Modeling Urban Water Demand Using Agent-based Simulation: A Case Study in Salvador

Edmilson dos Santos de Jesus[1], **Célia Ghedini Ralha**[1]
Gecynalda Soares da Silva Gomes[1], **Karen Auzidéa dos Santos Pereira**[1]

[1]Department of Computer Science, Federal University of Bahia (UFBA)
Avenida Milton Santos, s/n - Campus de Ondina
PAF 2 – CEP: 40.170-110 – Salvador – Bahia – Brazil

edmilsonjesus@ufba.br, celiaralha@ufba.br

gecynaldassg@ufba.br, karen.asnt@gmail.com

Abstract. *This paper uses a data-driven reasoning model and inference rules to propose a Multi-Agent System (MAS) for water demand forecasting in the Metropolitan Region of Salvador (MRS). For implementation, the GAMA platform (GIS Agent-Based Modeling Architecture), its language, GAML (GAMA Modeling Language), and Python were used for simulation, preprocessing, and normalizing input data. The model considers population growth, average consumption per household, and housing type, enabling more individualized medium- and long-term forecasting. The results demonstrate the feasibility of using Multi-Agent Systems to support agencies and entities responsible for water resource management, providing valuable contributions to water supply's strategic and managerial planning.*

1. Introduction

The global population growth rate records an annual increase of approximately 80 million individuals, equivalent to more than 200,000 people daily, reinforcing the imperative for a growing supply of water intended for human consumption [Rosa et al. 2022]. The phenomenon of global warming and global population expansion result in an increased demand for potable water [Alves 2019]. In this context, the effective management of public water supply services is essential to ensure the uninterrupted provision of potable water to the population.

Galán et al. (2009) highlight limitations in traditional water demand forecasting methods, including the neglect of spatial interactions in per capita models, high data requirements and static nature of end-use models, the limited adaptability of extrapolation methods to structural changes, and the poor explanatory power of neural networks despite their predictive capability.

Frequently in the literature, water demand forecasting is classified into two main groups: short-term forecasting (forecasting for hours, days, and weeks) and long-term forecasting (months and years). Short-term forecasting is generally addressed using computational and statistical machine learning models, with an emphasis on the application of hybrid models and neural networks. Long-term forecasting tends to be more challenging to model, as it requires considering a larger number of variables, such as population growth, housing types, technological diffusion, and family behavior, where the application of multi-agent systems may prove to be a viable alternative for this type of forecasting.

Long-term water demand forecasting is essential to ensure the continuous supply of water to communities, as the installation and expansion of water supply systems need to be planned well in advance, given that these are typically engineering projects involving significant financial and operational resources, as well as requiring time for implementation.

The city of Salvador has over 2.4 million inhabitants [IBGE - Instituto Brasileiro de Geografia e Estatística 2022], and the system responsible for supplying this population is relatively complex. It consists of 3 intake points and 4 water treatment plants, where the Pedra do Cavalo intake, located 110 km from Salvador, is the primary source of raw water, responsible for more than 60% of the water supplied to the city and the metropolitan region.

Due to the need to ensure future water supply for communities, anticipating the need for expansion of water sources and their infrastructure, it is necessary to conduct an analytical study on long-term water demand growth forecasting to ensure the prediction of necessary investments for this expansion.

According to Norvig and Russell (2021), an intelligent agent is an entity or software that perceives its environment through sensors and acts upon that environment through actuators, aiming to achieve a specific objective or task. In other words, an intelligent agent is any system capable of perceiving, processing information, and making decisions to perform actions that maximize its performance to achieve a specific goal.

In this scenario, the objective of this work is to present an ABMS (Agent-Based Modeling and Simulation) that can predict water demand in the metropolitan region of Salvador, based on the combination of historical consumption, geographic location, housing categories, population growth, and the number of residents, thereby aiding in domestic water management. Given the large volume of data and the high processing load required for agent execution and simulations, a data subset was selected for simulation purposes, focusing on the Itapuã neighborhood, the largest and most representative in terms of the number of consumer units, environmental diversity, terrain, and social contrast.

