diagnostics. Standardised residuals were examined for: (1) zero mean and constant variance, (2) absence of autocorrelation (white noise property), verified through ACF plots and Ljung–Box tests with the test statistic:

$$Q = n\left(n+2\right)\sum_{m=1}^{k}\left(\rho_m^2/\left(n-m\right)\right) \tag{8}$$

where n is the sample size, ρ_m is the sample autocorrelation at lag m, and k is the number of lags tested. Under the null hypothesis of white noise, Q follows a χ^2 distribution with k–p–q degrees of freedom. Normality of residuals was assessed through histograms and Q–Q plots. All analyses reported in this paper—formal unit-root testing, ARMA(2,1) on the level series as the conservative parametric baseline, SARIMA variants, ARIMAX with calendar exogenous regressors, rolling-origin cross-validation, and Prophet hyperparameter sensitivity—were implemented in Python (statsmodels v0.14, prophet v1.3) for transparency and reproducibility.

2.3.2. Prophet model

Prophet is an additive regression model developed by Facebook's Core Data Science team for business time series forecasting (Taylor and Letham 2018). The model decomposes time series into trend, seasonal, and holiday components, providing interpretable forecasts with flexible handling of missing data, outliers, and structural breaks.

The Prophet model is specified as

$$y\left(t\right) = g\left(t\right) + s\left(t\right) + h\left(t\right) + \varepsilon_t \tag{9}$$

where $g(t)$ represents the trend function modelling non-periodic changes, $s(t)$ represents periodic changes (seasonality), $h(t)$ represents the effects of holidays or special events, and ε_t represents the error term assumed to be normally distributed.

Trend modelling: Prophet employs a piecewise linear growth model with automatic changepoint detection. The trend component is defined as:

$$g\left(t\right) = \left(k + a(t)^{\mathrm{T}}\delta\right)t + \left(m + a(t)^{\mathrm{T}}\gamma\right) \tag{10}$$

where k is the growth rate, δ contains rate adjustments at changepoints, m is the offset parameter, and $a(t)$ is a vector of indicator variables, where the jth element $a_j\left(t\right) = 1$ if $t \geqslant s_j$ and 0 otherwise, with s_j

denoting the location of the jth changepoint. Changepoints are automatically selected from the first 80% of the time series, though manual specification is also supported.

Seasonality modelling: Periodic effects are represented using Fourier series. For a period P (e.g. $P = 365.25$ for yearly seasonality or $P = 7$ for weekly seasonality), the seasonal component is:

$$s(t) = \sum_{n=1}^{N} [a_n \cos(2\pi nt/P) + b_n \sin(2\pi nt/P)] \tag{11}$$

where N determines the number of Fourier terms and thus the flexibility of seasonality, and a_n and b_n are the Fourier coefficients estimated from the data. Prophet can model multiple seasonalities simultaneously (e.g. both weekly and annual patterns), with parameters estimated using least squares regression.

Holiday and event effects: Irregular events are incorporated through indicator functions:

$$h(t) = \sum_{i=1}^{L} \kappa_{\mathrm{i}}.\mathbb{1}\{t \in D_{\mathrm{i}}\} \tag{12}$$

where κ_i represents the impact of the ith holiday, D_i is the set of dates for that holiday (including surrounding windows), and $\mathbb{1}\{t \in D_i\}$ is an indicator function that equals 1 when time t falls within the window of holiday i (i.e. $t \in D_i$) and 0 otherwise. For this study, no specific holiday effects were incorporated, as water supply and demand patterns were primarily driven by seasonal weather variations rather than discrete events.

Model estimation: Prophet uses a Bayesian framework with Stan for parameter estimation, providing posterior distributions that quantify uncertainty in forecasts. The model generates prediction intervals at specified confidence levels (typically 80% and 95%), accounting for uncertainty in trend, seasonality, and observation noise.

In this study, Prophet models were specified with a linear trend, the default automatic changepoint detection (25 potential changepoints), and both yearly and weekly seasonality components; daily seasonality was deliberately excluded, as preliminary analyses showed that it did not improve forecast accuracy. Prophet models were fitted in Python (prophet v1.3).

2.4. Model application and forecasting procedure

Both ARIMA and Prophet models were fitted separately to the Total Water Supply and Total Consumption series using data from 1 January 2016, to 29 December 2024 (training period). Models were then used to generate daily forecasts for the period from 1 January 2025, to 31 December 2050, providing a 26 year forecast horizon suitable for strategic water infrastructure planning.

For ARIMA models, the complete Box-Jenkins procedure was followed: (1) stationarity assessment and differencing, (2) ACF/PACF analysis for parameter identification, (3) maximum likelihood estimation of candidate models, (4) model selection using AIC and BIC, and (5) residual diagnostics for validation. Daily forecasts were generated with prediction intervals calculated using the estimated parameter covariance matrix.

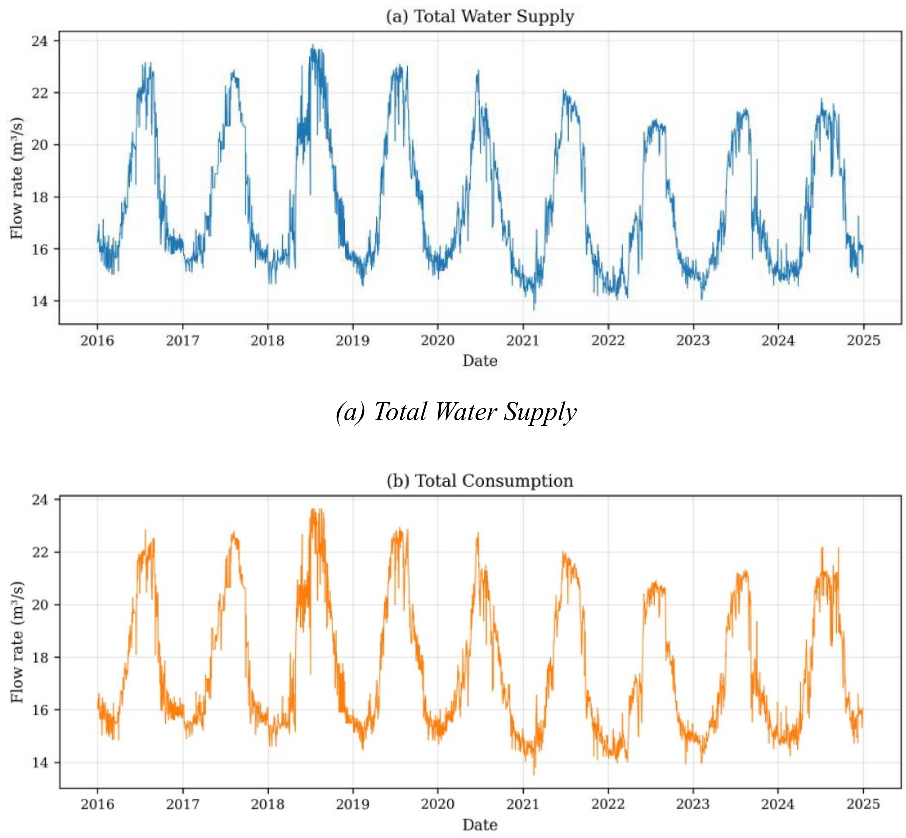
For Prophet models, the training data were provided in the required format (date column 'ds' and target variable column 'y'), and models were fitted with default priors adjusted only for seasonality components. Forecasts were generated using the model's built-in prediction method, which provides point forecasts and uncertainty intervals based on posterior sampling.

Daily forecasts from both models were aggregated to annual totals to facilitate interpretation and comparison with water resource planning horizons. The annual supply-demand balance was calculated as:

$$\Delta S_{\mathrm{a}} = \Sigma_{=1}^{\mathrm{d}365} P^{\mathrm{d}} - \Sigma_{=1}^{\mathrm{d}365} C^{\mathrm{d}} \tag{13}$$

where ΔS_{a} is the annual surplus/deficit ($\mathrm{m}^3\ \mathrm{yr}^{-1}$), and P^{d} and C^{d} are daily production and consumption volumes ($\mathrm{m}^3\ \mathrm{d}^{-1}$) computed from the daily mean flow rates ($\mathrm{m}^3\ \mathrm{s}^{-1}$) by multiplication by 86 400 s d^{-1}. Positive values indicate supply surplus, while negative values would indicate potential water scarcity requiring intervention. This metric provides a strategic indicator for long-term water security assessment.

All analyses reported in this paper were performed in Python 3.12. ARIMA-family models were fitted with statsmodels (v0.14) and Prophet models with prophet (v1.3); data manipulation used pandas and NumPy, and all figures were produced with Matplotlib.



(a) Total Water Supply

(b) Total Consumption

Figure 2. Daily water production and consumption time series (2016–2024).

3. Results and discussion

3.1. ARIMA-family model analysis

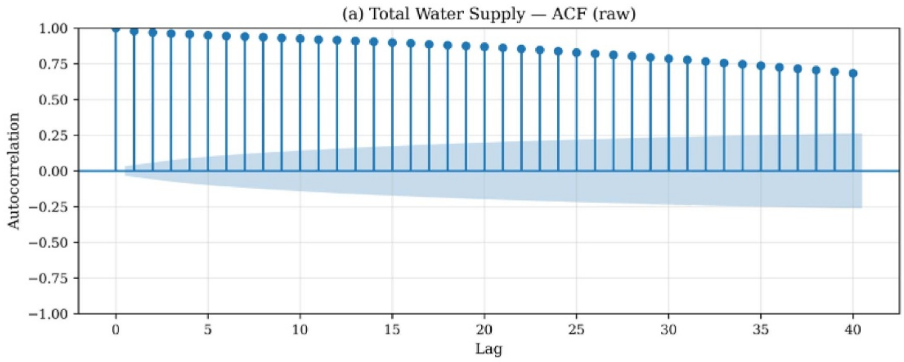
3.1.1. Data characteristics and preprocessing

The dataset consists of daily water production and consumption values measured between 1 January 2016, and 29 December 2024. Total Water Supply (total daily production) and Total Consumption (total daily consumption) were used as the primary variables for analysis. Following data cleaning procedures, including removal of NA values and conversion of numeric formats, a continuous daily time series of 3286 observations was obtained for each variable. Examination of the data revealed that Total Water Supply and Total Consumption ranged from approximately 15–23 m^3 s^{-1}, with consumption comparable to production.

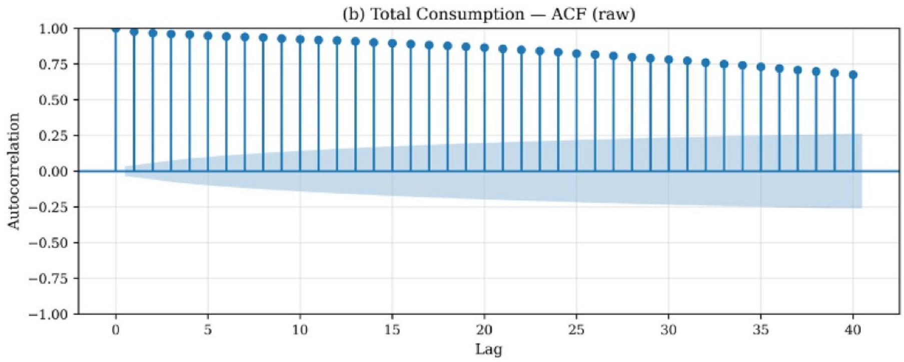
3.1.2. Time series structure

Visual inspection of the raw time series, as displayed in figure 2, revealed several key structural features essential for model specification. Both the Total Water Supply and Total Consumption series fluctuated around a relatively stable level, with no discernible long-term deterministic trend, although short-term fluctuations were evident in certain periods. This pattern indicated strong persistence and motivated formal stationarity testing before deciding whether differencing was required.

Distinct and recurring seasonal fluctuations were observed across both series, with regular patterns repeating during specific periods of the year. Production and consumption both increased markedly during the summer months (June–August), likely driven by elevated cooling and irrigation demands, whereas they remained lower and more stable during winter. The amplitude of seasonal variation was approximately 5–7 m^3 s^{-1} between seasonal peaks and troughs, representing roughly 30%–35% of the mean level.



(a) Total Water Supply



(b) Total Consumption

Figure 3. Autocorrelation function (ACF) plots of raw time series.

