Journal of Environmental Management 351 (2024) 119680

Contents lists available at ScienceDirect

Journal of Environmental Management

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

Research article

Modeling spatiotemporal domestic wastewater variability: Implications for measuring treatment efficiency

Néstor DelaPaz-Ruíz [a,*], Ellen-Wien Augustijn [a], Mahdi Farnaghi [a], Raul Zurita-Milla [a]

[a] *Department of Geo-Information Processing (GIP), Faculty of Geo-Information Science and Earth Observation (ITC), University of Twente, Drienerlolaan 5, 7522 NB, Enschede, the Netherlands*

A R T I C L E I N F O

Keywords:
Agent-based model
Water quality
Time series
Sanitation regulations
Spatiotemporal resolutions
Decentralized treatment plants

A B S T R A C T

Continuously measuring the efficiency of wastewater treatment plants is crucial to progress in sanitation management. Regulations for decentralized wastewater treatment plants (WWTP) can include rudimentary specifications for sporadic sampling, unencouraging continuous monitoring, and missing crucial domestic wastewater (DW) variability, especially in low- and middle-income countries. However, few studies have focused on modeling and understanding spatiotemporal DW variability. We developed and calibrated an agent-based model (ABM) to understand spatial and temporal DW variability, its role in estimated WWTP efficiency, and provide recommendations to improve sampling regulations. We simulated DW variability at various spatial and temporal resolutions in Santa Ana Atzcapotzaltongo, Mexico, focusing on chemical oxygen demand (COD) and total suspended solids (TSS). The model results show that DW variability increases at higher spatiotemporal resolutions. Without a proper understanding of DW variability, treatment efficiency can be overestimated or underestimated by as much as 25% from sporadic sampling. Sensor measurements at 6-min intervals over 3 hours are recommended to overcome uncertainty resulting from temporal variability during heavy drinking water demand in the morning. Reporting of sewage catchment areas, population sizes, and sampling times and intervals is recommended to compare WWTP efficiencies to overcome uncertainty resulting from spatiotemporal variability. The proposed model is a useful tool for understanding DW variability. It can be used to estimate the impact of spatiotemporal variability when measuring WWTP efficiencies, support improvements to sampling regulations for decentralized sanitation, and alternatively for designing and operating WWTPs.

1. Introduction

Despite untreated wastewater being a serious hazard for humanity and the environment, the global percentage of treated domestic wastewater (DW) is low. According to the United Nations, only 56% of DW flow worldwide is treated (UN-Water, 2020). The percentage is lower in lower middle-income countries (28%) and even lower in low-income countries (8%) (Sato et al., 2013), meaning that 3.5 billion people worldwide are not using safely managed sanitation (The World Bank, 2020). Financing is a critical issue for DW management (Ujang and Mogens, 2006) because adequate financing for the operation of DW infrastructure is essential to avoid abandonment and deterioration of DW treatment facilities and inefficiencies in DW treatment. The United Nations Statistics Division (2020) suggests that efforts to increase and sustain sanitation must quadruple, and reporting wastewater treatment plant (WWTP) efficiency is vital for tracking progress.

Measuring WWTP efficiency can be defined by the difference in percentage between the pollutant's concentrations at the inflow and outflow of a treatment plant (Di Cicco et al., 2021), which is not a trivial task. Because of rudimentary (or lacking) regulations for DW sampling, the infeasibility of continuous monitoring through laboratories, and the substantial DW dynamics of pollutant loads, WWTP efficiencies are mostly unknown and undocumented (UNSTATS, 2022). First, sampling regulations often do not consider using specific sensors that fully cover higher temporality to capture pollutant loads of DW in WWTPs. The regulation gap exists in European countries and Mexico (European Union law, 2014; European Commission, 2019; SEGOB, 2022a), where sampling is mostly based on laboratories. Specifically, such regulations should be complemented in low-middle income countries, allowing DW sensors to make monitoring WWTP efficiency economically feasible in the long term. Second, laboratory analysis for continuous monitoring is economically unfeasible for many countries. Decentralized WWTPs

* Corresponding author.
E-mail address: n.delapazruiz@utwente.nl (N. DelaPaz-Ruíz).

https://doi.org/10.1016/j.jenvman.2023.119680
Received 31 January 2023; Received in revised form 7 November 2023; Accepted 10 November 2023
Available online 5 December 2023

increase the complexity of measuring treatment efficiencies, which can have a dozen WWTPs at peri-urban and rural localities without access to certified laboratories. In contrast, centralized urban treatment at larger and gauged WWTPs allows better access to continuous monitoring. DW sampling frequencies must differentiate centralized and decentralized treatment to provide long-term monitoring of efficiencies. Lastly, the substantial DW variability (Ujang and Mogens, 2006; Dubois et al., 2022) questions the representativeness of a few laboratory-based samples. These variations are driven by population behaviors (the use of water appliances) and population mobility in time (Atinkpahoun et al., 2018), i.e., the different numbers of people at a certain location. Hence, it is crucial to understand the significant implications of spatial and temporal DW variability when assessing DW treatment efficiencies. Focusing on temporal variability, Di Cicco et al. (2021) showed that strong TSS variability over time calls into question whether sporadic sampling truly describes the operating conditions of treatment plants. Penn et al. (2017) analyzed small-scale diurnal patterns using data from 170 households and showed that the flow across a wastewater network is highly variable over time. Zhou et al. (2019) demonstrated high wastewater variability using two case studies with WWTPs serving 680, 000 and 141,500 inhabitants, with the smaller treatment plant exhibiting higher variability. Similarly, high variability has been shown for COD (Rodríguez et al., 2013). Ideally, a sampling regulation of treatment efficiency must clarify the differences in spatiotemporal resolutions between (de)centralized treatment systems and standardized sampling resolutions. Modeling DW variability at multiple spatiotemporal resolutions can contribute substantially to wastewater sanitation.

Although DW modeling is important for improving wastewater management, few studies have focused on modeling DW at multiple spatiotemporal resolutions to clarify variability in space and time. Dubois et al. (2022) analyzed DW pollution, providing further knowledge about DW variability. Still, their study is limited to the household's spatial resolution without modeling DW pollution. De Keyser et al. (2010) modeled DW time series, providing different temporal resolutions for designing WWTPs without analyzing the relevance of time series resolutions and DW variations for measuring WWTP efficiencies. Penn et al. (2017) modeled DW volume, which allows for reporting multiple spatial sections across the sewage. However, DW pollutants and a comprehensive understanding of the resolutions are not provided. Although Rodríguez et al. (2013a,b) modeled DW quality and quantity and considered sub-catchment areas, they did not analyze multiple spatiotemporal resolutions. Jia et al. (2021) reviewed 110 papers published in the last decade on modeling water quality in sewage and suggested that multiple resolutions should be considered in future wastewater modeling studies.

Mexico is a middle-income country that faces many of the problems mentioned above. It encounters challenges in sustaining DW treatment, with 63% of wastewater treated (Tabla-Vázquez et al., 2020). Since 1996, a norm for the maximum pollutant concentration of wastewater discharge (SEGOB, 2022a) has been applied by the National Water Commission (CONAGUA). Operation rules for sanitation also exist (SEGOB, 2022b), quantifying DW flow discharge but ignoring the efficiency of pollutant removal. The percentage of safely treated wastewater (PSTW) is an indicator that refers to pollutant removal in WWTPs. The closest sampling regulation linked to PSTW defines instantaneous sampling every six months carried out by accredited laboratories with a temporal interval of up to 4 h on the sampling day (SEGOB, 2022a). Mexico has a combination of centralized and decentralized treatment systems. In addition, the lack of cost recovery through user fees puts a large financial restriction on wastewater management (OECD, 2010).