Agent-based modeling can be useful for reproducing scenarios where various characteristics can affect water demand. In the proposed model, the future consumption of each household will be inferred based on historical consumption, housing categories, the number of residents, and family growth projections.

The methodology used is represented in Figure 1, carried out in stages, which are described throughout this study. It begins with the definition and contextualization of the problem in relation to the current scenario of the Integrated Water Supply System of the Metropolitan Region of Salvador (MRS) in Section 1, the introduction. Section 2 presents a brief history of related works. The proposal is presented in Section 3, including the pre-design of agents (PEAS). The modeling and preliminary architecture are presented in Section 4. The prototype, along with data collection, is explored in Section 5. Results and discussions are presented in Section 6. Finally, the conclusions are presented in Section 7.

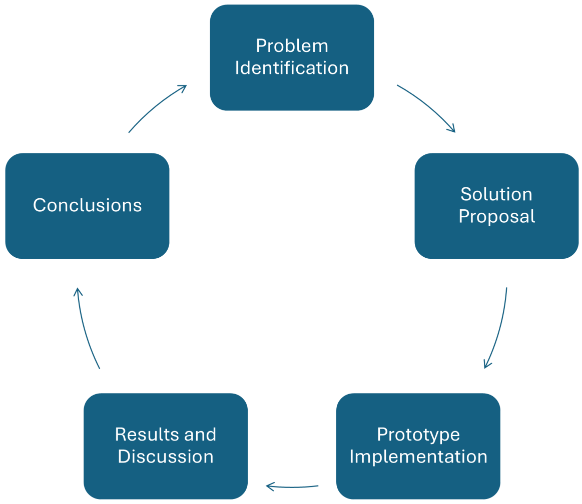


Figure 1. Work methodology

2. Related Work

In the literature, numerous studies emerged to accurately predict water demand by applying AI and ML techniques, developing methods and models capable of handling distinct time horizons: short, medium, and long-term. These advancements aim to improve water resource management by optimizing processes and reducing operational costs. However, few papers explore the agent-based approach [Ehsanitabar et al. 2025], [Vidal-Lamolla et al. 2024], [Koutiva and Makropoulos 2016], and [Galán et al. 2009]. Most papers focus on statistical, optimization, and ML models that incorporate randomness or uncertainty in their forecasts, as seen in [Iwakin and Moazeni 2024], [Kavya et al. 2023], [VanBerlo et al. 2021], and [Rees et al. 2020].

[Ehsanitabar et al. 2025] simulated demand management on a reservoir, showing that incentive-based scenarios (e.g., drought awareness and water-saving facilities) significantly increased water availability. [Vidal-Lamolla et al. 2024] explored Demand-Side Management (DSM) policies under water scarcity, combining price and non-price strategies across 125 simulations, revealing reductions of 8.1–15.6

Similarly, [Koutiva and Makropoulos 2016] modeled domestic consumption in Athens during a drought (1988–1994), combining behavioral and demand models to guide effective public campaigns. In [Galán et al. 2009], an ABMS linked with GIS simulated water use in Valladolid, capturing the impact of urban, technological, and social dynamics. These works underscore the capacity ABMS to integrate complex variables for realistic scenario modeling.

In this context, this paper propose ABM-WP for water demand forecasting in metropolitan Salvador, Brazil. Our spatially explicit model uses autonomous agents to simulate consumption patterns and test adaptive strategies. This contributes to building more resilient, efficient, and data-driven urban water supply systems.

3. Proposal

The proposal of this work consists of the development of an ABMS (Agent-Based Modeling System) capable of predicting water demand in the Metropolitan Region of Salvador (RMS). To achieve this, the model will integrate multivariate data, including historical consumption, geographic location, household categorization, population growth projections, and the number of residents per household. The implementation of this system aims to enhance domestic water management, providing valuable insights for optimizing water supply and mitigating challenges related to scarcity and efficient distribution of water resources.