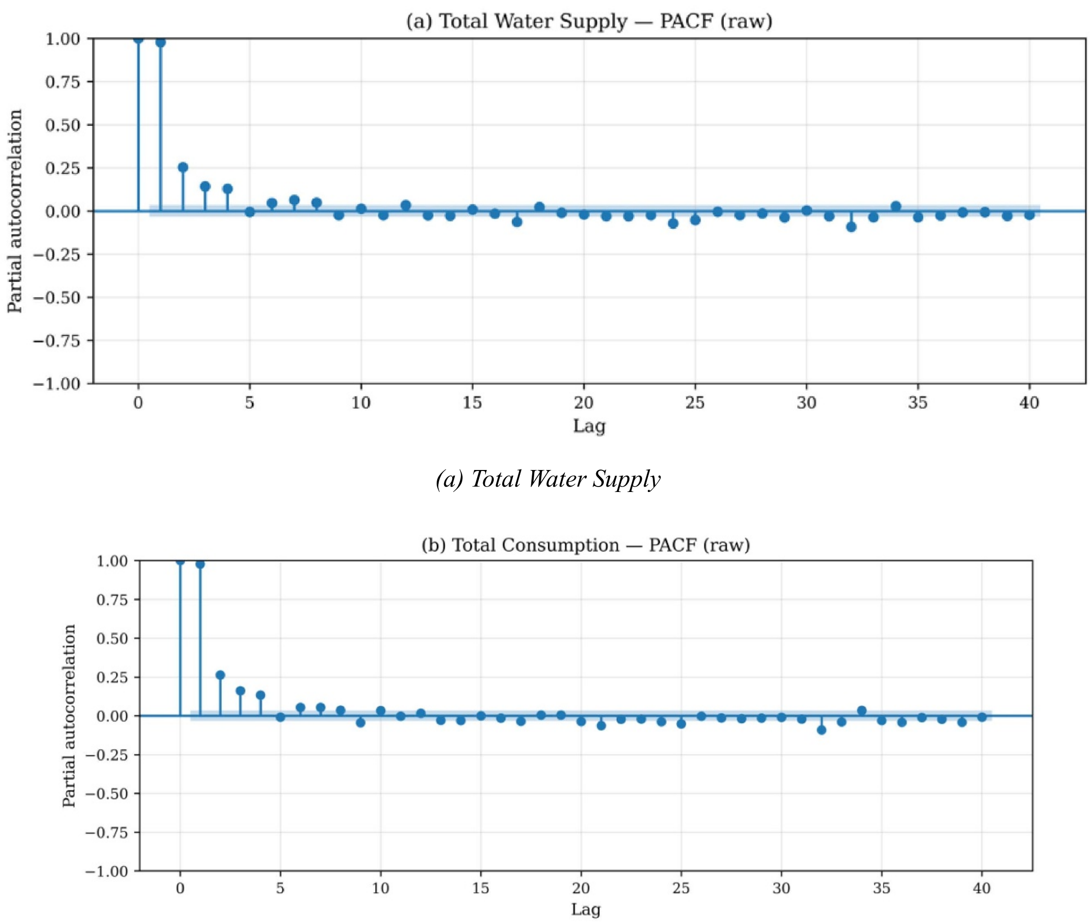
High synchronicity between the two series was evident throughout the observation period, with increases and decreases in consumption largely mirrored by corresponding changes in production patterns. This tight coupling indicates coordinated short-term water-supply and demand management, suggesting that production is actively adjusted to meet daily consumption requirements. Sharp spikes and sudden drops occasionally appeared in the daily frequency data, particularly during transition periods between seasons, indicating short-term volatility that must be accounted for in model error terms. These irregularities, while relatively infrequent, highlight the influence of weather anomalies and system adjustments on daily water flows.

3.1.3. Stationarity testing and differencing

ACF analysis revealed that both series exhibited high positive ACF values at initial lags, which decreased gradually and regularly as the lag increased as shown in figure 3. For Total Water Supply, the ACF remained above 0.75 for the first 10 lags, declining to approximately 0.50 at lag 20 and 0.25 at lag 30. Total Consumption demonstrated nearly identical behaviour, with ACF values tracking those of production within ± 0.05 at most lags.

This slow, hyperbolic decay pattern, combined with the absence of a sharp cutoff and statistical significance extending beyond lag 35, indicated strong persistence and motivated formal unit-root and stationarity testing; visual ACF behaviour alone was not treated as conclusive evidence of non-stationarity. In a stationary series, the ACF typically drops to near-zero within a few lags; the persistence observed here indicates changing mean levels over time. The substantial similarity between ACF patterns for Total Water Supply and Total Consumption suggested parallel short- and medium-term dynamics in supply and demand processes, reinforcing the observation of tight operational coupling between production and consumption.

Formal unit-root testing was conducted using the Augmented Dickey-Fuller (ADF) and Kwiatkowski-Phillips-Schmidt-Shin (KPSS) procedures. For Total Water Supply, the level series yielded ADF $= -3.91\,(p = 0.002)$ and KPSS $= 0.23(p > 0.10)$; for Total Consumption, ADF $= -3.96\,(p = 0.002)$ and KPSS $= 0.18(p > 0.10)$. Under the conventional decision rule (stationary if ADF $p < 0.05$ and KPSS $p > 0.05$), both series are stationary in levels by both tests, indicating that first differencing ($d = 1$)



(a) Total Water Supply

(b) Total Consumption

Figure 4. Partial autocorrelation function (PACF) plots of raw time series.

is not strictly required to achieve stationarity. After first differencing, the ADF statistics strengthen to -9.80 and -9.84 (both $p < 0.001$) and the KPSS statistics fall to 0.02, consistent with the differenced series remaining stationary. On the basis of these tests, ARMA(2,1) was fitted directly to the level series as the conservative parametric baseline, while ARIMA(2,1,1) was included as a comparison model to assess the effect of first differencing; the strongly negative lag-1 ACF of the differenced series ($\rho \approx -0.32$) further indicates that differencing is not appropriate for already-stationary data, while the moving-average term in ARIMA(2,1,1) only partly compensates for this specification issue.

PACF analysis identified statistically significant values at the first three lags for both series as presented in figure 4. For Total Water Supply, PACF values were approximately 0.98 at lag 1, 0.27 at lag 2, and 0.16 at lag 3, with all subsequent lags falling within the 95% confidence bounds. Total Consumption exhibited similar values: 0.98, 0.26, and 0.14 at lags 1, 2, and 3, respectively.

The extremely high PACF value at lag 1 (near unity) and the slowly decaying ACF visually suggested strong persistence consistent with—but not by itself diagnostic of—non-stationarity. The formal stationarity tests reported in the next subsection (section 3.1.3) resolve this ambiguity in favour of stationarity in levels. The significance of lags 2 and 3, though substantially smaller in magnitude, indicated a potential AR component of order 2 once the level series is modelled directly. The rapid decay after lag 3, with all values remaining within confidence intervals, provided no evidence supporting higher-order AR terms.

Although the high lag-1 PACF and slowly decaying ACF visually suggested strong persistence, the formal ADF and KPSS tests reported above indicated stationarity in levels. First differencing was therefore not required to achieve stationarity. The differenced series and its ACF/PACF diagnostics (figures 5–7) are reported only to document the consequences of the original ARIMA(2,1,1) specification and in particular to demonstrate the over-differencing pattern (lag-1 ACF≈ -0.32, characteristic of an unnecessarily differenced stationary series) rather than as evidence of a successful unit-root removal.

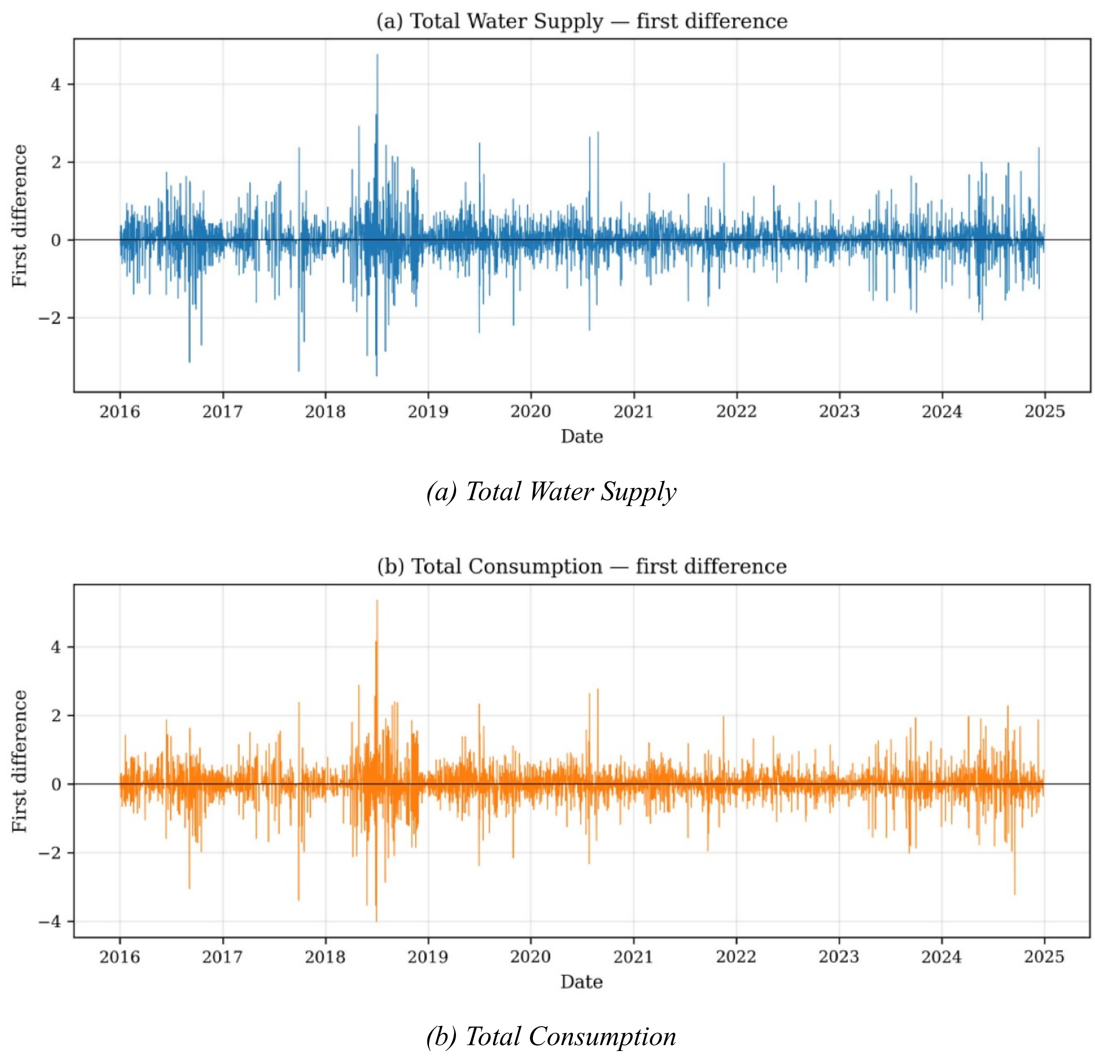


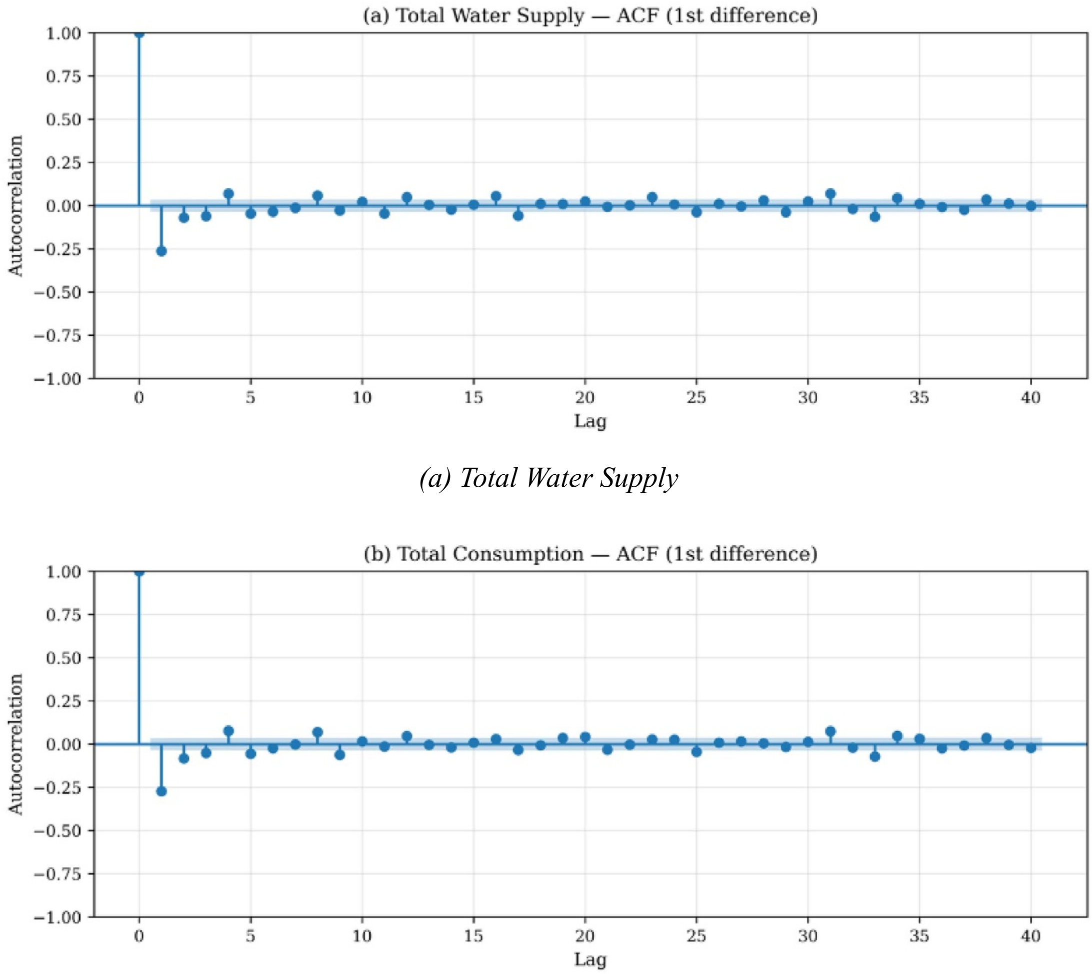
Figure 5. First-differenced time series.

Examination of figure 5 revealed that the differenced Water Supply series exhibited several key characteristics of stationarity. The data oscillated symmetrically around zero (mean≈0.001), with no apparent drift or long-run trend. Variance appeared relatively constant over time, with no obvious heteroscedasticity or systematic changes in volatility. Occasional large excursions (±2.5 to ±4 m^3 s^{-1}) occurred during specific periods, notably in early 2018 and mid-2019, corresponding to major system adjustments or unusual weather events. These outliers, while noteworthy, did not dominate the series behaviour and could be adequately captured by the model's error structure.