To better understand how the efficiency of decentralized wastewater treatment plants should be monitored, we propose creating a model to simulate the DW variability focusing on Chemical oxygen demand (COD) and total suspended solids (TSS). COD and TSS are among the PSTW that are not continuously monitored, thereby preventing the generation of statistics on DW treatment status worldwide (WHO et al., 2020). In this study, i) we investigate spatial and temporal DW variability, ii) illustrate the implications of the spatiotemporal DW variability when measuring WWTP efficiency (percentage of safely treated DW), and iii) introduce basic recommendations for sampling regulations to measure the efficiency of decentralized and ungauged WWTPs.

2. Methods

Fig. 1 presents an overview of the main steps in modeling and evaluating spatiotemporal DW dynamics. An agent-based model (ABM) is conceptualized and subsequently used to study DW variability. The ABM is complemented by spatial microsimulation (SMS), generating the required synthetic population for the model. The model is based on the DW literature and calibrated with data on sampled DW. Postprocessing generates DW time series at various temporal resolutions. The ABM model is validated by comparing the results with data collected in situ. The following subsections provide more details on each of these steps. Using this ABM allowed us to quantify and understand spatiotemporal DW variability and discuss implications and recommendations for considering DW variability in assessing WWTP efficiencies. Case studies with combined sewer systems should consider the potential effects of rainfall, as discussed in Section 4.

2.1. Study area and data

The selected study area is Santa Ana Atzcapotzaltongo, a rural locality in Hidalgo, Mexico (Fig. 2). Santa Ana has two decentralized WWTPs (Shaded areas in Fig. 2). The targeted WWTP is in the southwestern catchment area and is based on biodigesters as a secondary sanitation treatment so that treated DW can be used for agriculture. The census (INEGI-DENUE, 2017) shows that most DW comes from domestic activities, where wastewater from economic activities is minimal, and there is no production of industrial wastewater.

Population data were obtained from a census generated by the Mexican federal census survey (INEGI, 2020) through the INEGI microdata lab. The census contains information related to population and neighborhood blocks. Santa Ana has 1678 inhabitants, and an estimated 758 people on the southwestern side of the locality are discharging DW at the targeted WWTP catchment area that consists of a separated sewer system, as seen in Fig. 2. The sewage design was digitized via a dedicated field data collection campaign in 2022, in which pipe sections, collectors, accessible maintenance holes, catchment coverage, and network connectivity were verified. For the digitalization, we used topographic data from a digital elevation model based on *SRTM Tiles acquired* via *the NASA Server* QGIS plugin (Duester, 2021) and the *Drinking Water, Sewerage and Sanitation Manual* (CONAGUA, 2019) that provides the design and regulation criteria applicable for Mexico.

Data on DW from water fixtures and appliances (toilets, kitchen and bathroom sinks, showers, and washing machines) were taken from Almeida et al. (1999) and Rose et al. (2015). The data represents water fixture and appliance discharges (l) and DW pollutants (mg/l), split into COD and TSS. Preparations of individual events of pollutant loads for feces and urine are required based on literature (Almeida et al., 1999; Rose et al., 2015), as literature does not represent events after flushing the toilet. Details of estimated urine and feces events that represent pollutant loads after flushing the toilet are included in the supplementary materials. DW was sampled between March 19 and 23, 2022, matching the study area's dry season, thus allowing the assessment of the pollutant variability from human sources without dilution. The sampling focused on simultaneous spatial and temporal measurements of COD and TSS at catchment and sub-catchment coverage. Spatially, the WWTP catchment and sub-catchments 258 and 39 measured at maintenance holes (see Fig. 2 delineated areas) were selected to study differences between DW variability based on population and area size. Temporally, targeted catchment and sub-catchments were synchronously sampled to compare the DW variability at the same time slots.

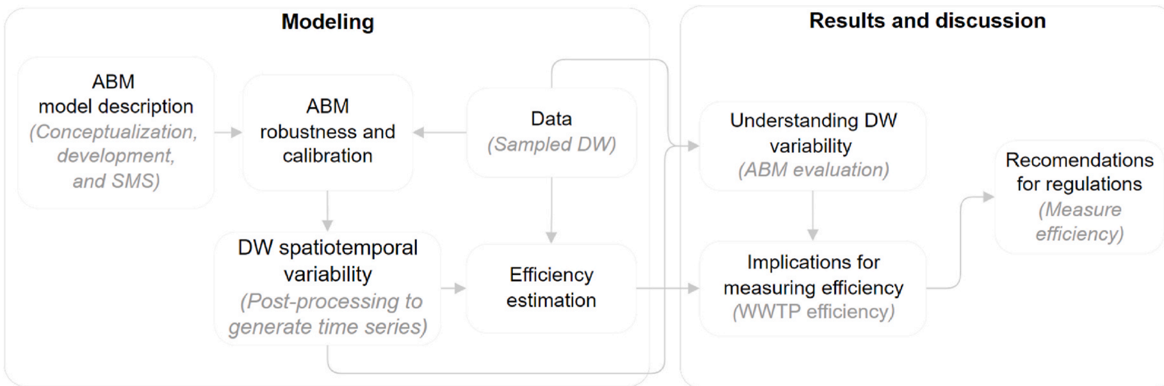


Fig. 1. Main steps and components to model and evaluate DW dynamics at multiple resolutions linked to results and discussion.

Sampled DW was measured in situ at a frequency of 6 min with a calibrated spectrometer sensor (scan, 2022).

2.2. Agent-based modeling

2.2.1. ABM model description

ABMs are computational simulations of individual, dynamic, and adaptive behavior agents with multiple characteristics providing interactions between them and their environment (McLane et al., 2011). The ABM model was first conceptually designed (Fig. 3) with a practical methodological guide for modeling with ABMs based on Auchincloss et al. (2015) and then translated into a computational model in NetLogo 6.1.1. Information on the production and dynamics of DW in the literature was used to propose the ABM (Almeida et al., 1999; Von Sperling, 2007; Henze and Comeau, 2008; Mesdaghinia et al., 2015; Penn et al., 2017; Atinkpahoun et al., 2018).

Because detailed data from the population census are unavailable because of privacy protection, this study uses SMS to generate a synthetic population imported in the ABM model (i.e., the number of residents, their gender, and education). Lovelace and Dumont (2018) describes SMS as an approach for studying phenomena at multiple spatial levels using georeferenced microdata, including the provision of detailed geolocated synthetic inhabitants. SMS uses two datasets collected by INEGI (2020): i) a non-geolocated inhabitant survey that has the advantage of providing detailed individual information on mobility (for studying and working purposes), age, and gender, and ii) a geolocated inhabitants survey that has the disadvantages of not containing individual information but aggregated data by neighborhood blocks. The output of SMS is a geolocated inhabitant dataset at the individual level that matches the two initial datasets to provide the best of both.

The description of the ABM is based on a synthesized version of the Overview, Design Concepts, and Details protocol (ODD) for describing ABMs provided by Grimm and Ayllón (2020). The overall purpose of the DW ABM is to simulate the DW variability of pollutants and volume discharge at multiple spatial and temporal resolutions. Specifically, we are addressing the following question: How understanding can be gained by modeling spatial and temporal DW variations that inhabitants produce? To consider our model realistic enough for its purpose, we evaluate and analyze time series patterns, plotting together the simulations and observations and showing Pearson correlations.

The ABM includes the following entities. Agents: Inhabitants and DW particles. Environment: houses, neighborhood blocks, economic points, sewage network, maintenance holes, and treatment plant. The information on the state variables characterizing those entities is listed in Table 1. Inhabitants are provided with information to produce DW particles and follow working or studying schedules interacting with the environment of houses, schools, and economic points. While inhabitants move around the locality (i.e., from houses to schools), generate DW particles. After producing particles, we simulate their motion in the sewer, maintenance holes (nodes), and finishing in the WWTP. All environments are static during the simulation. The sewage consists of pipes and maintenance holes (nodes) where DW particles move, simulating DW flow.