For the development of this project, a preliminary survey of information about the target audience was necessary:

- The **Municipality of Salvador** was chosen as it is the most representative in terms of the number of inhabitants in the entire Metropolitan Region (13 municipalities).
- According to IBGE estimates, in 2024, Salvador had approximately 2,568,928 million inhabitants, out of a total of 3,623,647 in the RMS.
- From this universe, **Residential** water connections with **Occupied** properties were selected, resulting in **542,148 consumer units**.
- The 15 most representative neighborhoods in Salvador in terms of the number of water connections were listed, as shown in Table 1.

Table 1. List of the most representative neighborhoods in terms of the number of water connections.

Water connections	Neighborhoods
16,153	Itapuã
13,145	Paripe
12,787	Sã Cristóvão
12,770	Faz G do Retiro
11,718	Pernambués
11,633	São Caetano
10,553	Periperi
10,240	Beiru T Neves
9,972	Boca do Rio
9,988	Liberdade
9,463	Valéria
9,455	Brotas
8,248	Pirajá
8,213	Uruguai
8,181	Plataforma

Due to the large volume of data and the high processing load required for agent instantiation and simulation execution, a subset of data was selected for simulation purposes. The neighborhood of Itapuã was chosen, with 16,193 connections, being the largest and most representative in terms of the number of consumer units, environmental adversity, terrain, and social contrast.

The neighborhood of Itapuã is located in the eastern zone of Salvador, Bahia, bordering the neighborhoods of Stella Maris and Jardim Armação. It encompasses the Abaeté Metropolitan Park, an environmental preservation area. Figure 2 shows the neighborhood of Itapuã, where it is possible to observe that a large portion of the residences are concentrated around the park area, located on the right side of the map. This has resulted in a high concentration of inhabitants per hectare in specific points of the neighborhood.

Figure 2. Demarcation of the Itapuã neighborhood, where the Abaeté Metropolitan Park, located in an environmental preservation area, can be observed on the right.
Source: Adapted from Machado, Barros and Moreira (2022).

As evidenced in Figures 3 and 4, the Itapuã neighborhood exhibits a significant population concentration in higher altitude areas. This geographic distribution has direct implications for the water supply system, as water pressure varies with altitude. In lower altitude regions, pressure is naturally higher due to gravitational force, while in higher areas, pressure is reduced. As a result, elevated regions often receive water only after the network is saturated and the lower areas, where pressure is higher, are fully supplied. This phenomenon can lead to operational challenges, such as the need for additional pumping systems to ensure adequate supply in all areas.

For water demand forecasting, a multi-agent simulation model was proposed, based on a modeling approach with three interconnected agents, each with specific functions:

- **Analyzer Agent (AA)**: Responsible for collecting and analyzing consumption, geographic, and socioeconomic data. This agent consolidates and groups data by family, providing a structured basis for the next stage.
- **Predictor Agent (AP)**: Uses the data consolidated by the Analyzer Agent to simulate family behavior. Through predictive models, this agent reproduces the consumption function based on the provided variables, generating individualized forecasts.
- **Communicator Agent (AC)**: Consolidates the forecasts generated by the Predictor Agent and issues consumption forecast reports by family. This agent acts as

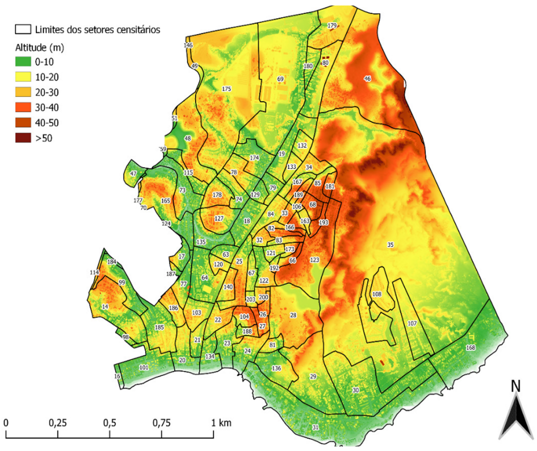


Figure 3. Hypsometry by census sector in the Itapuã neighborhood. Source: Adapted from Machado, Barros and Moreira (2022).