The differenced Consumption series, shown in figure 5, displayed characteristics nearly identical to those of production, with symmetric oscillation around zero, constant variance, and similar timing of major excursions. The close correspondence between figure 5 quantitatively confirmed the visual synchronicity observed in the raw series, demonstrating that day-to-day changes in consumption are tracked by corresponding changes in production. This parallel behaviour in the differenced series suggested that a single ARIMA specification would likely be appropriate for both variables.

Following first differencing, ACF and PACF plots were re-examined to identify appropriate model orders (figures 6 and 7). The ACF plots of the differenced series revealed a marked negative spike at lag 1 (approximately −0.32 for Water Supply and −0.28 for Consumption), followed by smaller positive and negative values at subsequent lags that remained predominantly within confidence bounds.

The strong negative autocorrelation at lag 1 in figure 6 is a diagnostic signature of overdifferencing or, more commonly in this context, the presence of a MA component. This pattern indicates that positive shocks tend to be followed by negative shocks and vice versa, a characteristic captured by MA(1) or higher-order MA models. The absence of additional large spikes beyond lag 1 suggested that a low-order MA process ($q = 1$ or $q = 2$) would be sufficient. Weak but statistically significant autocorrelations



(a) Total Water Supply

(b) Total Consumption

Figure 6. Autocorrelation function (ACF) plots of first-differenced series.

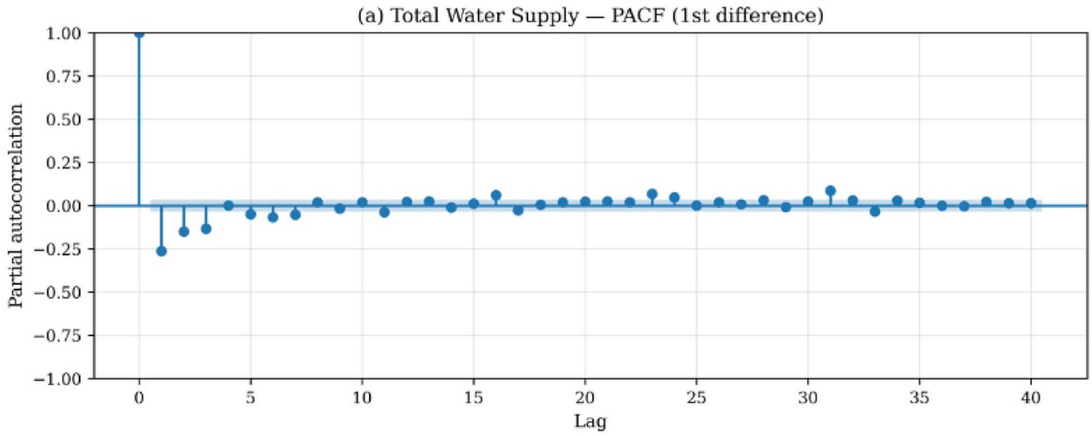
at several later lags (lags 7, 14, 21, 28) exhibited a subtle weekly pattern, reflecting minor residual periodicities not fully captured by simple differencing. However, these correlations were small in magnitude ($|\rho| < 0.05$) and were expected to be adequately handled by the MA component without requiring seasonal extensions.

PACF plots of the differenced series (figure 7) show statistically significant negative spikes at lags 1, 2, and 3 for both Water Supply and Consumption. Beyond lag 3, most partial autocorrelations fall within the 95% confidence bounds, although a few isolated higher-order lags exhibit small but statistically significant deviations without indicating a persistent or systematic structure.

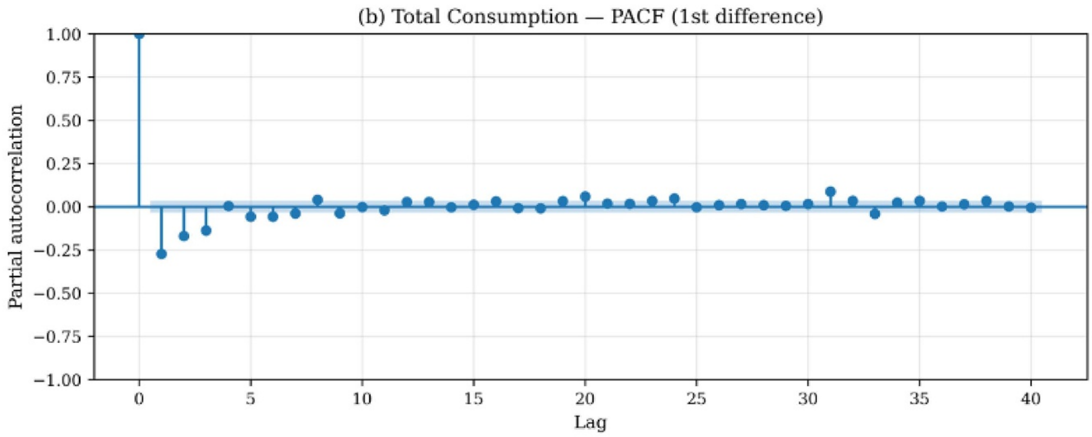
The pattern in figure 7 indicated an AR component of order 2 or 3. The significance of the first three lags suggested that today's value depends on values from the previous three days, even after accounting for the stochastic trend through differencing. Taken together, figures 6 and 7 guided the selection of candidate models ARIMA(2,1,1), ARIMA(3,1,1), ARIMA(2,1,2), and ARIMA(3,1,2) for formal evaluation.

3.1.4. Model selection and validation

Model comparison using AIC and BIC was conducted for candidate specifications. For Total Water Supply, ARIMA(2,1,1) performed better, with AIC = 4066.1 and BIC = 4090.0, compared with ARIMA(3,1,1), which had AIC = 4068.5 and BIC = 4098.4. Both models exhibited nearly identical error variance ($\sigma^2 \approx 0.24$), indicating that the additional AR parameter in ARIMA(3,1,1) increased model complexity without meaningful improvement in fit. Higher-order specifications ARIMA(2,1,2) and ARIMA(3,1,2) yielded marginally higher information criteria, confirming overparameterisation.



(a) Total Water Supply

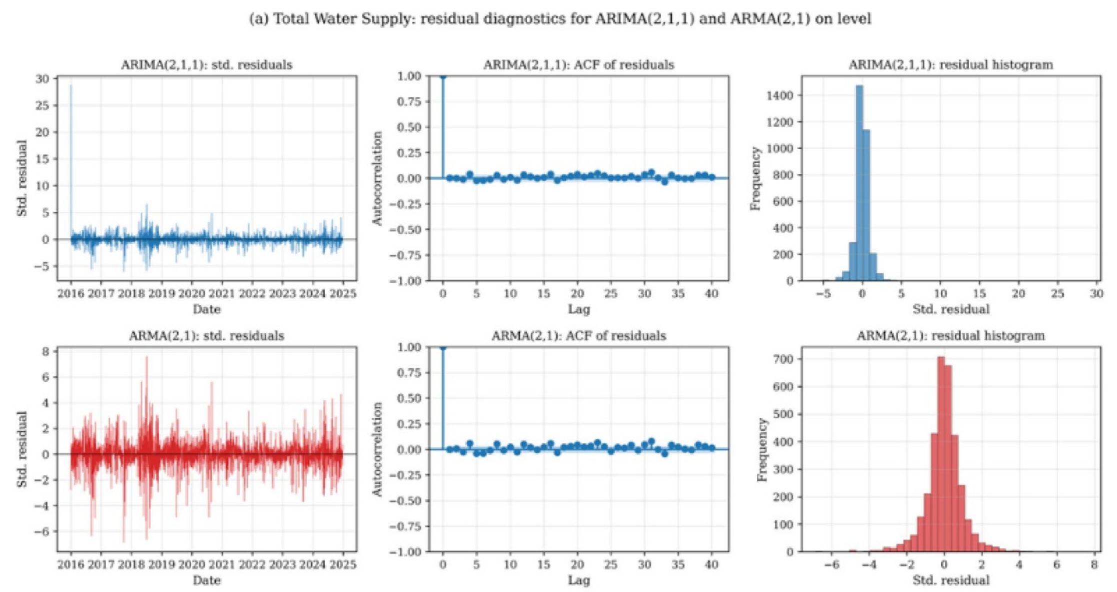


(b) Total Consumption

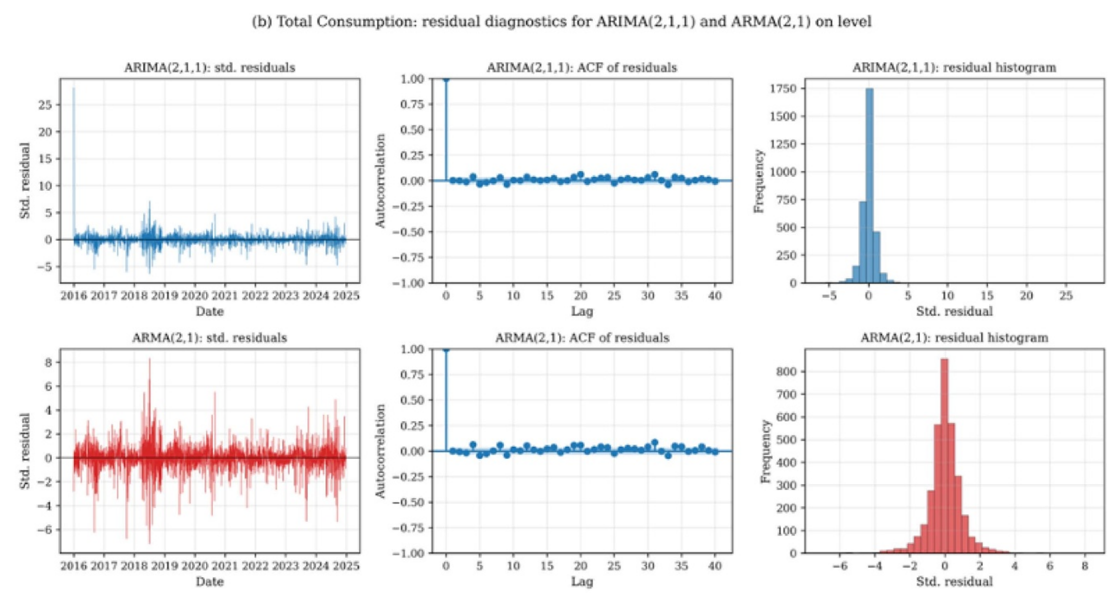
Figure 7. Partial autocorrelation function (PACF) plots of first-differenced series.

Similarly, for Total Consumption, ARIMA(2,1,1) outperformed differenced alternatives with AIC = 4168.3 and BIC = 4192.2 versus ARIMA(3,1,1) at AIC = 4170.6 and BIC = 4200.5. Error variance was essentially identical ($\sigma^2 \approx 0.25$). The consistent selection of ARIMA(2,1,1) among differenced specifications confirms the parallel short-term dynamics observed in the ACF and PACF analyses, reinforcing the conclusion that water supply and demand processes share fundamentally similar temporal structures.

However, when ARMA(2,1) is fitted directly to the (already-stationary) level series, it achieves a lower AIC than ARIMA(2,1,1) (4061.3 versus 4066.1 for Supply, and 4163.3 versus 4168.3 for consumption) and substantially better held-out predictive accuracy (RMSE 2.30 $m^3\ s^{-1}$ versus 3.20 $m^3\ s^{-1}$ for Supply, and 2.34 versus 3.19 for consumption). On both information-theoretic and out-of-sample grounds, and in the absence of any unit root in the data, ARMA(2,1) is therefore used as the conservative parametric baseline for the analyses that follow; ARIMA(2,1,1) is included as a comparison model to assess the consequences of the differencing choice. Note that BIC marginally favours ARIMA(2,1,1) (4090.0 vs 4091.2 for Supply on the training series) because of its one fewer parameter; this difference of approximately 1.2 units is negligible on the Kass–Raftery scale. The use of ARMA(2,1) as the conservative parametric baseline does not, however, rest on information criteria: it follows from the formal ADF and KPSS evidence of stationarity in levels (section 3.1.3), the substantially lower out-of-sample RMSE, and the physically bounded prediction intervals (section 3.1.5), each of which favours the level specification over the differenced one. Seasonal SARIMA(2,1,1)$(P,D,Q)_7$ variants with weekly seasonal AR and MA components were also tested, motivated by the small but statistically significant residual autocorrelations at lags 7, 14, 21 and 28 noted in section 3.1.3; the best of these, SARIMA(2,1,1)$(0,0,1)_7$, attained



(a) Total Water Supply: standardised residuals, ACF of residuals, and residual histogram



(b) Total Consumption: standardised residuals, ACF of residuals, and residual histogram

Figure 8. Residual diagnostics for the conservative ARMA(2,1) on level models.

a marginally lower training-set AIC than ARIMA(2,1,1) (4060.0 versus 4066.1 for supply; a six-unit gap that, on the Kass–Raftery scale, constitutes positive evidence for the weekly seasonal term) but produced an identical held-out 2024 RMSE (3.20 m^3 s^{-1} for supply), so the weekly seasonal component adds no out-of-sample forecasting accuracy. A full annual-period SARIMA model ($s = 365$) was not pursued because the seasonal period is long relative to the available daily record and imposes a substantial computational burden; the annual seasonal cycle is instead captured by the Prophet specification through Fourier-series decomposition.

Residual diagnostics were performed to validate model adequacy (figure 8). Each diagnostic plot consists of three panels: the time plot of innovation residuals (top), the ACF of residuals (bottom left), and the histogram of standardised residuals (bottom right).