Concerning the spatial and temporal resolutions, which are the most important design concepts of the model, the model uses the NetLogo time extension (Sheppard et al., 2022) for explicitly reporting temporality and generating dynamic discrete event scheduling for simulating DW production, mobility, and discharge (see Fig. 3.a). All the events apply a probability schedule defining timestamps when an event takes place. The model executes every event with dynamic timestamp schedules based on hourly probabilistic distributions contrasting with the fixed temporal resolution of “ticks” in NetLogo. For instance, DW production events for each inhabitant differ from day to day, weekday to weekend, hour to hour, and even recorded minutes and seconds of execution. The dynamic timestamps allow the processing at multiple temporal resolutions of the DW variability. Additionally, the model is run and designed to simulate a minimum range of two days.

Spatially, the simulation coverage is defined by the shapefiles vector data from the INEGI census, where the execution of water appliances takes place. DW particle timestamps are registered at the catchment and sub-catchment through the WWTP and sewage maintenance holes. Due to the spatiotemporal records of the dynamic events, each ABM output of DW particles at the catchment and sub-catchment levels (two spatial resolutions) is aggregated at various temporal resolutions to evaluate the DW variability of COD and TSS in mg/l. The postprocessing of DW particles consists of aggregating each (sub)catchment and time windows (i.e., 60 min) with mean functions resulting in outputs comparable to sampled DW. The simulated variance and the analysis between the spatial and temporal resolutions are the basis for evaluating results, highlighted in the plots as shadow areas, as shown in the results section.

The process of the model initializes with creating the inhabitant agents based on the SMS outcomes. They are located at the houses of their respective neighborhood blocks, where households are created from the inhabitants of the same house. The inhabitant agent has variables like age and location (see Table 1). Inhabitant agent behavior includes using various water-using fixtures and appliances that produce DW events, such as toilets (for urination and defecation), sinks (in kitchens and bathrooms), showers, and washing machines. Events related to these fixtures and appliances follow occurrence probabilities and maximum and minimum uniform distributions of DW events (pollutant loads and discharge in liters) according to typical sequences of daily water usage activities. For inhabitants’ mobility, first, the agents remember their assigned household. The mobility of inhabitant agents includes going to and returning from work or school. Work and school locations are inside the whole locality (Fig. 2). Inhabitant agents store their home, work, and school locations in memory, ensuring they return to the same location. The model differentiates between working days

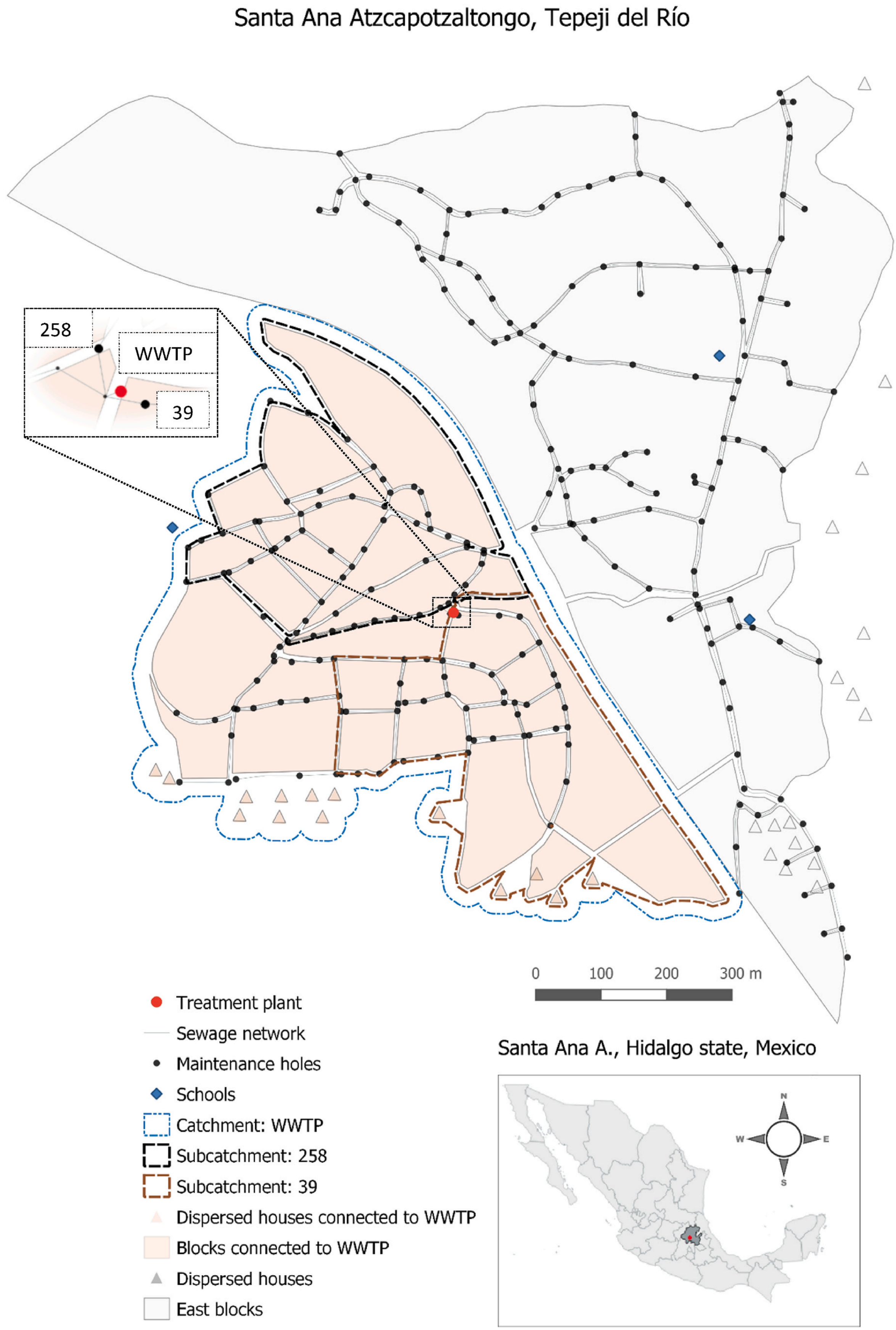


Fig. 2. Study area. Left: Numbers 39 and 258 are sampled maintenance hole IDs. A highway divides the locality's West (brown) and East (grey) sides. West side discharges at the WWTP. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