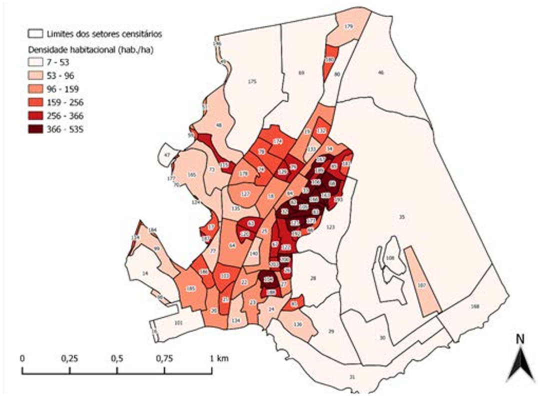


Figure 4. Number of inhabitants per hectare in the Itapuã neighborhood. Source: Adapted from Machado, Barros and Moreira (2022).

the final interface, ensuring that information is transmitted clearly and accessibly to decision-makers.

Table 2 present PEAS (Performance Measures, Environment, Actuators, and Sensors) for Predictor Agent (AP).

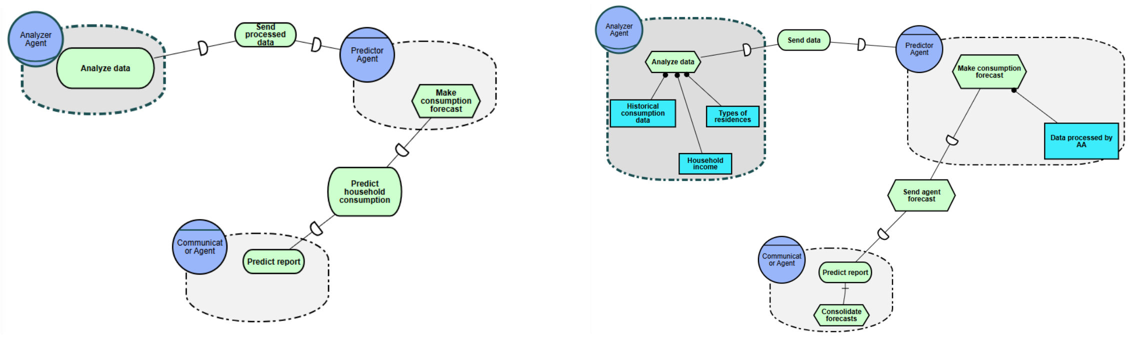
Table 2. PEAS description for the AP.

PEAS	Description
Performance	Maintain the quality of predictions, with the lowest possible MAPE.
Environment	Virtual with georeferenced addresses. Classified as stochastic (future events cannot be predicted with certainty), dynamic (changes independently of the agent's intervention), continuous, partially observable, and sequential (previous decisions influence the future state of the environment).
Actuators	Prepare data; choose the best prediction model; perform consumption forecasts by family; make forecasts available to the AC.
Sensors	Data generated by the AA.

4. Modeling

TROPOS is a methodology for agent-oriented software development that employs a goal-based and incremental design approach, aiming to align the system with the needs of stakeholders through the identification, refinement, and achievement of goals. According to Bresciani et al. (2004), TROPOS divides the development process into phases ranging from requirements analysis, focusing on identifying stakeholders and their goals, to architectural and detailed design, where agents, their responsibilities, and interactions are defined. The methodology is particularly suitable for complex systems involving multiple autonomous agents, promoting modeling that reflects both dependencies among agents and their collaborations to achieve common goals.

The TROPOS modeling approach was used to structure and coordinate the interaction between agents, each with specific functions, such as data collection, predictive analysis, and communication of results. Figures 5a and 5b show, respectively, the initial and final requirements defined for the project.



(a) Initial requirements in TROPOS modeling, using PiStar.