In figure 8, the time plot of standardised residuals showed fluctuation around zero with no apparent trends, systematic patterns, or clusters of large residuals. Variance appeared stable throughout the period, with only occasional spikes exceeding ±3 standard deviations during the 2018 and 2020 periods previously identified as containing unusual events. These isolated excursions did not indicate systematic model failure but rather reflected genuine extreme observations in the underlying water system.

The ACF of residuals (figure 8, bottom left) exhibited small but statistically significant autocorrelations at multiple lags. Values remained predominantly below ±0.05 in magnitude, with the largest spike occurring at lag 21 ($\rho \approx 0.07$). While the Ljung–Box test formally rejected the white noise hypothesis ($p = 5.90 \times 10^{-10}$), this result must be interpreted with caution. With a sample size of 3286 observations, even trivial autocorrelations become statistically significant due to high statistical power. The practical magnitude of these correlations was negligible, and their pattern did not suggest obvious model misspecification. This residual structure likely reflects high-frequency microstructure in daily operations (weekly scheduling, system maintenance cycles) that, while statistically detectable, contributes minimally to forecast error.

The histogram of standardised residuals (figure 8) exhibited a symmetric, approximately bell-shaped distribution centred near zero. The distribution showed slight positive skewness and marginally heavier tails than a pure Gaussian, evidenced by the small number of observations beyond ±3 standard deviations. However, the overall shape was sufficiently close to normality to support the use of prediction intervals based on the Gaussian assumption, particularly for the central 95% of the distribution, which is most relevant to forecasting applications.

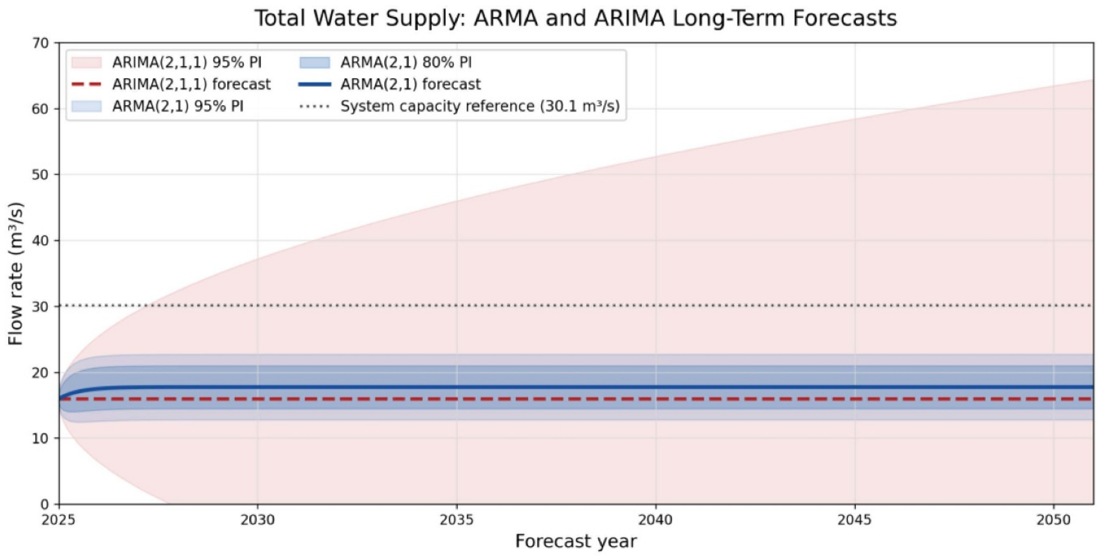
Residual diagnostics for Total Consumption (figure 8) closely paralleled those for Water Supply. The time plot showed centred fluctuations with stable variance and similar timing of extreme residuals. The ACF of residuals displayed comparable small-magnitude autocorrelations (largest $\rho \approx 0.06$ at lag 14), yielding a similar Ljung–Box test result ($p = 5.23 \times 10^{-9}$). The histogram likewise exhibited near-normality with slight departures in the tails. The remarkable similarity between figure 8 quantitatively confirmed that the two series not only share similar autocorrelation structures but also exhibit comparable forecast error properties, validating the adequacy of the conservative ARMA(2,1) specification for both variables.

3.1.5. Long-term forecasts (2025–2050)

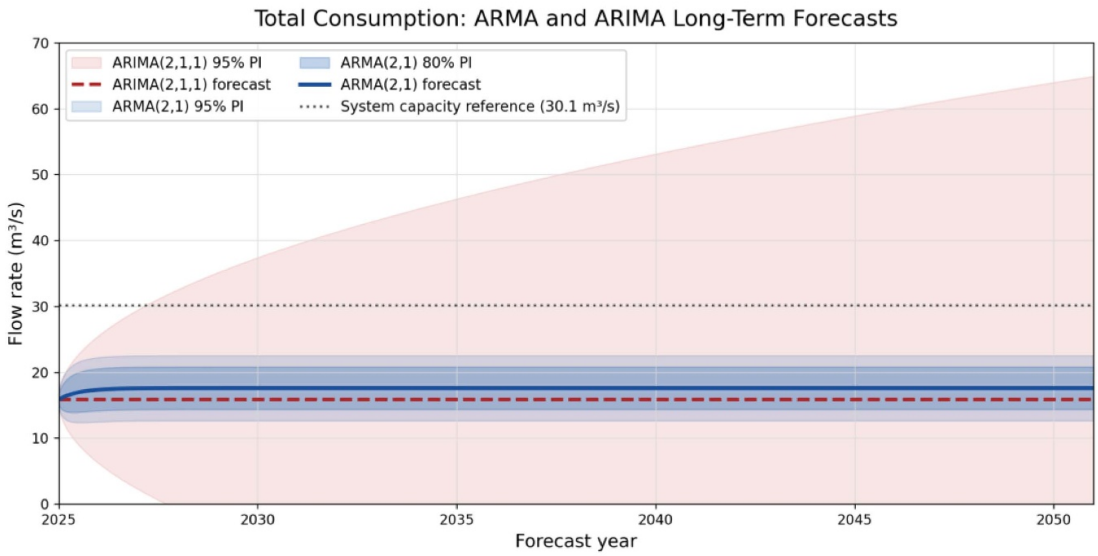
The validated ARMA(2,1) models on the level series—the conservative parametric baseline—were used to generate daily forecasts for 2025–2050, with ARIMA(2,1,1) included as a comparison model; both were aggregated to annual totals for strategic planning purposes. For Total Water Supply, ARMA(2,1) forecasts converge rapidly to the unconditional mean of approximately 17.6 $m^3\ s^{-1}$ and remain stationary thereafter (figure 9), consistent with the formal stationarity tests reported in section 3.1.3.

Figure 9 illustrates the qualitatively different behaviour of the two specifications. The ARMA(2,1) point forecasts (solid blue) converge rapidly to the unconditional mean and remain bounded; the ARIMA(2,1,1) point forecasts (dashed red), shown for comparison, also stabilise at approximately the last observed level. The forecasts illustrate several characteristic features of integrated and stationary time-series specifications. The point forecasts (solid line for ARMA(2,1), dashed for ARIMA(2,1,1)) converge rapidly to a stable level by approximately 2027 and remain essentially constant thereafter. An integrated ARIMA process does not, strictly, possess an unconditional mean, the long-run ARIMA(2,1,1) forecast level depends on the last observed values rather than the historical average, but in the present case the last-observed levels and the historical means coincide closely, so the practical interpretation in terms of 'mean reversion' is essentially correct for both ARMA(2,1) and ARIMA(2,1,1). The qualitative difference between the two specifications lies not in the central trajectory but in the prediction-interval behaviour discussed below. The mean-reverting behaviour of both specifications reflects their lack of a deterministic trend component; absent new information about structural changes, their best guess for the long-term future is simply the historical level.

For ARMA(2,1), prediction intervals expand only over the first few years and then asymptote to a fixed width determined by the unconditional variance of the stationary process: the 95% band stabilises at approximately ±4.9 $m^3\ s^{-1}$ and the 80% band at approximately ±3.2 $m^3\ s^{-1}$, both bounded for all horizons (figure 9, blue shading). For the ARIMA(2,1,1) comparison, intervals widen progressively with horizon as expected for an integrated process: the 95% band grows from approximately ±2 $m^3\ s^{-1}$ in 2025 to ±48 $m^3\ s^{-1}$ by 2050 (figure 9). This contrast is one of the principal substantive reasons for using ARMA(2,1) as the conservative parametric baseline: the bounded ARMA intervals remain physically meaningful across the entire projection horizon, whereas the ARIMA intervals lose physical interpretability beyond the first decade. For the ARIMA(2,1,1) comparison, the unconstrained Gaussian 95% prediction interval extends below zero by approximately 2035, which is physically impossible for a non-negative quantity such as water flow. The lower bound has therefore been truncated at zero in figure 9; the upper bound by 2050 reaches approximately 64 $m^3 s^{-1}$, which itself exceeds the documented



(a) Total Water Supply



(b) Total Consumption

Figure 9. Long-term forecasts (2025–2050) with 80% and 95% prediction intervals: ARMA(2,1) on level (blue) and ARIMA(2,1,1) (dashed red, lower bound zero-floored).

system capacity of $30.1\,m^3 s^{-1}$ (table 1) and should therefore be regarded as physically implausible for the water-supply application. The ARMA(2,1) intervals are not subject to either of these problems: the bounded interval remains within $[\approx 12.7, \approx 22.5]\,m^3 s^{-1}$ across the entire 2025–2050 horizon, comfortably above zero and below system capacity, providing physically meaningful uncertainty bounds at all forecast horizons.

The contrast in interval-width behaviour follows a well-defined mathematical relationship. For stationary ARMA models, the forecast error variance is bounded by the unconditional variance of the process and asymptotes within a few horizons. For integrated ARIMA models, by contrast, the forecast error variance grows without bound as a function of forecast horizon h according to the model's impulse response weights, with the rate of expansion determined by the persistence of shocks. The observed contrast in figure 9 is the textbook manifestation of this distinction and a direct consequence of the inappropriate differencing in ARIMA(2,1,1) for an already-stationary series.

Total Consumption forecasts exhibited characteristics nearly identical to those of production. The ARMA(2,1) daily point forecasts stabilised at the unconditional mean of approximately 17.4 $m^3\ s^{-1}$, with corresponding 95% bands of approximately $\pm$4.9 $m^3\ s^{-1}$; these are daily-scale flow values from

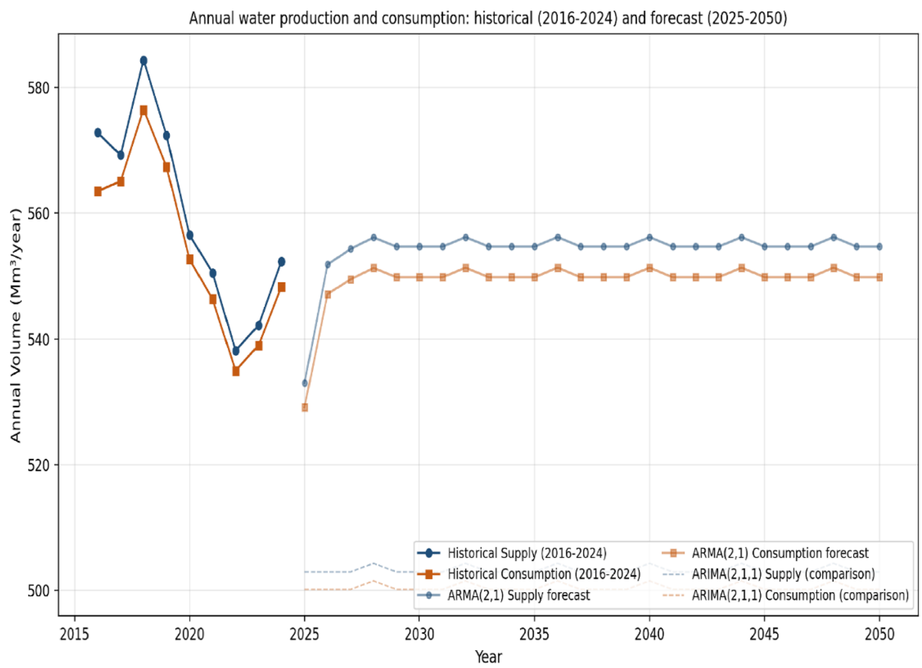


Figure 10. Annual water production and consumption: historical (2016–2024) and forecast (2025–2050) by ARMA(2,1) (solid) and ARIMA(2,1,1) (dashed).

the underlying ARMA(2,1) output, whereas figure 10 displays the annual aggregated totals. The close correspondence in figure 10 demonstrates that the two variables exhibit similar forecast uncertainty, a natural consequence of their parallel time-series structures and similar historical volatility.

Comparing figure 10 reveals that consumption forecasts are marginally more uncertain than production forecasts (wider bands by approximately 3%–5%), consistent with the slightly higher error variance estimated for the consumption ARMA(2,1) model ($\sigma^2 \approx 0.246$ vs 0.238). This small difference likely reflects the fact that consumption is influenced by user behaviour and external factors (weather, economic activity) that introduce additional unpredictability beyond the production-side operational variability.

Annual aggregated forecasts revealed that production consistently exceeded consumption throughout the entire 2025–2050 period, with the conservative ARMA(2,1) on level projecting an annual surplus of approximately 4.9 Mm^3 yr^{-1} ($\approx$0.87% of annual supply) and the ARIMA(2,1,1) comparison projecting approximately 2.8 $Mm^3 yr^{-1}$ ($\approx$0.55%); see figure 10 and table 2. Both projections indicate near-balance rather than substantial reserve capacity, assuming historical patterns remain unchanged.