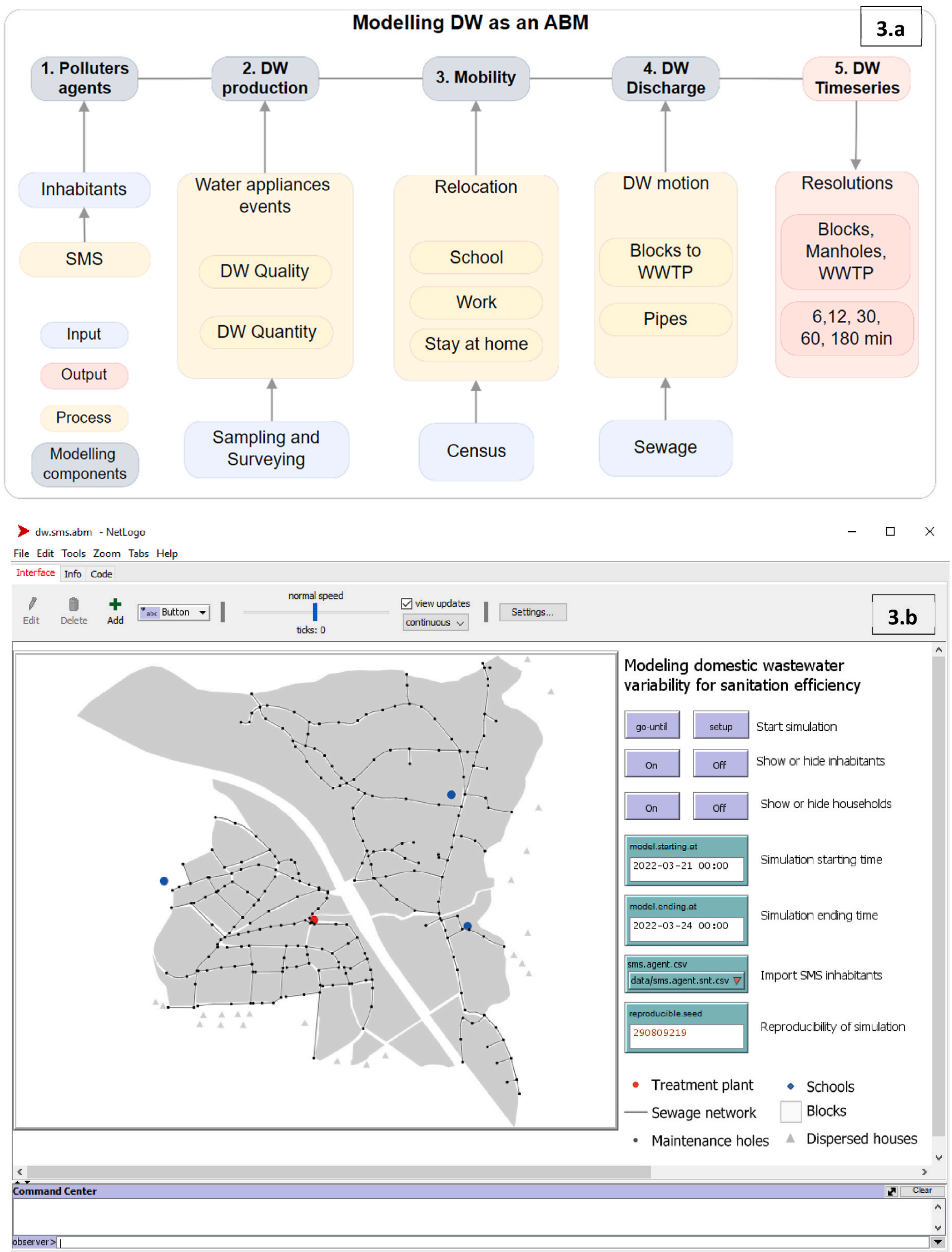


Fig. 3. Theoretical and computational conceptualizations of domestic wastewater. 3.a: Systemic components of the domestic wastewater dynamics 3.b: Translation of theory into a Netlogo simulation of the target locality.

and weekends.

The processes related to the DW particle agents are DW motion and discharge to the WWTP. DW particles are simulated to travel from houses (at the DW production event) via the sewage network to the WWTP, the discharge point where particles disappear. DW particles move with an average design velocity of 1.8 m/s in the sewage pipelines following Mexican regulations (CONAGUA, 2019). DW particle agents store timestamps records and locations in memory, including hatching (in households and blocks), dying (in WWTP catchment), and sub-catchment (when reaching nodes of maintenance holes). DW particles that reach the WWTP are generated from the western section of the locality, as shown in Fig. 2.

Table 1
List of model entities for the ABM. The "Possible value" column shows an example of the information available in the ABM for the entities. For example, quantitative and qualitative values are shown for the selected inhabitant based on census data and the SMS method. Note that qualitative values, e.g., "Female," are defined as static.

Entity	Variable	Description	Possible value	Units
Inhabitants	Age	Age category	18–24	years
	Study	Attend school	Yes	–
	School level	School level	Highschool	–
	Work	Actively working	Yes	–
	Gender	Inhabitant sex	Female	–
	CVEGEO	Block location	1,302,700,010,105,004	–
	Ind id	Individual ID	01,750	–
	Education level	Education grade	2	–
DW particle	DW type	Type of water appliance	Toilet urine	–
	DW quantity	DW liters	9	l
	DW quality	DW pollutants: COD and TSS	453	mg/l
	CVEGEO	Block location	1,302,700,010,105,004	–
	Ind id	Individual ID	01,750	–
	DW speed	Traveling particle speed	1.8	m/s
	Traveling path	Sewage network node	Node 2020	–
Houses	House id	House ID	304	–
	CVEGEO	Block location	1,302,700,010,105,004	–
Blocks	CVEGEO	Block location	1,302,700,010,105,004	–
Economic points	ID	Economic point ID	7,996,840	–
	CVEGEO	Block location	1,302,700,010,105,004	–
	Avg. workers	Average workers	3, 5, 15 ….	–
	School exists	School in this point	High school	–
Sewage network	Node	Connecting nodes	Node 11	–
	Travel path	Path to move DW particle	Node 684	–
Treatment plant	Final station	The final destination of the traveling path	–	–

2.2.2. Robustness analysis and calibration

The robustness analysis addresses how many simulations are required to obtain stable results. The accumulated average method is used to evaluate the model stability, targeting the variable production of DW particles by inhabitants. After running the model multiple times, the accumulated average across all runs is plotted to find the expected number of runs for stable results. Cumulative distributions and average values from several simulations demonstrate variability effects in ABMs, including ABMs with water applications (Zechman, 2011; Gilbert, 2011). The simulation for robustness quantifies the DW particles produced by inhabitants during five days. At the end of each simulation, the DW particles produced are quantified. Fig. 4 shows the cumulative average (CA) for a simulation (n) based on the total DW particles (TP). The Y axis is the CA at each simulation (n). For instance, from Heckert (2003), the first simulations are: $CA_{n=1} = (TP_{n=1})$; $CA_{n=2} = (TP_{n=1} + TP_{n=2})/2$; $CA_{n=3} = (TP_{n=1} + TP_{n=2} + TP_{n=3})/3$, etc. Fig. 4 shows that the model does not significantly change between 25 and 100 runs of the simulations. We reckon that the model is stable starting from 25 simulations.

The calibration requires calculating the Pearson correlation between the simulated and observed DW time series. The Pearson correlations are also useful for analyzing minimum and maximum input value ranges. Calibration deals with identifying ranges of realistic input values to improve the results. The variation between DW pollutants of a simulated and observed day is analyzed by considering data inputs of urine and feces pollutant events based on input data from the literature (See section 2.1), which describe loads and distributions of occurrence during the day. Pollutant loads from urine and feces (see Section 2.1 and the supplementary material) are selected for testing input values, which are parameters highly variable in the model. The selection of urine and feces is a parameter prioritization testing subsets of inputs, which facilitates the analysis of results, as Borgonovo et al. (2022) suggest. Simulated time series are calibrated with hourly resolutions to match the model's design of hourly probabilistic distributions and to maintain a low-to-moderate computational processing cost.

The simulation runs as many times as suggested by the robustness analysis. Simulated time series outputs are averaged to provide a representative time series for each pollutant. Average, maximum, and minimum values are provided to visualize the variance in the simulations (see shadow areas in Fig. 5.b). We changed the parameter values and evaluated the results to obtain values that fit the time series best. When simulated and observed DW time series are similar, with correlation values greater than 0.6 (Schober and Schwarte, 2018), we consider there to be a match in patterns and the model to be calibrated. Calibrated values that provide the expected correlations are provided in the supplementary material. During calibration, it was sufficient to focus on feces and urine distributions; pollutant ranges were maintained as suggested by references cited in Section 2.1.