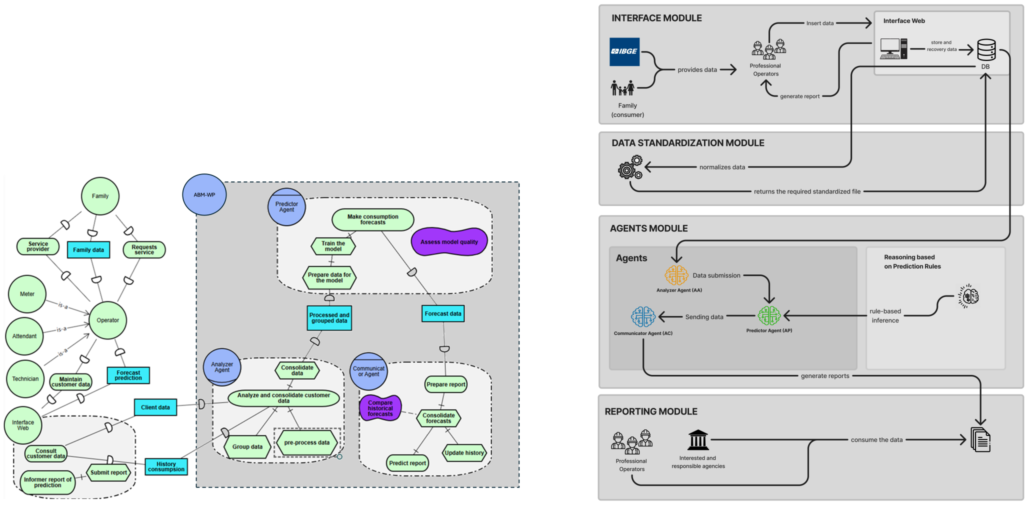
(b) Final requirements in TROPOS modeling, using PiStar.

Figure 5. TROPOS modeling for early and late requirements using PiStar.

Figure 6a shows the overall architecture of the proposed system, named Agent-Based Model for Water Prediction (ABM-WP), encompassing the interaction between actors and agents. Built based on TROPOS modeling, the process begins with the Family, which interacts with the system through collaborators (meter readers, attendants, or technicians), who consult and register information in a Web Interface, providing consumption data. The Analyzer Agent (AA) collects and processes user data, grouping and preparing it for analysis. Next, the Predictor Agent (AP) uses this processed data to generate individualized consumption forecasts per connection, while also evaluating the quality of the predictive model. Finally, the Communicator Agent (AC) consolidates the forecasts, compares them with historical data, and generates reports that are made available to collaborators through a reporting interface. This multi-agent architecture allows for a modular and scalable approach, where each agent plays a specific role, ensuring the accuracy of forecasts and the efficiency of water resource management.

The software architecture used was that of reactive agents, which employ a reasoning model based on inference rules, with a vertically layered structure. The architecture is organized into 04 layers, structured as modules:

- **Interface Module:** focuses on the role of data input providers, actors responsible for manual data collection, and the web interface that receives and stores multiple data inputs.
- **Data Normalization Module:** currently responsible for processing location and housing type data, but scalable as needed by the project.
- **Agent Module:** handles communication between the involved agents and the flow of exchanges based on their responsibilities.
- **Reporting Module:** responsible for generating reports based on the analyses performed by the system using the provided data.



(a) TROPOS architectural model of the ABM-WP System.

(b) Architectural overview of the ABM-WP System implementation.

Figure 6. Architectural representations of the ABM-WP System from different modeling perspectives.

The reasoning model based on data and inference rules uses historical consumption data, population growth forecasts, housing type, and the number of residents to estimate future consumption. An inductive approach was used to identify consumption patterns from historical data, but deduction is also applied through predefined rules to infer consumption behaviors. For example, if the number of residents increases, consumption increases.

5. Prototype Implementation

The prototype was implemented using the GAMA platform (GIS Agent-Based Modeling Architecture), an open-source and user-friendly modeling and simulation environment designed for creating agent-based simulations with explicit spatial representation. Developed for application in various domains, such as urban mobility, climate change adaptation, epidemiology, disaster evacuation planning, and urban planning, GAMA offers a general and flexible approach to agent-based modeling. This flexibility is enhanced by its high integration capability, allowing the development of plugins for specific needs and interaction with other languages and software, such as R and Python. Prior to the prototype implementation in GAMA, data preparation was conducted, including preprocessing and normalization using Python, primarily for geographic coordinate conversion and data cleaning.