Figure 10 presents annual totals rather than daily values, aggregating daily forecasts to the planning-relevant annual scale. The regular oscillations visible in the figure represent the projection of historical seasonal patterns across future years. ARMA(2,1) production forecasts (solid blue, upper) and consumption forecasts (solid orange, lower) maintain parallel trajectories separated by approximately 4.9 Mm^3 yr^{-1} ($\approx$0.87% of total annual supply); the corresponding ARIMA(2,1,1) trajectories (dashed) lie roughly 50 Mm^3 below them, reflecting the integrated model anchoring its long-run forecast at the lower 2024 levels rather than the historical mean, demonstrating a persistent but small structural surplus.

The small magnitude of this surplus (equivalent to $\approx$3 d of total production per year for ARMA(2,1) and $\approx$1.8 d for ARIMA(2,1,1)) indicates that the system operates near supply–demand balance rather than maintaining substantial reserve capacity. While both univariate models project that this balance will persist, the extremely narrow margin highlights vulnerability to disruptions. Any increase in consumption growth rates, reduction in production capacity, or change in external conditions could quickly transform the modest surplus into a deficit. The forecast should therefore be interpreted as a baseline projection contingent on continuation of historical patterns, not as evidence of robust long-term water security.

3.2. Discussion and model limitations

Both univariate specifications generated conservative, mean-reverting forecasts: ARMA(2,1) on the level series (the conservative parametric baseline) projects an annual surplus of approximately 4.9 Mm^3 yr^{-1}

Table 2. Comparative performance of forecasting models, including accuracy, information criteria, and surplus forecasts.

Model	n	AIC	BIC	RMSE 2024	RMSE CV (5-fold)	Surplus 2030 (Mm^3 yr^{-1})	Surplus 2050 (Mm^3 yr^{-1})	95% PI 2030 (m^3 s^{-1})	95% PI 2050 (m^3 s^{-1})
SUPPLY series									
ARIMA(2,1,1)—comparison (over-differenced)	4	4066.1	4090.0	3.20	3.16 ± 0.27	2.78	2.78	38.2	63.9
ARMA(2,1) on level—conservative baseline	5	4061.3	4091.2	2.30	2.31 ± 0.20	4.98	4.98	9.9	9.9
Prophet (defaults)	$\sim$30	—	—	0.75	1.00 ± 0.27	6.31	14.77	35.6	193.7
Seasonal-naive baseline	0	—	—	1.07	1.06 ± 0.18	—	—	—	—
CONSUMPTION series									
ARIMA(2,1,1)—comparison (over-differenced)	4	4168.3	4192.2	3.19	3.14 ± 0.27	—	—	38.4	64.5
ARMA(2,1) on level—conservative baseline	5	4163.3	4193.2	2.34	2.32 ± 0.19	—	—	9.9	9.9
Prophet (defaults)	$\sim$30	—	—	0.78	0.96 ± 0.20	—	—	28.3	154.7
Seasonal-naive baseline	0	—	—	1.12	1.06 ± 0.18	—	—	—	—

(≈0.87% of annual supply), and the ARIMA(2,1,1) comparison projects approximately 2.8 Mm^3 yr^{-1} (≈0.55%). While these projections provide valuable baseline scenarios, they are subject to important limitations that must be considered when using results for policy and planning decisions.

Both univariate specifications (ARMA(2,1) on level and the ARIMA(2,1,1) comparison) share a fundamental constraint: they learn solely from historical production and consumption patterns. Neither incorporates external drivers known to affect water supply and demand, including population growth, climate change impacts, temperature trends, economic development, regional demand shifts, or infrastructure investments. The model assumes that all factors influencing water flows will continue to operate as they have historically, an assumption increasingly questionable over 26 year projection horizons.

The flat long-term forecasts (figure 10) reflect this structural limitation. By construction, univariate ARMA/ARIMA models without exogenous predictors cannot anticipate shifts in mean levels arising from external factors. The convergence to a stable level (the unconditional mean for ARMA(2,1), the last observed values for ARIMA(2,1,1)), therefore, represents a statistical artefact of the model specification rather than a substantive prediction about future water system behaviour. If consumption grows due to population increase or economic expansion, or if production capacity contracts due to drought or infrastructure ageing, the actual trajectory will deviate systematically from these projections.

To assess the value of including exogenous regressors directly within an ARIMA framework, an ARIMAX(2,1,1) specification with calendar exogenous regressors (month-of-year and day-of-week dummies) was additionally fitted to the same training window. Calendar regressors were chosen rather than temperature because daily ERA5 or comparable temperature data for Baku were not available within the present analysis; calendar variables capture the same underlying seasonal forcing as temperature, although less directly. The calendar-ARIMAX achieved a 2024 hold-out RMSE of approximately 2.99 m^3 s^{-1} for supply and 3.05 m^3 s^{-1} for consumption, compared with 2.34 and 2.38 m^3 s^{-1}, respectively, for the conservative ARMA(2,1) on level—that is, calendar regressors did not improve predictive accuracy beyond what the AR(2) structure on the level series already delivers. This is consistent with the seasonal cycle being already partially absorbed by the AR structure once the inappropriate first-differencing is removed. Temperature, which would carry additional information about cooling/irrigation demand and source-water availability beyond pure seasonality, is more likely to add value and is flagged as an immediate priority for follow-up work using daily ERA5 temperature for Baku (40.4° N, 49.9° E).

The widening prediction intervals (figure 10) quantify forecast uncertainty arising from parameter estimation error and the stochastic nature of the processes, but they do not account for model specification uncertainty or structural breaks. The true uncertainty surrounding long-term water balance is likely substantially larger than suggested by the statistical intervals, particularly beyond 2030 when cumulative changes in external conditions become more significant.

Despite these limitations, the analysis provides several practical findings. The persistent positive supply-demand balance in all forecast scenarios, while modest in magnitude, offers reassurance regarding near-term (2025–2030) water security under business-as-usual conditions. The strong synchronisation between production and consumption series indicates effective short-term operational management and demand responsiveness.

The seasonal patterns documented in the raw series (figure 2), showing 30%–35% variation between summer peaks and winter troughs, offer actionable information for operational planning and demand-side management interventions. The timing and magnitude of these patterns could inform storage management, maintenance scheduling, and public water conservation campaigns.

Future modelling efforts should address current limitations through several extensions. ARIMAX models incorporating exogenous variables (temperature, precipitation, population, economic indicators) would allow explicit modelling of external drivers and scenario-based forecasting under different assumptions. Vector autoregression (VAR) or vector error correction frameworks could capture dynamic relationships between supply, demand, and external variables while allowing for long-run equilibrium relationships. Machine learning approaches (XGBoost, Random Forest, LSTM networks) may better capture non-linear relationships and long-range dependencies, particularly when trained on multivariate datasets including climate, socioeconomic, and hydrological variables.

Integration with hydrological models of source water systems (Kura and Samur rivers) would allow coupling of demand-side forecasts with supply-side constraints, enabling a comprehensive water security assessment. Incorporation of climate change scenarios from regional climate models would permit evaluation of supply reliability under different warming pathways. Demographic and economic projections could translate into spatially and temporally resolved demand forecasts that account for urban growth patterns and sectoral shifts.

While current ARIMA results indicate adequate water supply capacity under historical patterns, prudent water resource management for the Absheron Peninsula requires scenario planning that accounts

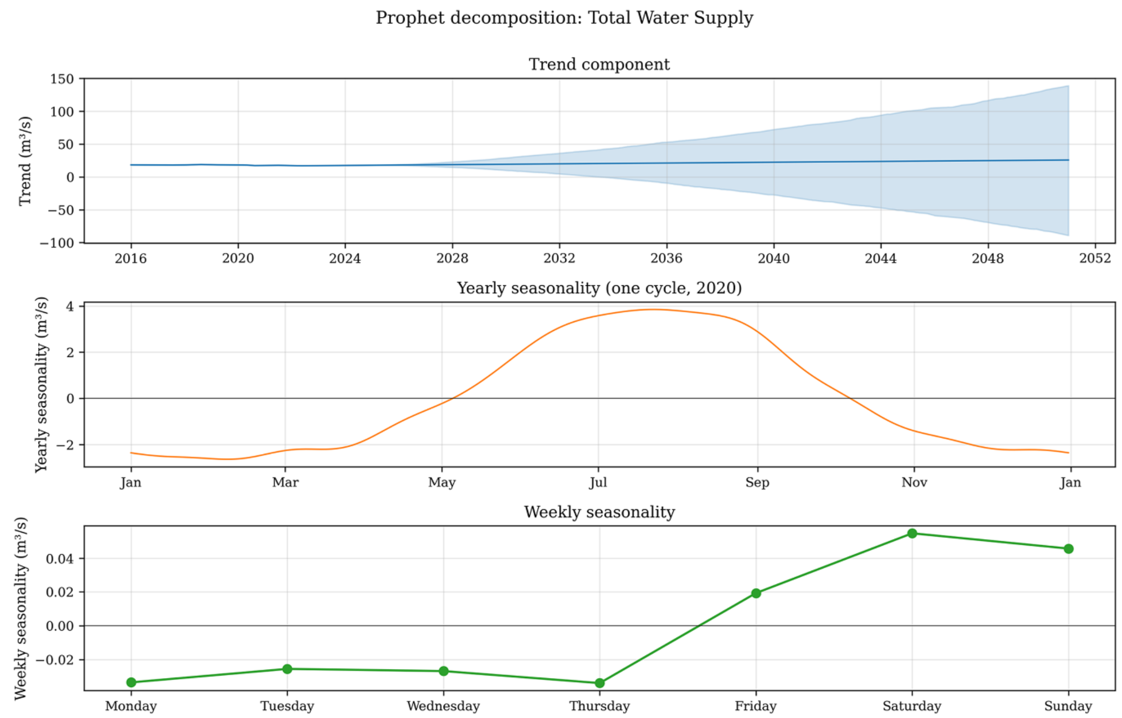


Figure 11. Prophet model decomposition for total water supply: trend, weekly seasonality, and yearly seasonality components.

for climate variability, demand growth, and infrastructure constraints not captured in univariate statistical models. The baseline projections presented here provide a reference point against which alternative scenarios and policy interventions can be evaluated, but should not be interpreted as definitive predictions of future water system conditions.

3.3. Prophet model results and long-term projections

Complementing the ARIMA analysis, Prophet models were developed to capture long-term trends and seasonal patterns through an additive decomposition framework. Prophet's explicit treatment of trend and seasonality components provides interpretable insights into the structural drivers of water supply and demand dynamics.

Prophet model decomposition reveals the underlying structure of water production patterns as shown in figure 11. The trend component indicates gradual long-term changes. In both decomposition figures (figures 11 and 12), the shaded band around the trend denotes Prophet's uncertainty interval for the trend component, obtained from the posterior distribution over changepoint magnitudes; it widens where the historical data constrain the local trend less strongly and should not be interpreted as a forecast prediction interval, while yearly seasonality captures the strong summer peak (June–August) with approximately 4 m^3 s^{-1} amplitude. Weekly seasonality shows elevated weekend production, particularly on Saturday, reflecting operational scheduling and increased weekend demand.

Prophet decomposition of consumption series, as visualised in figure 12, exhibits similar seasonal structures with notable differences. The trend component shows steady growth, suggesting increasing baseline demand over the historical period. Yearly seasonality mirrors production with summer peaks of 4 m^3 s^{-1} above baseline. Weekly patterns indicate higher weekend consumption, with Sunday showing the maximum weekly effect at approximately 0.05 m^3 s^{-1} above the weekly average.

Comparison of production and consumption decompositions reveals coordinated patterns. Both series exhibit identical yearly seasonality amplitude (± 4 m^3 s^{-1}), confirming supply responsiveness to demand fluctuations. However, consumption's stronger positive trend component suggests demand growth outpacing production growth historically, though production maintains adequate levels through operational adjustments.

Long-term Prophet forecasts for water production through 2050, as shown in figure 13, maintain seasonal oscillations while incorporating gradual trend continuation.

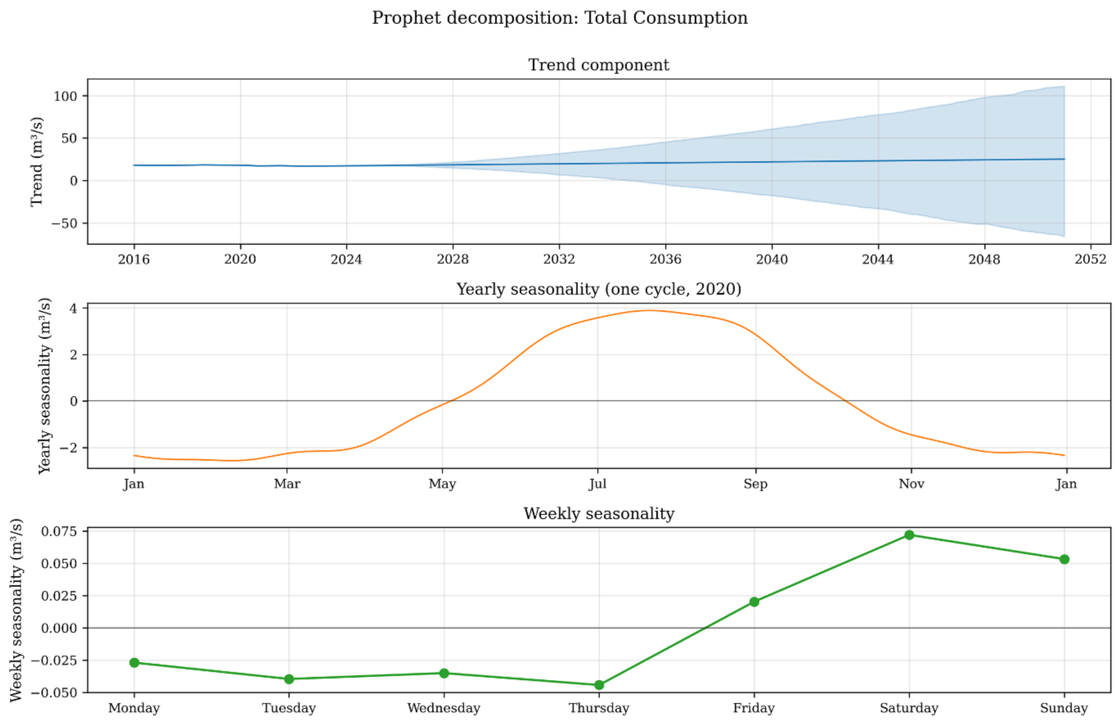


Figure 12. Prophet model decomposition for total consumption: trend, weekly seasonality, and yearly seasonality components.