Fig. 5.b shows the correlations obtained at the inflow of the WWTP. At 60 min resolution, the Pearson correlations are 0.89 and 0.86 for COD and TSS, respectively. P-values and the number of observed points also confirm the reliability of the results during calibration. It is noted that the calibration criteria focused on analyzing the WWTP for COD and TSS where DW flow (l) was not possible to evaluate for the lack of observed

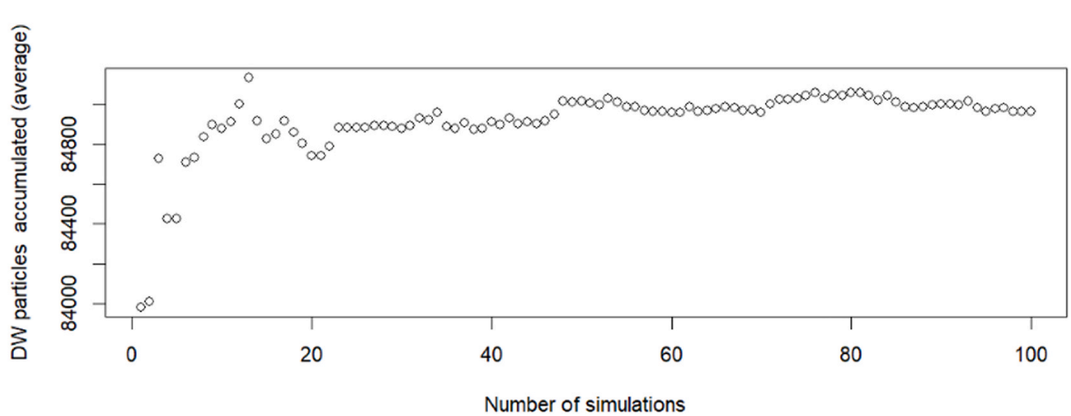
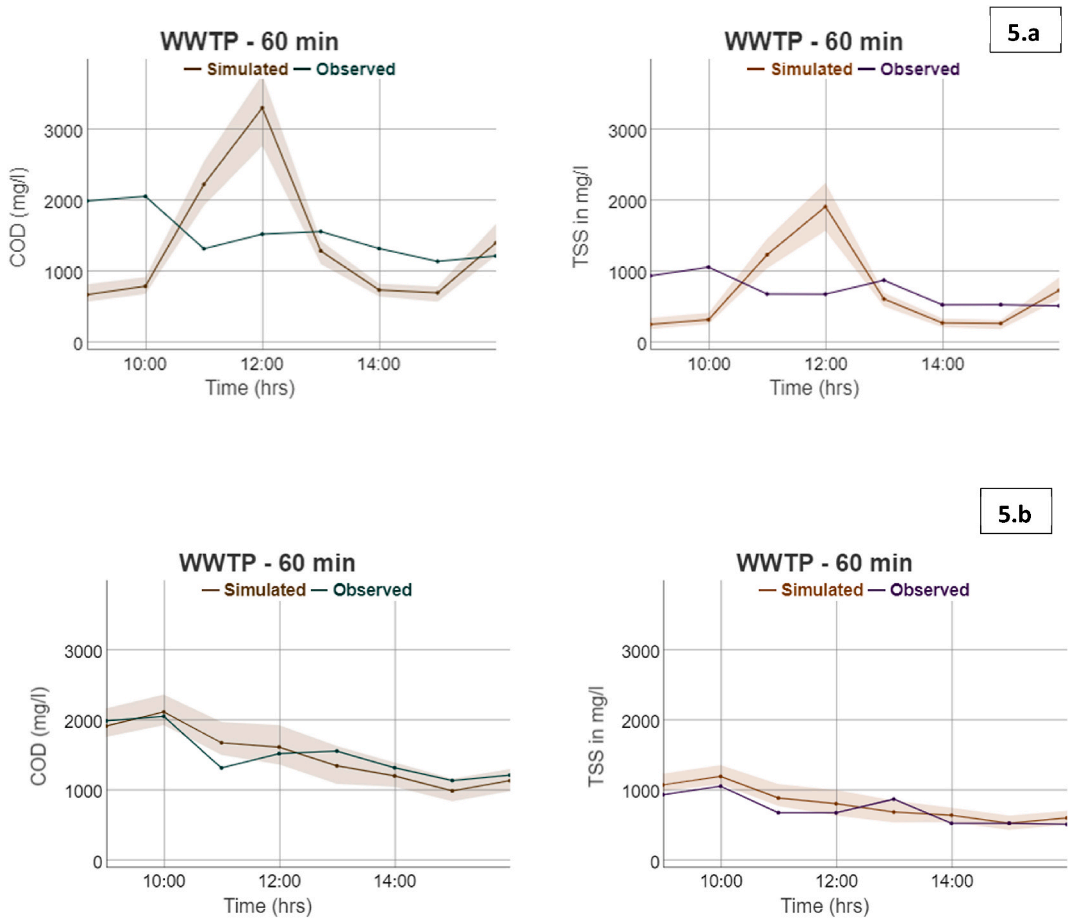


Fig. 4. Finding robustness of the DW simulation.



Plot	Min.	ID	Pollutant	r	p	N_Obs
5.a Not calibrated	60	WWTP	COD	-0.192	0.6480	8
5.a Not calibrated	60	WWTP	TSS	-0.174	0.6794	8
5.b Calibrated	60	WWTP	COD	0.888	0.003	8
5.b Calibrated	60	WWTP	TSS	0.857	0.007	8

Fig. 5. ABM before (5.a) and after (5.b) calibration. WWTP catchment with a temporal resolution of 60 min. The shaded areas show the DW variability determined from 50 simulation runs. r: Pearson correlation, p: P value, N_Obs: number of observed samples.

data. Fig. 5.a shows an example of the model before the calibration process to appreciate the improvement of the modeling process. The WWTP is also less susceptible to the strong variability at maintenance holes and the possibility of uncertain sewage connections in the proposed sewage.

2.3. Spatiotemporal domestic wastewater variability

Two pollutants (COD, TSS), five temporal resolutions (6, 12, 30, 60, and 180 min), and two spatial resolutions (WWTP catchment and sub-catchments 258-39) are evaluated. A 3-hr temporal resolution is selected to reflect the sampling interval of a continuous discharge that applies to DW from NOM-001-SEMARNAT. The 6-min resolution represents instantaneous samples, and resolutions between 6 min and 180 min (i.e., 12, 30, and 60 min) reveal details about the trends of the temporal resolutions. Spatially, we evaluate the catchment and sub-catchment levels. Fig. 2 shows that the catchment splits into two sub-catchment areas, including the location of the WWTP and the maintenance holes for the sub-catchments. The WWTP catchment covers approximately 0.35 $km^{2,}$ with 743 inhabitants. Each sub-catchment covers approximately 0.1 km^2, and sub-catchments 258 and 39 are estimated to have 393 and 215 inhabitants, respectively. We used the differences between the DW time series linked to the (sub)catchment areas and the number of inhabitants to describe the spatial DW variability (see Fig. 2).

We examined DW variability by quantifying and comparing the differences in pollutant loads (COD and TSS) between spatial levels and temporal resolutions. Temporally, the magnitude of DW variability was analyzed by comparing the number of events of maximum pollutant loads (peaks in time series) between temporal resolutions for a targeted (sub)catchment. The spatial magnitude of DW variability was analyzed by comparing the areas of variance (indicated by the shaded areas in Fig. 6) between the catchment and sub-catchment levels for a given temporal resolution. Spatiotemporally, the minimum and maximum load ranges were compared between resolutions. The variance areas illustrate the spatiotemporal DW variability quantitatively.

For the comparison, we also refer to spatiotemporal DW variability properties visible in the DW time series. One property is the temporal resolution trend, where the higher the temporal resolution, the higher the DW variability that can be expected, especially at decentralized