In the context of this work, GAMA was used for model definition and simulation, utilizing the GAML (GAMA Modeling Language), which provides robust features for creating agent behaviors and interactions. In the data preprocessing and normalization stage, Python played a crucial role in converting geographic coordinates, as shown in Figure 7a, and in data cleaning, ensuring the quality and consistency required for modeling. This combination of tools enabled an efficient and precise prototype implementation, aligning with the study's objectives and demonstrating GAMA's versatility in scientific applications.

(a) Python script used for converting geographic coordinates from EPSG:4326 to EPSG:31984, compatible with GAMA.

(b) GAML script defining the `Residencia` agent responsible for consumption prediction.

Figure 7. Scripts used in the implementation of the simulation model.

Initially, data characterizing the consumer unit were collected, including a unique anonymized primary key for user identification, residence type, number of residents, and geolocation. Subsequently, the consumption history for the last 12 months, from January 2024 to December 2024, was also obtained. In this case, the dataset contained the anonymized foreign key of the user, consumption in liters, and the reference month/year. In the next stage, the data were imported into a Python editor, specifically Microsoft@ Visual Studio, for handling null and invalid values, converting geographic coordinates to a format compatible with the GAMA platform, and extracting the average consumption per user primary key. All data processed in Python were temporarily saved as CSV (Comma-Separated Values) files, a widely used plain text file format for storing and exchanging tabular data, with a semicolon used as the delimiter in this case. Population growth data were obtained from IBGE statistical projections [IBGE - Instituto Brasileiro de Geografia e Estatística 2022] for the years 2024 to 2060 for the State of Bahia, with the annual total consolidated and the annual growth percentage calculated for the next 10 years. This percentage was then imported into the GAMA platform and used as input for the consumption projection function.

In the GAMA platform, the processed consumer unit data, including geolocation, residence type, number of residents, and average consumption per unit, were loaded and instantiated into an agent responsible for predicting individual household consumption, as shown in Figure 7b, which depicts the `Residencia` species with its `reflex` method responsible for consumption prediction.

6. Results and Discussion

After implementing the prototype, several simulation tests were conducted to evaluate and refine the prediction model. Two scenarios were modeled and simulated:

- Scenario I - Water demand forecast for a 10-year period, considering only the population growth rate.
- Scenario II - Water demand forecast for a 10-year period, considering the population growth rate, residence category, and the presence of a swimming pool.

Comparing the simulations between Scenarios I and II, it is evident that when considering only the population growth rate for the 10-year period (2025 to 2034), the consumption forecast grows at a lower rate in Scenario I. While Scenario I shows an average annual growth of approximately 0.15%, Scenario II grows at an average rate of 1.19% per year. This results in a significant cumulative difference by the end of the period, with Scenario II reaching a consumption of 163,025.07 m^3 in 2034, compared to 148,734.33 m^3 in Scenario I, as shown in Table 3.

Table 3. Forecasted Consumption for Scenarios I and II (2025-2034)

Year	Scenario I Consumption (m^3)	Δ	Scenario II Consumption (m^3)	Δ	Growth Rate (%) CI vs CII
2025	146,551.39	-	146,551.39	-	-
2026	146,947.08	0.27%	148,176.40	1.11%	0.84%
2027	147,314.45	0.25%	149,851.83	1.13%	0.88%
2028	147,638.54	0.22%	151,566.88	1.14%	0.92%
2029	147,919.05	0.19%	153,325.69	1.16%	0.97%
2030	148,170.51	0.17%	155,148.17	1.19%	1.02%
2031	148,377.95	0.14%	157,024.00	1.21%	1.07%
2032	148,541.17	0.11%	158,958.32	1.23%	1.12%
2033	148,660.00	0.08%	160,956.69	1.26%	1.18%
2034	148,734.33	0.05%	163,025.07	1.29%	1.24%

Spatial analysis of consumer units in urban areas is essential to understanding the distribution and consumption of water across different regions. Using the GAMA platform, it was possible not only to model the multi-agent system but also to visualize the spatial distribution of these units. Figure 8 provides an overview of the neighborhoods in Salvador, highlighting the concentration of consumer units in Itapuã, represented in blue. Figure 9 offers a more detailed view of public roads and the spatial distribution of consumer units in the neighborhood, where each unit corresponds to a Residential agent instantiated in the GAMA platform. This approach allows for a more precise analysis of urban dynamics and water consumption, contributing to efficient water resource planning and management.