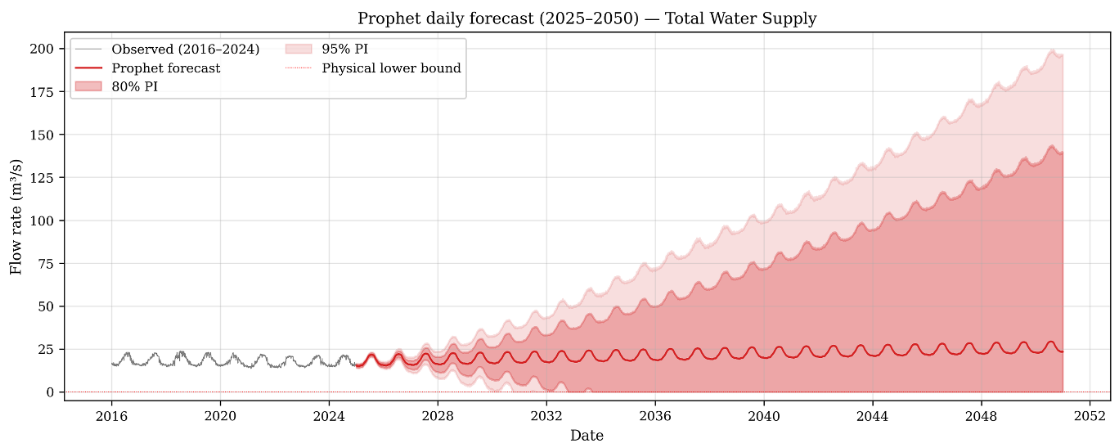


Figure 13. Prophet daily water production forecast (2025–2050) with uncertainty intervals.

Figure 13 presents Prophet's 26 year production forecast exhibiting continued seasonal cycling with modest trend extrapolation. The forecast centreline maintains a near-horizontal trajectory around historical means, with seasonal amplitude preserved throughout. The 95% interval is zero-floored and therefore asymmetric: by 2050 it spans from zero to approximately 194 $m^3\ s^{-1}$ against a centreline close to 25 $m^3\ s^{-1}$, giving a full width of about 193.7 $m^3\ s^{-1}$. This upper bound is not a physically attainable operating flow; it is the unconstrained statistical upper tail of Prophet's long-horizon trend uncertainty and exceeds the 30.1 $m^3\ s^{-1}$ system-capacity reference in table 1, as does the widened ARIMA comparison interval. The interval should therefore be read as evidence of large extrapolation uncertainty rather than as a feasible capacity scenario.

Prophet's consumption forecast (figure 14) displays similar structural characteristics to production but with slightly steeper trend extrapolation. The seasonal amplitude remains consistent, maintaining summer peaks and winter troughs throughout the projection period. Uncertainty intervals for consumption expand marginally faster than production, reaching an upper bound near 155 $m^3\ s^{-1}$ by 2050

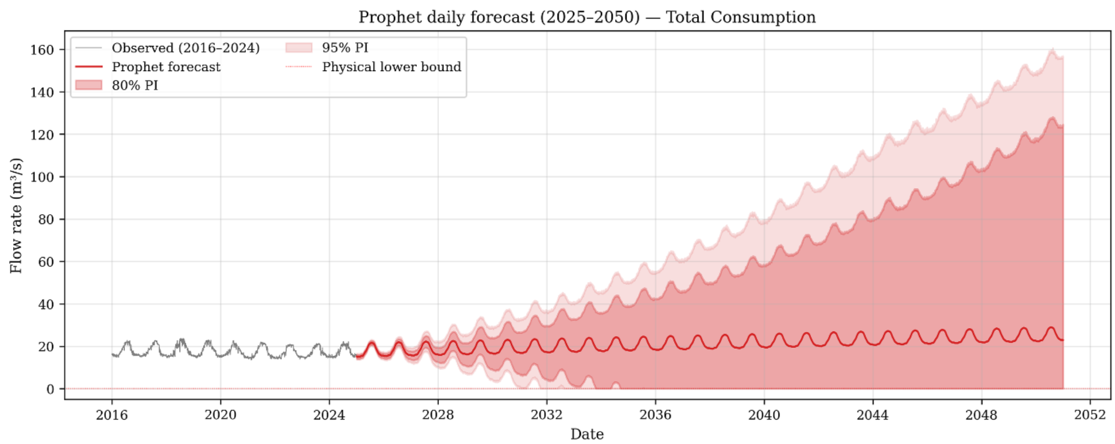


Figure 14. Prophet daily water consumption forecast (2025–2050) with uncertainty intervals.

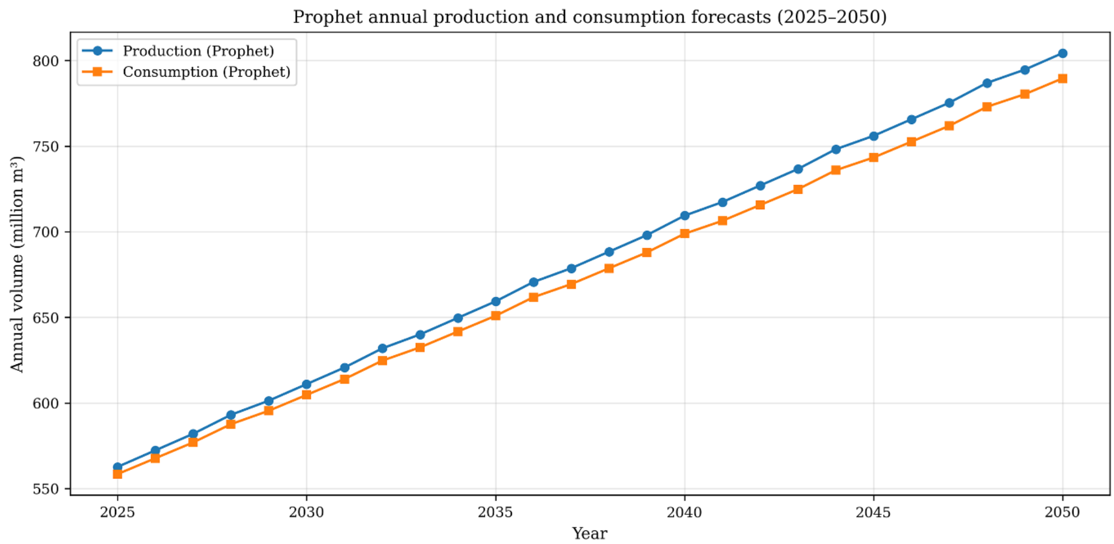


Figure 15. Prophet annual water production and consumption forecasts (2025–2050).

(the lower bound likewise floored at zero). This asymmetry suggests Prophet identifies greater historical variability in consumption patterns, potentially reflecting demand-side volatility from weather and socioeconomic factors not explicitly modelled.

For strategic planning purposes, daily Prophet forecasts were aggregated to annual totals. This transformation facilitates policy-relevant insights and direct comparison with ARIMA projections.

Figure 15 presents annual aggregated forecasts showing production rising from approximately 562 Mm^3 yr^{-1} in 2025–804 Mm^3 yr^{-1} by 2050, with consumption rising in parallel from 558 to 790 Mm^3 yr^{-1}. Both series exhibit near-linear growth with production consistently exceeding consumption. The annual growth rate averages 1.44% for production and 1.39% for consumption (compound annual growth, computed from the fitted 2025 and 2050 annual totals), substantially higher than ARIMA's stable projections. This divergence stems from Prophet's trend extrapolation mechanism, which continues historically observed growth patterns into the future.

The supply-demand balance, shown in figure 16, reveals a critical difference from ARIMA projections. Prophet forecasts an expanding surplus from approximately 4.2 Mm^3 yr^{-1} in 2025 to 14.8 Mm^3 yr^{-1} by 2050, contrasting sharply with the corrected ARIMA(2,1,1) surplus of approximately 2.8 Mm^3 yr^{-1}. This roughly 3.5-fold increase relative to 2025 (and approximately 5-fold relative to ARIMA) reflects Prophet's compound growth assumption: production's slightly higher growth rate (1.44% vs 1.39%) compounds over 26 years to generate widening surplus. However, this optimistic scenario depends critically on the assumption that historical growth trends persist unmodified—an assumption that may not withstand demographic, climatic, or infrastructural changes.

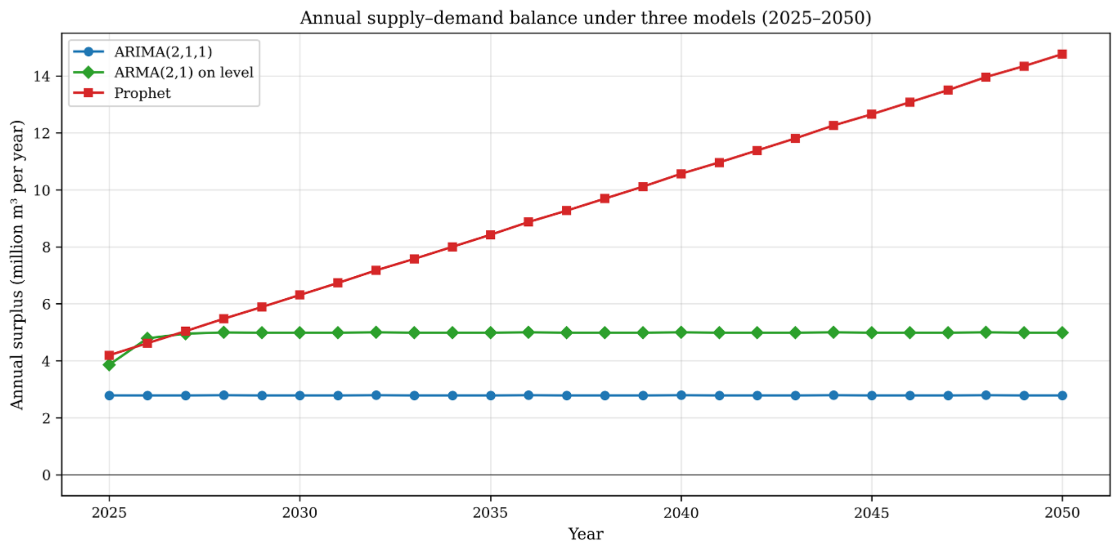


Figure 16. Prophet annual supply-demand balance (2025–2050).

A 5 × 4 hyperparameter sensitivity grid was conducted to assess the robustness of the Prophet defaults, varying changepoint prior scale (cps) across 0.001, 0.01, 0.05, 0.1, and 0.5 and seasonality prior scale (sps) across 0.1, 1, 10, and 100. Held-out 2024 RMSE varied modestly across this grid: from 0.73 to 1.04 $m^3\ s^{-1}$ for supply (default cps = 0.05, sps = 10 producing 0.75) and from 0.77 to 1.03 $m^3\ s^{-1}$ for consumption (default 0.78). The very stiff trend setting (cps = 0.001) underperformed for short-horizon accuracy, while cps values from 0.05 to 0.5 produced similar 1 year RMSE. The long-horizon balance was therefore checked directly at the changepoint-prior endpoints while holding sps = 10: cps = 0.001 gives a 2050 balance of approximately −7.1 $Mm^3\ yr^{-1}$ (a deficit), whereas cps = 0.5 gives approximately +30.5 $Mm^3\ yr^{-1}$ (a surplus). Short-horizon RMSE stability therefore does not imply long-horizon surplus stability; the Prophet projection is interpreted below as a trend-continuation scenario rather than as a hyperparameter-invariant forecast.

3.4. Comparative analysis of forecasting specifications

The fundamental divergence between ARMA(2,1)/ARIMA(2,1,1) and Prophet forecasts—approximately 4.9/2.8 $Mm^3\ yr^{-1}$ versus 14.8 $Mm^3\ yr^{-1}$ annual surplus by 2050—encapsulates different modelling philosophies and assumptions about system dynamics. This section systematically compares the two approaches across forecasting methodology, structural assumptions, uncertainty quantification, and policy implications.

Methodological differences underlie the contrasting projections. The original ARIMA(2,1,1) comparison model employs differencing, even though the formal ADF and KPSS tests indicate that the level series are already stationary. This differencing removes low-frequency persistence and leads to wider, physically less interpretable long-horizon intervals, effectively removing trends and modelling only short-term dependencies. The resulting forecasts exhibit mean reversion: after accounting for autocorrelation structure, predictions gravitate toward historical means. In contrast, Prophet explicitly models the trend as a separate component using piecewise-linear or logistic growth functions. When fitted to data showing historical growth, Prophet extrapolates this trend forward, generating forecasts that continue rising (or falling) rather than reverting to means.