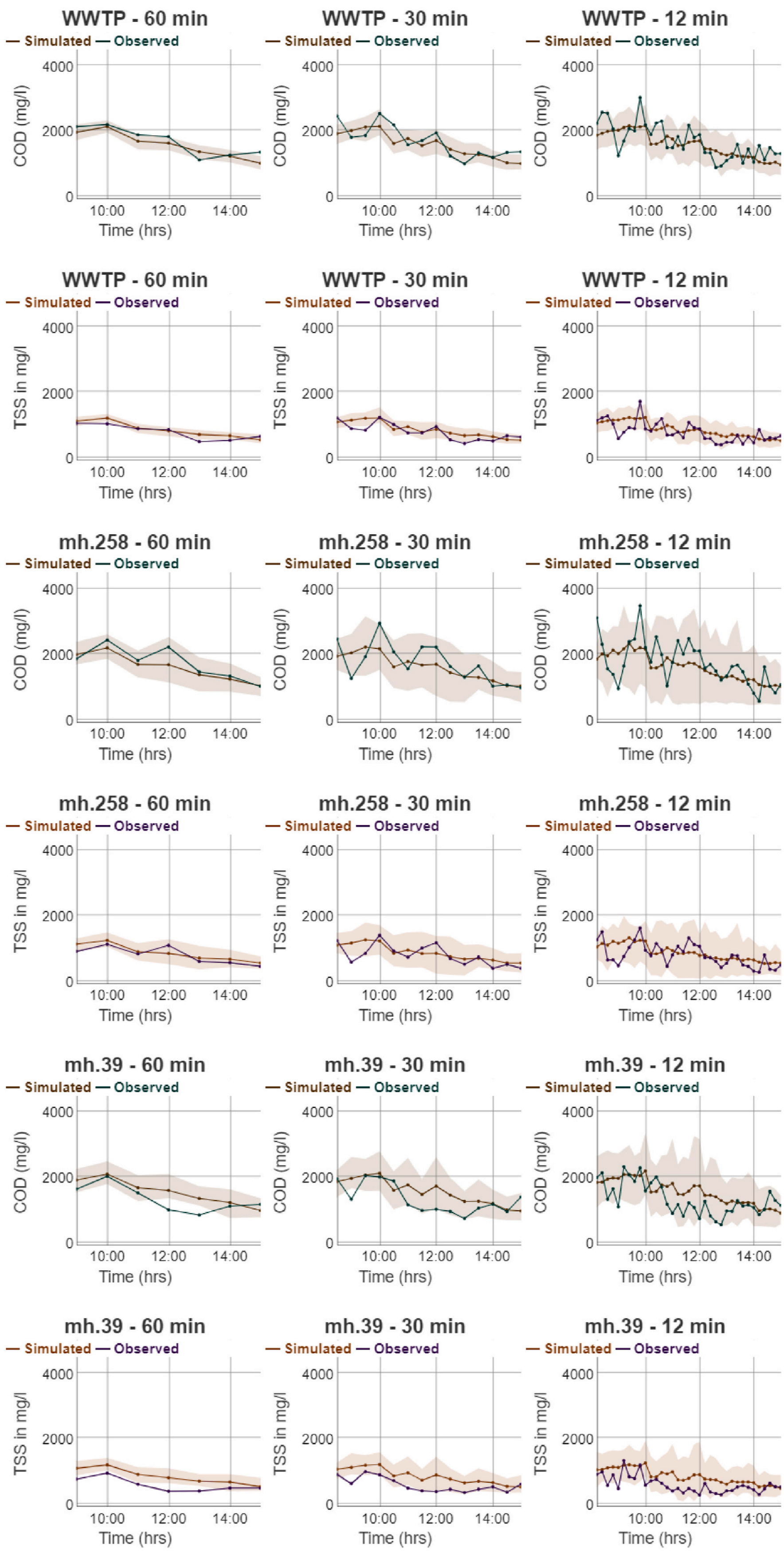


Fig. 6. Model results for DW variability. DW pollutant patterns were simulated at multiple spatial resolutions (sub-catchment areas 258 and 39 and the WWTP catchment area) and temporal resolutions (60, 30, and 12 min). The shaded areas show the variance coverage of 50 simulations.

treatment. A second property is the spatial resolution trend. The lower the spatial resolution is (i.e., from sub-catchment to catchment), the lower the DW variability is. The spatial resolution trend is also relevant to understanding the relationship between the DW load variability's relations to population and (sb)catchment size.

2.4. Spatiotemporal variability on treatment efficiency

Two inputs are required to estimate the treatment efficiency as the percentage of safely treated DW at WWTP: The DW modeling results of pollutant concentration (COD, TSS) and the discharge limits for instant values in the category of infiltration and irrigations from NOM-001, which values are TSS: 140 mg/l, COD: 210 mg/l (SEGOB, 2022a). Equation (1) defines a treatment plant efficiency as the difference between the pollutant's concentrations at inflow and outflow, considering a constant treatment performance (discharge limit) on a regular day—where *pstw* is the percentage of safely treated DW, and w_{out}, w_{in} are outflows and inflows of a pollutants (mg/l), respectively:

$$pstw = 100 - (100 w_{out} / w_{in}) \tag{1}$$

We demonstrate the implications of DW variability by quantifying the PSTW between the catchment and sub-catchment levels at multiple temporal resolutions. The following two assumptions were made: i) high temporal resolutions (e.g., 12 min) were assumed to represent instantaneous sampling to resemble NOM-001, and ii) by comparing sub-catchment and catchment levels, we can represent smaller and larger sanitation systems. PSTW was calculated for each catchment and sub-catchment area (see Fig. 2) for temporal resolutions of 12, 30, and 60 min, and the differences were calculated.

We introduce basic recommendations by incorporating the quantifiable effects of DW variability in the efficiency estimation at decentralized and ungauged WWTP. The relevance of the recommendations is provided in the context of current regulatory documentation linked to monitoring the sanitation situation, i.e., PSTW worldwide and NOOM-001 in Mexico (UNSTATS, 2022; SEGOB, 2022a). Temporally, we provide proper sampling duration and intervals to complement laboratory sampling limitations and efficiently capture DW variability for future regulation updates. The recommendation highlights the early morning DW variability as the most pronounced event (Von Sperling, 2007). Spatially, we highlight proper variables linked to spatial DW variability to make comparable efficiencies between multiple treatment systems.

3. Results

3.1. Understanding spatiotemporal variability

Quantifying the magnitude change between the spatial and temporal DW variability at multiple resolutions shows substantial differences. When comparing the temporal resolutions (12, 30, 60 min) in the sub-catchment 258 from Fig. 6, a low temporal resolution cannot show the number of peaks with significant maximum loads. The time series of the sub-catchment 258 show two and six significant loads at 60 and 12 min, respectively. The six maximum COD loads (12 min resolution) are between 3400 and 3900 mg/l with a high temporal resolution. In comparison, maximum COD loads are reduced from six to two with a low resolution (at 60 min) with maximum loads between 2500 and 3000 mg/l.

The simulated variances (indicated by the shaded areas in Fig. 6) exhibit substantial differences between the spatial levels of the WWTP catchment and sub-catchment 258. The variance area for COD loads is up to three times smaller for the catchment than for sub-catchment 258 at a temporal resolution of 60 min, which reflects the spatial variability trend (see Fig. 6). The WWTP catchment has a bigger sewage coverage resulting from the dilution of the combined flow from the multiple sewage sections, including DW collected from the sub-catchment 258 (see Fig. 2).

When comparing DW variability in space and time simultaneously (spatiotemporal), we find the biggest difference in the magnitude change of pollutant loads. Fig. 6 shows the variance areas where pollutant loads' minimum and maximum ranges change depending on the spatial level and temporal resolutions. The difference between a low and a high spatiotemporal resolution of the variance area can be up to five times less when comparing the WWTP catchment and sub-catchment 258. At the same time, the difference between the minimum and maximum load ranges can be around double with COD ranges that go from 800 to 2200 mg/l at the WWTP (60 min) compared to ranges from 400 to 3700 mg/l at sub-catchment 258 (12 min). The spatiotemporal resolution is especially relevant when comparing multiple WWTPs. Ensuring that spatiotemporal resolutions are similar allows a fair comparison between treatment systems**.**

Table 2 shows the evaluation metrics of the observed and simulated spatiotemporal DW variability. Table 2 allows comparisons of the COD and TTS pollutant loads for spatial catchment and sub-catchment resolutions at various temporal resolutions based on Pearson correlations (r) and p-values for a confidence interval of 95%. The most significant correlations are at 180- and 60-min temporal resolutions, which get the best correlation results above 0.75 with a decreasing trend when the temporal resolution increases. The best results are r values between 0.7 and 0.99 for the WWTP and maintenance holes 258 and 39, respectively, at a 180-min resolution. P-values also serve as reliability indicators at various resolutions. It is noted that the r significantly changes between the spatial and temporal resolutions considering (sub)catchment areas and the number of inhabitants. No significant change is noticed between the DW pollutants (COD and TSS).