7. Conclusions

Initial tests indicated that agent-based modeling allows for a more granular and adaptable prediction of water demand. Each residence was modeled as an autonomous predictive agent, enabling the observation of consumption patterns at different scales applied to each household. It also allows for an analysis of how population density and residence category influence water distribution.

In this study, the model estimates household consumption based on residence type and population growth trends, with projections made for the next 10 years for the Itapuã neighborhood, as shown in Figure 10.

As suggestions for future work, it is worth highlighting the need to expand the scenarios for simulation purposes, with the possibility of parameterizing other characteristics that also influence the water supply system, such as:

- The latitude of each consumer unit;
- The geographic distribution of household monthly income;

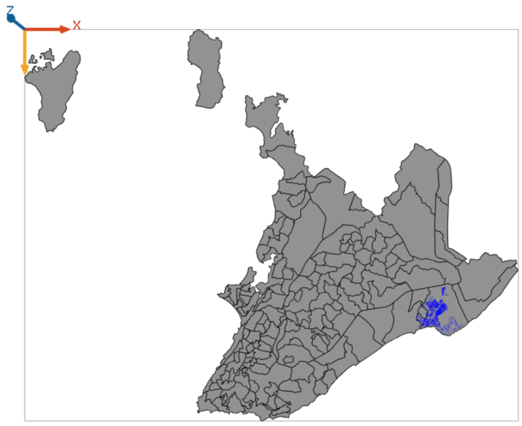


Figure 8. Plot of Salvador neighborhoods in the GAMA platform, showing the spatial distribution of consumer units concentrated in blue in the Itapuã neighborhood.

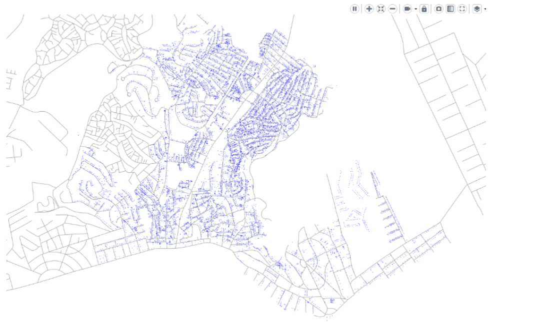


Figure 9. Zoomed-in plot of public roads in the Itapuã neighborhood, showing the spatial distribution of consumer units, where each unit represents a `Residencia` agent instantiated in the GAMA platform.