For the Absheron Peninsula data, this philosophical difference manifests starkly. ARIMA identifies no significant deterministic trend after differencing, interpreting the 2016–2024 period as fluctuations around stable means with strong seasonal and AR components. Prophet, however, detects subtle positive trends in both production (growth rate: 1.44% annually) and consumption (growth rate: 1.39% annually) and projects these forward. Over 26 years, even modest growth rates compound substantially: a 1.44% annual increase compounds to approximately 43% total growth by 2050, explaining Prophet's expanding surplus.

Structural assumptions further differentiate the models. ARIMA's univariate framework treats water production and consumption as autonomous processes, learning only from their own histories. Prophet similarly operates univariately but adds flexibility through changepoint detection (automatically identifying 25 potential regime shifts) and Fourier-based seasonality (10 yearly terms, 3 weekly terms). This flexibility allows Prophet to capture more complex seasonal patterns but also increases parameter count and potential for overfitting, particularly when extrapolating far beyond training data.

Uncertainty quantification reveals critical differences in forecast confidence. ARMA(2,1) prediction intervals, derived from a stationary process, are bounded for all horizons: the 95% interval stabilises at approximately ± 4.9 m^3 s^{-1} and remains within physical limits across the 2025–2050 window. ARIMA(2,1,1) prediction intervals, by contrast, widen progressively with horizon as expected for an integrated process: the unconstrained 95% half-width expands from approximately ± 2 m^3 s^{-1} at one year to approximately ± 48 m^3 s^{-1} at 26 years, crossing physical bounds before the projection horizon ends. Prophet's Bayesian framework generates intervals from posterior predictive distributions, accounting for uncertainty in trend, seasonality, and observation noise. For Prophet, however, the corrected intervals are zero-floored and asymmetric rather than symmetric $\pm$ half-width bands. The production 95% full width grows from 35.6 m^3 s^{-1} in 2030–193.7 m^3 s^{-1} in 2050, and the corresponding consumption widths grow from 28.3 to 154.7 m^3 s^{-1} (table 2). The resulting upper bounds exceed the 30.1 m^3 s^{-1} supply-capacity reference and should be interpreted as unconstrained statistical uncertainty from multi-decade trend extrapolation, not physically feasible operating levels. This difference underscores the profound unpredictability of 26 year extrapolations.

Policy implications of the model divergence warrant careful consideration. The conservative ARMA(2,1) surplus of approximately 4.9 Mm^3 yr^{-1} ($\approx$0.87% of annual production)—and the smaller ARIMA(2,1,1) comparison surplus of 2.8 Mm^3 yr^{-1} ($\approx$0.55%)—provides minimal buffer for demand growth, supply disruptions, or forecast errors. Prudent water managers would interpret either figure as a warning signal necessitating supply augmentation or demand management measures. Prophet's $\approx$14.8 Mm^3 yr^{-1} surplus ($\approx$1.84% of annual production) might be misinterpreted as providing comfortable margin, potentially delaying necessary investments. However, Prophet's corrected uncertainty intervals are very wide and asymmetric: by 2050 the supply interval has a full width of approximately 193.7 m^3 s^{-1} and the consumption interval approximately 154.7 m^3 s^{-1}, with lower bounds floored at zero and upper bounds far beyond the documented capacity reference. The point forecast is therefore less informative than the uncertainty envelope, which indicates that substantial deficit outcomes cannot be excluded.

The appropriate interpretation synthesises all three perspectives. The ARMA(2,1) mean-reverting forecast represents a conservative parametric baseline with physically bounded uncertainty; the ARIMA(2,1,1) comparison provides the same central message under an integrated assumption; the Prophet trending forecast captures potential continuation of recent growth patterns but should not be interpreted as a prediction—rather as a scenario conditional on sustained historical trends. The approximately 2.8–14.8 Mm^3 yr^{-1} range thus bounds plausible futures under different assumptions about trend persistence. Given the 26 year projection horizon against a 9 year training window, roughly three times the training period, the 2050 figures should be regarded as illustrative scenarios rather than direct operational forecasts; the quantitative discussion of operational adequacy in this paper concentrates on the near-term window of approximately 2025–2030, where models retain meaningful predictive skill.

Validation metrics computed on held-out data are reported in full in table 2 below and visualised in figures 17 and 18. Table 2 compares the ARMA(2,1) model fitted to the level series, the ARIMA(2,1,1) comparison specification, Prophet, and the seasonal-naive benchmark. The ARMA(2,1) model is treated as the conservative parametric baseline because it is consistent with the formal stationarity testing in section 3.1.3, provides physically bounded prediction intervals as discussed in section 3.1.5, and improves held-out accuracy relative to the ARIMA(2,1,1) comparison. For the 2024 hold-out (training on 2016–2023), the conservative ARMA(2,1) model fitted to the level series achieves on supply (and 2.34 on consumption), the ARIMA(2,1,1) comparison achieves 3.20 m^3 s^{-1} (3.19), Prophet (default settings) achieves 0.75 m^3 s^{-1} (0.78), and a seasonal-naive baseline ($y_t = y_{t-365}$) achieves 1.07 m^3 s^{-1} (1.12). Across rolling-origin cross-validation with five hold-out years (2020–2024), the same ranking holds on average: Prophet (mean RMSE 1.00 $\pm$ 0.27 m^3 s^{-1})$\approx$seasonal-naive (1.06 $\pm$ 0.18 m^3 s^{-1}) $\ll$ ARMA(2,1) on level (2.31 $\pm$ 0.20 m^3 s^{-1}); ARIMA(2,1,1) (3.16 $\pm$ 0.27 m^3 s^{-1}). AIC and BIC are reported on the training set, while surplus values correspond to annual forecast volumes over 2025–2050. Prophet is not assigned comparable AIC/BIC values because it is fitted using Stan under a different objective function. Two implications follow. First, the simple seasonal-naive baseline is competitive with Prophet for 1 year-ahead forecasting, so the operational value added by the parametric models in the near term is modest. Second, the ARIMA(2,1,1) comparison performs worse than all alternatives on held-out data, including the trivial baseline, principally because the $d = 1$ differencing eliminates the strong annual cycle that the held-out year exhibits. The structural value of the ARMA(2,1), ARIMA(2,1,1), and Prophet specifications lies not in short-horizon forecast skill, where naive methods are competitive, but in their contrasting assumptions about long-horizon system behaviour (mean reversion to the unconditional mean

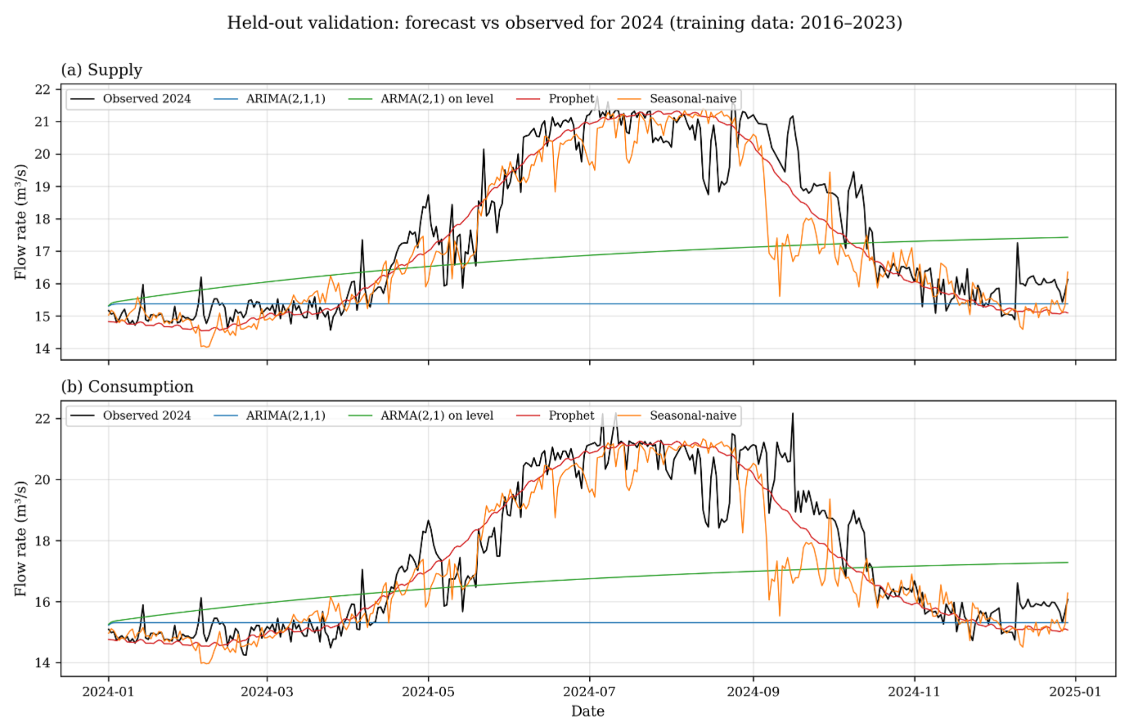


Figure 17. Held-out validation: forecast vs observed for 2024, training on 2016–2023. Panels show (a) supply and (b) consumption.

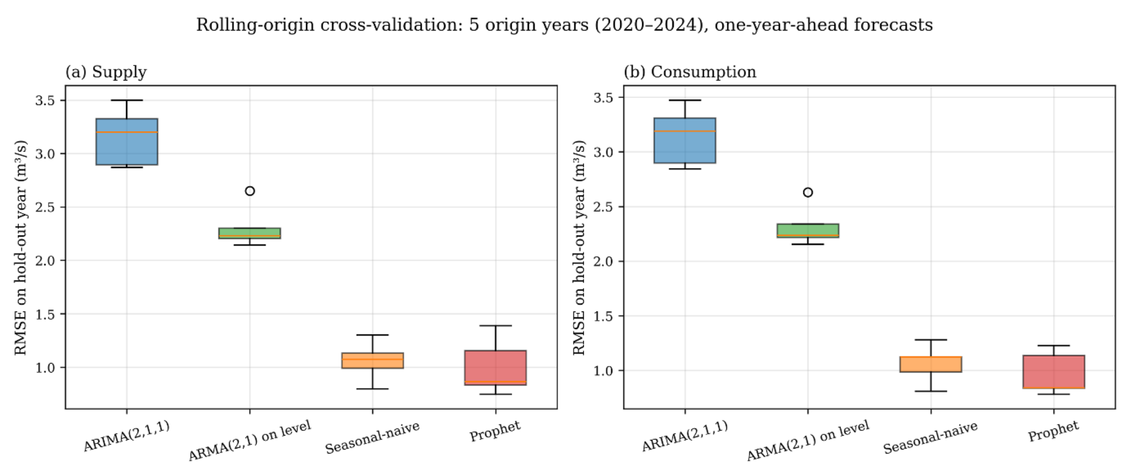


Figure 18. Rolling-origin cross-validation across 5 origin years (2020–2024) for one-year-ahead forecasts.

for ARMA, drift from the last observed values for integrated ARIMA, trend extrapolation for Prophet), which is the comparison the present study is designed to make.

This near-equivalence in short-term out-of-sample accuracy contrasts sharply with the substantial divergence observed in long-horizon projections, highlighting the risks associated with extrapolation beyond the range of observed data. While ARMA(2,1), ARIMA(2,1,1), and Prophet all capture short-term (daily-to-monthly) dynamics with roughly comparable skill, they rely on fundamentally different assumptions regarding long-term structural evolution over multi-decadal horizons.

For water resource planning, this analysis suggests a portfolio approach: use ARMA(2,1) to establish conservative baseline requirements under mean-reverting conditions, employ Prophet to explore trend continuation scenarios, and develop contingency plans for the uncertainty envelope spanning both projections. The divergence itself provides valuable information—it quantifies model uncertainty and underscores the need for adaptive management approaches that can respond as future observations reveal which model's assumptions better approximate reality.

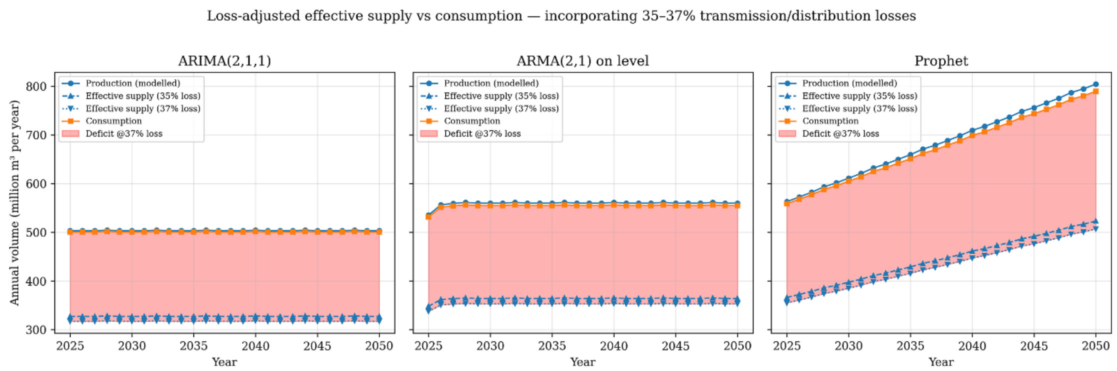


Figure 19. ARMA(2,1) loss-adjusted effective supply versus consumption (2025–2050), incorporating the documented 35%–37% transmission and distribution losses, with the operational-gap comparison panel.

Table 3. Model-specific loss-adjusted structural deficits by 2050 under the 35%–37% non-revenue-water stress-test scenario.