3.2. Implications on treatment efficiency

The DW variability generates substantial differences when estimating treatment efficiency (*pstw*) based on the proposed scenario (see section 2.4). The variance (indicated by shaded bands in Fig. 6) highlights the variability of input values for equation (1) (w_{in}). As a result, the estimation of treatment efficiency given in percentages is significantly affected.

As seen in Table 3, temporal resolutions implicate that treatment efficiency can be inaccurate when ignoring the temporality of DW variability. Considering a high temporal resolution (12 min) of COD instantaneous sampling at the catchment, the treatment efficiency (*pstw*) can range between 70% and 95% (min.-max.). The instantaneous sampling can provide an (over)under-estimated treatment efficiency with an uncertainty of up to 25%, depending on the hour that a sample is taken. In contrast, when considering a lower temporal resolution (60 min) at the catchment, the uncertainty of treatment efficiency reduces. Possible COD *pstw* values only range between 80% and 90% (min.-max) with an uncertainty up to 10%. The treatment efficiency uncertainty can be reduced from 25% (12 min) to 10% (60 min), showing that temporal resolutions provide significant differences in treatment efficiency measurements.

The spatial component also impacts reporting treatment efficiency, mainly when considering spatiotemporal resolutions. The spatiotemporal implication relates to an inadequate comparison of treatment efficiencies between locations. For instance, assuming the efficiency estimation from the sub-catchment 258 (higher spatial resolution than the catchment) with a COD instantaneous sampling (high temporal resolution), the treatment efficiency can vary between 50% and 95% at the sub-catchment 258 and between 70% and 95% at the catchment. A 20% difference between the two spatial levels of catchment and sub-catchment misleads the comparison of treatment efficiency of locations with different spatial resolutions. It is required to provide information on the spatial resolutions regarding the (sub)catchment areas and inhabitants covered by WWTPs to compare multiple decentralized or centralized WWTPs objectively.

Table 2
Metrics to evaluate simulated DW variability. "r" refers to Person correlations, and "p" for p values.

	Temporal resolution (minutes)	r (COD)	r (TSS)	p (COD)	p (TSS)	Observations
Catchment spatial resolution (WWTP: 0.35 km^2 743 inhabitants)	180	0.942	0.942	0.2184	0.2171	3
	60	0.904	0.876	0.0053	0.0097	7
	30	0.769	0.755	0.0013	0.0018	14
	12	0.684	0.625	0.0000	0.0001	35
	6	0.646	0.595	0.0000	0.0000	70
Sub-catchment spatial resolution (mh.258: 0.1 km^2 393 inhabitants)	180	0.988	0.994	0.0970	0.0702	3
	60	0.909	0.838	0.0046	0.0186	7
	30	0.698	0.638	0.0055	0.0141	14
	12	0.592	0.560	0.0002	0.0005	35
	6	0.538	0.476	0.0000	0.0000	70
Sub-catchment spatial resolution (mh.39: 0.1 km^2 215 inhabitants)	180	0.921	0.868	0.2546	0.3308	3
	60	0.635	0.866	0.0146	0.0116	14
	30	0.635	0.738	0.0146	0.0026	14
	12	0.558	0.607	0.0005	0.0001	35
	6	0.469	0.485	0.0000	0.0000	70

Table 3
Variability of treatment efficiency linked to spatiotemporal resolutions. Treatment efficiency columns (*pstw*) are given in percentages.

	Temporal resolution (minutes)	PSTW Min (COD, %)	PSTW Max (COD, %)	PSTW Uncertainty (COD, %)
Catchment spatial resolution (WWTP: 0.35 km^2 743 inhabitants)	60	80	90	10
	12	70	95	25
Sub-catchment spatial resolution (mh.258: 0.1 km^2 393 inhabitants)	12	50	95	45

4. Discussion

4.1. Modeling spatiotemporal variability

The modeling outcomes provide comprehensive insights for explaining the DW variability from their source, the human polluters. By its detailed spatiotemporal design, we targeted to analyze where, when, and how much the DW variability is expected to be higher or lower. Temporally, Teerlink et al. (2012) mention that a high sampling frequency is required at small spatial scales to detect significant wastewater pollutant concentrations. Our results confirm that significant pollutant concentrations are better detected at high temporal resolutions. Spatially, Penn et al. (2017) highlight that a low population (linked to high spatial resolutions from small catchment areas) relates to intermittent DW variations. Friedler and Butler (1996) affirm that DW at the source (the high spatial resolution of households) is highly variable. Our study also shows that as the spatial resolution increases (from catchment to sub-catchment with lower population), the DW pollutants concentrations will provide intermittent and strong variations.

We also find that the ABM performance is comparable with the current state of the art, where the reported modeling performance has been moderate in the last decade (Jia et al., 2021). As Jia et al. (2021) mentioned, studies on modeling DW mentioning spatial resolutions are uncommon. Our findings contradict a note from Jia et al. (2021), who suggested that once a model is calibrated, it can be applied to multiple spatial resolutions. We advise evaluating the methodological principles for modeling DW variability to understand which spatial characteristics a modeling method can cover, specifically when the method must be replicated in real-life conditions in low and middle-income countries. The DW variability is linked to sewage catchment features that significantly affect DW modeling results when having a significant presence, such as infiltration in the sewage (from drinking water, spring water, or rivers), DW leaking (from old sewage or broken pipes), and water shortages (reflected as moderate water appliance usage). We advise referring to the population size and catchment area to compare DW variability in future literature concerning spatial resolutions.

The ABM has non-trivial points during the modeling process that should be considered. These include that no mobility from/to areas outside the locality is simulated. The model focuses on domestic wastewater; not all wastewater from economic activities is included. However, as the model uses general behaviors, biological rules, and reference input data, the model can be replicated in other localities. The current model only considers the COD and TSS DW components. Other DW components are not evaluated, and the current modeling approach does not apply to components whose units are not measured in mg/l as ph. Studying DW variability across several weeks requires more data, which is beyond the scope of the current study. For instance, studying for several weeks allows one to analyze common water shortages that are common during the dry season. The potential effects of rainfall in this model are not developed as the case study targets a separate sewer network between rain and domestic wastewater. Fieldwork for studying DW variability at multiple spatiotemporal resolutions represents a challenge. Data scarcity and fieldwork challenges can be the strongest reasons for rarely studying spatiotemporal DW dynamics. As the model is based on hourly probabilistic distributions, the model usage for decision-making is not recommended with resolutions lower than 1 h. However, the model usage can be recommended to provide a better understanding of DW variability at any resolution (including lower than 1 h), i.e., for understanding the complexity of DW patterns and variability. We note that designing the model using higher probabilistic distributions of minutes substantially increases the computational cost, where high-performance computing is out of the scope of this study.

4.2. Recommendations about treatment efficiency

Our study allows us to propose two general recommendations for sampling and reporting treatment efficiency at decentralized and ungauged WWTP: i) For covering the temporal trend, sensors are recommended for higher-frequency sampling. Sensors should be regulated to guarantee correct use and properly complement current limitations with laboratory practices. A few instantaneous samples are not enough to represent the realistic temporal trend of DW pollutant loads. Especially for decentralized WWTP (smaller population and catchment areas), instantaneous values should be avoided when treatment efficiency must be reported. Instead, DW sampling at 6-min intervals over 3 h, reflecting the time range of major use of water fixtures and appliances (i.e., 7:00 to 10:00 a.m.), should be considered. The starting sampling time should be adjusted based on intensive water use and DW flow travel time from households to WWTP. This basic recommendation ensures that the significant DW peaks (produced by early morning population activities) are captured during the sampling, where a WWTP should demonstrate that high demands of water appliance usage do not compromise the WWTP efficiency.

ii) To cover spatiotemporal trends, the spatial and temporal resolutions of the DW sampling must be reported. This recommendation addresses standardizing the reporting of treatment efficiency to allow interpretation and comparison between multiple treatment systems. Temporally, the duration and sampled times and intervals must be documented. The spatial parameters of the sewage catchment area and the number of inhabitants must be recorded in parallel with sampling times and intervals. Such parameters are sensitive indicators of whether there is substantial variability in the estimated treatment efficiency. Reporting the number of inhabitants must consider, when possible, how significant population mobility is (entering and leaving the catchment). For instance, studying, working, and daily tourism can significantly change the population in the sewage catchment. It is also necessary to report the sewage conditions and maintenance periods of WWTPs, as infiltration from clean water or DW leaking from sewage affects pollutant concentrations.