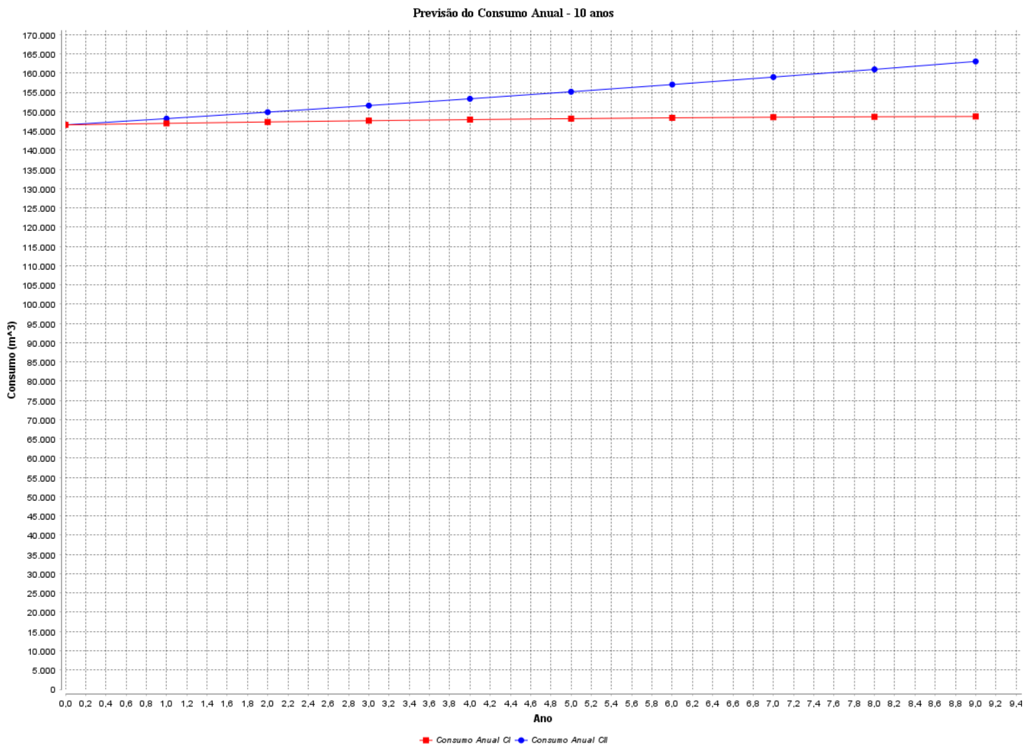


Figure 10. Consumption projection for the next 10 years.

- The possibility of adopting technologies that may influence consumption behavior, such as low-flow electric showers, as suggested by [Galán et al. 2009];
- Extending the simulation dataset to more powerful computational environments to enable city-wide simulations for Salvador.

We thank the Empresa Baiana de Águas e Saneamento (EMBASA) and its technicians for providing essential data and technical information crucial to this research.

We also thank the Federal University of Bahia (UFBA), through the Graduate Program in Computer Science (PGCOMP), for enabling this study by providing infrastructure, guidance, and the expertise of its faculty.

References

Alves, J. E. D. (2019). Crescimento populacional e mudanças climáticas vão aumentar o estresse hídrico. Accessed: Feb. 2025.

Ehsanitabar, A., Hassanzadeh, Y., Aalami, M., and Sadeghfam, S. (2025). Agent-based modeling for demand management of reservoirs considering social and hydrological interactions under uncertainty. *Journal of Environmental Management*, 380:125089.

Galán, J. M., López-Paredes, A., and Del Olmo, R. (2009). An agent-based model for domestic water management in valladolid metropolitan area. *Water Resources Research*, 45(5).

IBGE - Instituto Brasileiro de Geografia e Estatística (2022). Projeções de população. Accessed: Feb. 2025.

Iwakin, O. and Moazeni, F. (2024). Improving urban water demand forecast using conformal prediction-based hybrid machine learning models. *Journal of Water Process Engineering*, 58:104721.

Kavya, M., Mathew, A., Shekar, P. R., and P, S. (2023). Short term water demand forecast modelling using artificial intelligence for smart water management. *Sustainable Cities and Society*, 95:104610.

Koutiva, I. and Makropoulos, C. (2016). Modelling domestic water demand: An agent based approach. *Environmental Modelling & Software*, 79:35–54.

Rees, P., Clark, S., and Nawaz, R. (2020). Household forecasts for the planning of long-term domestic water demand: Application to london and the thames valley. *Population, Space and Place*.

Rosa, A. M. F., Souza, G. B. G., and Cavalcante, M. M. (2022). A urbanização do bairro de itapuã, salvador (ba). Accessed: Feb. 2025.

VanBerlo, B., Ross, M. A. S., and Hsia, D. (2021). Univariate long-term municipal water demand forecasting. `https://arxiv.org/abs/2105.08486`. arXiv preprint arXiv:2105.08486.

Vidal-Lamolla, P., Molinos-Senante, M., Oliva-Felipe, L., Alvarez-Napagao, S., Cortés, U., Martínez-Gomariz, E., Noriega, P., Olsson, G., and Poch, M. (2024). Assessing urban water demand-side management policies before their implementation: An agent-based model approach. *Sustainable Cities and Society*, 107:105435.