Model	Deficit at 35% loss (Mm^3 yr^{-1})	Deficit at 37% loss (Mm^3 yr^{-1})
ARIMA(2,1,1)—comparison	173	183
ARMA(2,1) on level—conservative baseline	191	202
Prophet (defaults)	267	283

3.5. Loss-adjusted stress-test scenario

While the baseline statistical results indicate adequate water supply capacity under historical patterns, prudent water resource management for the Absheron Peninsula requires scenario planning that accounts for climate variability, demand growth, and infrastructure constraints not captured in univariate statistical models. This paper distinguishes two quantities. The baseline statistical forecasts (ARMA(2,1), ARIMA(2,1,1), and Prophet) describe metered production and consumption and incorporate no loss factor, whereas the loss-adjusted scenario (figure 19 and table 3) applies the documented 35%–37% transmission and distribution losses to those same baseline production forecasts. The baseline forecasts therefore do not account for water losses occurring between source and end-user; it is the separate loss-adjusted stress-test scenario that incorporates them. Under this stress test, the modest surpluses projected by all three specifications are substantially eroded, potentially giving way to structural water deficits well before 2050. The forecasts presented here should therefore be interpreted not as reassuring predictions, but as baseline scenarios before explicit loss and demand-growth stress testing that highlight the urgency of proactive investment in loss reduction, supply diversification, and climate-adaptive water management strategies for the Absheron Peninsula.

It is informative to compare this literature-based loss assumption with the operational record contained in the source dataset. Using the complete daily supply and consumption series, the same-day metered operational gap was calculated as total source-side production (summed across the five supply points in table 1) minus total district-level consumption, expressed as a fraction of production. This calculated ratio is available for all 3 286 daily records and reports a substantially smaller mean supply-consumption gap of approximately 0.9% of supply (median 0.6%, 95th percentile 3.1%), corresponding to an aggregated annual operational gap of 3–9 Mm^3 yr^{-1} over 2016–2024. This is roughly two orders of magnitude smaller than the 35%–37% NRW figure cited from (Pasha *et al* 2023, World Bank 2024). The discrepancy most plausibly reflects the difference between metered operational throughput (which the calculated operational gap captures, essentially the imbalance between source-side production and end-user consumption recorded at meter points) and structural NRW (which additionally includes leakage between metering points, illegal connections, billing errors, and unbilled use). The two figures, therefore, measure different physical quantities. The 35%–37% NRW assumption is retained as the stress-test scenario in figure 19 because it is the strategically relevant figure for long-term infrastructure planning. Under this scenario, the ARMA(2,1) specification projects a structural deficit of approximately 191–202 Mm^3 yr^{-1} by 2050, the ARIMA(2,1,1) comparison projects approximately 173–183 Mm^3 yr^{-1}, and Prophet (with its higher long-term consumption projection) projects up to approximately 267–283 Mm^3 yr^{-1}; the 173–283 Mm^3 yr^{-1} combined range cited in the abstract and earlier in this section spans these three model-specific values. The observed 0.9% operational gap is shown as a complementary panel in the revised figure 19 to bound the optimistic edge of the loss-uncertainty envelope. For strategic-planning purposes, the 35%–37% NRW figure is adopted as a deliberately conservative stress-test assumption, recommended on the basis of the external (Pasha *et al* 2023, World Bank 2024) literature rather than derived from the present dataset; the contrast between it and the 0.9% operational gap quantifies a previously implicit assumption and identifies a high-value direction for further data acquisition (utility-side network audits to estimate true NRW directly from the operational data).

4. Conclusions

This study evaluated a set of complementary forecasting specifications, including an ARMA(2,1) model fitted to the level series as a conservative parametric baseline, an ARIMA(2,1,1) comparison model, Prophet models, and a seasonal-naive benchmark, to project water supply and demand for Azerbaijan's Absheron Peninsula through 2050. Using daily data from 2016–2024, all specifications identified strong seasonal patterns (amplitude of 30%–35%) and coordinated supply and demand management. However, they diverged fundamentally in long-term projections: the ARMA(2,1) level forecast exhibits mean reversion to the unconditional mean with an annual surplus of approximately 4.9 $Mm^3\ yr^{-1}$, the ARIMA(2,1,1) comparison drifts to the last observed values with a surplus of approximately 2.8 $Mm^3\ yr^{-1}$, and Prophet projects expanding surpluses reaching approximately 14.8 $Mm^3\ yr^{-1}$ by 2050. The approximately 3- to 5-fold difference between these projections reflects substantial model uncertainty regarding the persistence of trends versus mean reversion. Given a 26 year projection horizon against a 9 year training window, the 2050 figures should be regarded as illustrative scenarios under contrasting structural assumptions rather than direct operational forecasts; the quantitative discussion of operational adequacy focuses on the near-term 2025–2030 window, where models retain meaningful predictive skill.

Prediction intervals behave qualitatively differently across specifications. ARMA(2,1) on level produces bounded 95% bands that asymptote to approximately $\pm 4.9\ m^3\ s^{-1}$, remaining within physical limits across the entire horizon. ARIMA(2,1,1) bands expand from approximately $\pm 2\ m^3\ s^{-1}$ in 2025 to an unconstrained half-width of approximately $\pm 48\ m^3\ s^{-1}$ by 2050, crossing physical bounds well before the end of the projection period. Prophet's corrected intervals are wider and asymmetric after zero-flooring: the 95% full width reaches approximately 193.7 $m^3\ s^{-1}$ for supply and 154.7 $m^3\ s^{-1}$ for consumption by 2050. These Prophet upper bounds exceed the documented 30.1 $m^3\ s^{-1}$ capacity reference and should be interpreted as statistical extrapolation uncertainty rather than feasible operating ranges.

The loss-adjusted stress-test scenario presented in section 3.5 shows that the apparent baseline surpluses are not robust to the documented 35%–37% transmission and distribution loss assumption. Across the three forecasting specifications, the 2050 structural deficit range is approximately 173–283 $Mm^3\ yr^{-1}$, while the metered operational gap in the source records provides a much smaller optimistic lower-bound case. These results reinforce that the baseline forecasts should be interpreted as pre-loss statistical scenarios rather than evidence of secure long-term supply.

The model divergence itself provides policy-relevant information, bounding plausible futures under contrasting structural assumptions rather than offering definitive predictions. Enhanced forecasting capabilities should be pursued through several strategic developments. Multivariate frameworks such as ARIMAX and VAR models should incorporate temperature, precipitation, population growth, upstream water availability, and system loss rates for explicit scenario analysis. Hydrological modelling of the Kura and Samur basins should be integrated with demand forecasts under regional climate change scenarios. Ensemble approaches combining ARIMA, Prophet, and machine learning methods such as Random Forest and LSTM networks would improve forecast robustness, while spatially disaggregated projections would capture heterogeneous growth patterns across sectors and settlements.

Beyond forecasting methodology, the findings underscore that long-term water security for the Absheron Peninsula cannot rest on statistical projections alone. Proactive infrastructure investment to reduce the 35%–37% transmission and distribution losses represents the single most impactful near-term intervention available, as improvements here directly expand effective supply without requiring new source development. Climate-adaptive management strategies must be established now to prepare for declining inflows from transboundary rivers, and scenario analyses explicitly accounting for population growth should inform all infrastructure planning decisions. Supply diversification, demand-side efficiency programmes, and contingency protocols for deficit conditions should be developed proactively rather than reactively. Long-term planning for the Absheron Peninsula should therefore combine statistical forecasting with loss-reduction measures, source-diversification options, and periodic model updating as new operational and climate data become available.

Data availability statement

The daily water-supply and consumption records that support the findings of this study were provided by the regional water utility under operational data-sharing restrictions and are therefore not publicly available; they may be obtained from the corresponding author on reasonable request and with the permission of the data provider. The Python code used to perform the unit-root testing, the ARMA, ARIMA and Prophet modelling, the rolling-origin cross-validation and the figure generation is available from the corresponding author on reasonable request.

ORCID iDs

Mustafa Sundu 0000-0003-4168-9273
Mertol Tüfekci 0000-0002-5530-1471
İnci Pir 0000-0002-0540-5387
Sevgi Güneş Durak 0000-0003-4273-7417
Güler Türkoğlu Demirkol 0000-0001-8222-5497
Neşe Tüfekci 0000-0002-7373-4131

References

Ahadov B, Gadirli F and Hajiyeva G 2025 Assessment of water quality and coastal changes under climate impacts using multi-sensor satellite techniques in the Southern Absheron Peninsula *Environ. Monit. Assess.* **197** 1312
Ahmadov E 2020 Water resources management to achieve sustainable development in Azerbaijan *Sustain. Futures* **2** 100030
Aliyev R and Veliyeva Z M 2024 Aspects of efficient use of water and land resources in Azerbaijan *Int. J. Earth Environ. Sci.* **9** 86
Calvin K *et al* 2023 IPCC, 2023: climate Change 2023: synthesis report *Contribution of Working Groups I, II and III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change* ed H Lee and J Romero (Core Writing Team) (IPCC) (https://doi.org/10.59327/IPCC/AR6-9789291691647)
Chen J L, Pekker T, Wilson C R, Tapley B D, Kostianoy A G, Cretaux J-F and Safarov E S 2017 Long-term Caspian sea level change *Geophys. Res. Lett.* **44** 6993–7001
Döll P, Trautmann T, Gerten D, Schmied H M, Ostberg S, Saaed F and Schleussner C-F 2018 Risks for the global freshwater system at 1.5 °C and 2 °C global warming *Environ. Res. Lett.* **13** 044038
General Country Assessment-Azerbaijan Azerbaijan General Water Security Assessment n.d. Water Security Diagnostic Initiative (available at: www.worldbank.org/en/topic/water/publication/water-security-diagnostic-initiative) (Accessed 24 January2026)
Huntington T G 2006 Evidence for intensification of the global water cycle: review and synthesis *J. Hydrol.* **319** 83–95
Huseynov A T, Hasanova G M, Aliyeva S I, Hasanova V Y, Abdullayeva S A, Tahirova E I, Aliyev S I, Suleymanov S S and Muradov P Z 2025 Microbiological regime of some transboundary Rivers of Azerbaijan *Int. J. Biosci.* **26** 202–8
Intergovernmental Panel on Climate Change (IPCC) 2022 *Global Warming of 1.5°C: An IPCC Special Report on the Impacts of Global Warming of 1.5°C above Pre-industrial Levels and Related Global Greenhouse Gas Emission Pathways, in the Context of Strengthening the Global Response to the Threat of Climate Change, Sustainable Development, and Efforts to Eradicate Poverty* (Cambridge University Press) (https://doi.org/10.1017/9781009157940)
Intergovernmental Panel on Climate Change (IPCC) 2023 *Climate Change 2021—The Physical Science Basis* (Cambridge University Press) (https://doi.org/10.1017/9781009157896)
Ismaylov N M, Nadjafova S I and Gasimova A 2015 Absheron industrial region—Environmental stress factors *Arid Ecosyst.* **5** 194–200
Karthe D, Chalov S and Borchardt D 2015 Water resources and their management in central Asia in the early twenty first century: status, challenges and future prospects *Environ. Earth Sci.* **73** 487–99
Klaphake A and Kramer A 2011 Kura-Aras river basin: burgeoning transboundary water issues *Turkey's Water Policy* (Springer) pp 263–75
Kosarev A N 2005 Physico-geographical conditions of the Caspian sea *The Caspian Sea Environment* vol 5P, ed A G Kostianoy and A N Kosarev (Springer) pp 5–31
Lüthi D *et al* 2008 High-resolution carbon dioxide concentration record 650,000–800,000 years before present *Nature* **453** 379–82
Mekonnen M M and Hoekstra A Y 2016 Four billion people facing severe water scarcity *Sci. Adv.* **2** e1500323
Namazov I, Karimov V, Sharifov J and Pashayeva S 2024 Modern hydrogeological conditions of the Absheron Peninsula Visnyk of Taras Shevchenko National University of Kyiv *Geology* **4** 114–21
Pasha N H, Zengin E and Baylarli A T 2023 Sustainability of water resources of Azerbaijan, future prospects and problems *J. Manage. Econ. Ind. Org.* **7** 24–37
Ramírez-Agudelo N A, Porcar Anento R, Villares M and Roca E 2020 Nature-based solutions for water management in peri-urban areas: barriers and lessons learned from implementation experiences *Sustainability* **12** 9799
Schewe J *et al* 2014 Multimodel assessment of water scarcity under climate change *Proc. Natl Acad. Sci.* **111** 3245–50
Schleussner C-F *et al* 2016 Differential climate impacts for policy-relevant limits to global warming: the case of 1.5 °C and 2 °C *Earth Syst. Dyn.* **7** 327–51
Taylor R G *et al* 2013 Ground water and climate change *Nat. Clim. Change* **3** 322–9
Taylor S J and Letham B 2018 Forecasting at scale *Am. Stat.* **72** 37–45
Tierney J E *et al* 2020 Past climates inform our future *Science* **370** eaay3701

Trenberth K E, Fasullo J T and Balmaseda M A 2014 Earth's energy imbalance *J. Clim.* **27** 3129–44
United Nations Environment Programme 2022 Emissions gap report 2022: the closing window—climate crisis calls for rapid transformation of societies (available at: https://wedocs.unep.org/handle/20.500.11822/40874)
Vörösmarty C J *et al* 2010 Global threats to human water security and river biodiversity *Nature* **467** 555–61
World Bank 2024 Azerbaijan: transitioning to efficient water sector institutions and programs for addressing water security challenges (available at: https://documents1.worldbank.org/curated/en/099092424174049861/pdf/P17890612f1e5908c18a4f15be388af51ff.pdf)
Zarghami M, Abdi A, Babaeian I, Hassanzadeh Y and Kanani R 2011 Impacts of climate change on runoffs in East Azerbaijan, Iran *Glob. Planet Change* **78** 137–46