This study contributes to a better understanding of the effect of DW variability on estimated WWTP efficiency. We describe the relevance of studying DW variations at different spatial and temporal resolutions. Highlighting DW variability contributes to addressing regulation issues linked to monitoring and improving understanding of the implications of reporting WWTP efficiency. Usually, in decentralized sanitation management, one or a few laboratory DW samples are collected at the WWTP catchment, which complies with current regulations. However, detailed simulation is needed to obtain insights into DW dynamics at different spatial and temporal resolutions. We detect inaccuracies of up to 25% in estimating WWTP efficiencies. We therefore suggest a sampling frequency of 6 min over 3 h in the morning to report WWTP efficiency, to account for the uncertainty introduced by DW variability. Regulations should be upgraded to report treatment efficiency using modern technologies like DW spectrometer sensors.

Studies on DW sampling for reporting treatment efficiency require more attention, and this research can be used as a starting point. Mexico is not alone in having regulations, such as NOM-001-SEMARNAT (SEGOB, 2022a), with DW sampling specifications that lack consideration of the effects of spatiotemporal DW variability. The European Union law (2014) has also documented regulations linked to sampling WWTP that allow sampling frequencies between 12 and 365 per year, omitting the relevance of DW spatiotemporal variability. This research shows that studying multiple resolutions is useful and that reporting spatiotemporal resolutions must be required, specifically for decentralized sanitation. A report by the European Commission (2019) on an evaluation of wastewater treatment mentions the promotion of research on the fit-for-purpose sampling frequency to assess whether WWTPs comply with regulations where our findings show a starting point.

Monitoring DW treatment efficiency is not affected by DW variability alone. Reporting WWTP efficiency requires considering the feasibility of continuous data collection, social aspects, and the local environment, which can determine how regulations should be updated. For instance, planning the logistics for sampling DW can require significant time coordination between technical teams, local WWTP design and management experts, municipal authorities, and citizens. Highlighting implications and recommendations for measuring WWTP efficiencies alone does not answer how to improve regulations to overcome the challenges of continuously reporting WWTP efficiencies.

5. Conclusions

We successfully developed, implemented, and evaluated an ABM to simulate DW variability reproduced at various spatial and temporal resolutions. The model provided correlations above 0.7 between COD and TSS at distinct locations and performed better at temporal resolutions of 30, 60, and 180 min. For high resolutions (less than 1 h), we recommend using the model to understand DW variability better and not for decision-making, for which high-performance computation can provide better correlations. The results of this study help to explain DW dynamics at multiple time resolutions and two spatial resolutions (catchment and sub-catchment). DW variability increases at higher temporal and spatial resolutions. Such knowledge is key for examining practices of instantaneous sampling and addressing the limitations of sampling with sporadic measurements when reporting treatment efficiency, which are common practices in Mexico and worldwide. Even with limited human and technical resources, it was possible to collect valuable data that support DW modeling for low- and middle-income rural and peri-urban localities. This is a relevant point for decentralized sanitation systems, where efforts should be increased to report treatment status. Our work highlights the relevance of understanding spatiotemporal DW variability at multiple resolutions. Without a proper understanding of DW variability, treatment efficiency can be overestimated or underestimated by up to 25%. We recommend sensor measurements at 6 min intervals over 3 h during heavy drinking water demand in the morning. Finally, data on sewage catchment areas, population sizes, sample times, and intervals are required to objectively compare WWTPs efficiencies and overcome misestimations due to DW variability. Because the current model is not easily transferable to other localities, our next research efforts will be focused on generating a model that can be easier to implement in other localities.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

This research implements open science practices. The research data and materials are preserved in the following repository: https://doi.org/10.5281/zenodo.10242567.

Acknowledgments

The authors wish to express their gratitude to the following for the valuable support and talks we received during the development and implementation of this study, without which our findings would not have been possible: Carlos Pailles, Ana Velasco, Rodrigo Tapia-McClung, Araceli Chávez, Diana Ramos, Daniela Goméz, José Luis, Lucía Guardamino, Paulo Raposo, Rosa Aguilar, Gustavo Garcia, Yuan Chen, Natalia Volkow, the anonymous facilitators from Mexico City, the citizens of Tepeji del Río, the INEGI Department of Microdata, and the scholarship funder, CONACYT-Alianza FiiDEM (Reference: 2019-000005-01EXTF-0027).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jenvman.2023.119680.

References

Almeida, M.C., Butler, D., Friedler, E., 1999. At-source domestic wastewater quality. Urban Water 1 (1), 49–55.
Atinkpahoun, C.N.H., Le, N.D., Pontvianne, S., Poirot, H., Leclerc, J.P., Pons, M.N., Soclo, H.H., 2018. Population mobility and urban wastewater dynamics. Sci. Total Environ. 622 (623), 1431–1437.
Auchincloss, A.H., Martin, L., Garcia, T., 2015. Brief Introductory Guide to Agent-Based Modeling and an Illustration from Urban Health Research, p. 14.

Borgonovo, E., Pangallo, M., Rivkin, J., Rizzo, L., Siggelkow, N., 2022. Sensitivity analysis of agent-based models: a new protocol. Comput. Math. Organ. Theor. 28 (1), 52–94.

Di Cicco, M.R., Masiello, A., Spagnuolo, A., Vetromile, C., Borea, L., Giannella, G., Iovinella, M., Lubritto, C., 2021. Real-time monitoring and static data analysis to assess energetic and environmental performances in the wastewater sector: a case study. Energies 14 (21).

CONAGUA, 2019. Volumen 20 Alcantarillado Sanitario [Online]. Available from: http://mapasconagua.net/libros/SGAPDS-1-15-Libro25.pdf.

Dubois, V., Falipou, E., Boutin, C., 2022. Quantification and qualification of the urban domestic pollution discharged per household and per resident. Water Sci. Technol. 85 (5), 1484–1499.

Duester, H., 2021. SRTM-downloader: Plugin for Download of SRTM Tiles from NASA.

European Commission, 2019. Evaluation of the urban waste water treatment directive [Online]. Brussels. Available from: https://ec.europa.eu/info/publications/evaluation-urban-waste-water-treatment-directive_en.

European Union law, 2014. EUR-lex - 01991L0271-20140101 - EN - EUR-Lex: Council Directive of 21 May 1991 Concerning Urban Waste Water Treatment (91/271/EEC), p. 40, 30 November 2022]. Available from: https://eur-lex.europa.eu/eli/dir/1991/271/2014-01-01.

Friedler, E., Butler, D., 1996. Quantifying the inherent uncertainty in the quantity and quality of domestic wastewater. Water Sci. Technol. 33 (2), 65–75.

Gilbert, N., 2011. Designing and developing agent-based models. In: SAGE Research Methods: Agent-Based Models. SAGE Publications, Inc., pp. 47–68

Grimm, V., Ayllón, D., 2020. The ODD protocol for agent-based and other simulation models: a second update to improve clarity, replication, and structural realism. Supplementary file S2 : summary ODD. J. Artif. Soc. Soc. Simulat. 23 (2), 1–8.

Heckert, A., 2003. National institute of standards and technology (NIST): statistical engineering division. Dataplot, CUMULATIVE AVERAGE (LET). [Online]. Available from: https://www.itl.nist.gov/div898/software/dataplot/refman2/auxillar/cumumean.htm.

Henze, M., Comeau, Y., 2008. 3. Wastewater Characterization *in: Biological Wastewater Treatment: Principles Modelling and Design*. LON: IWA.

INEGI-DENUE, 2017. Directorio Estadístico Nacional de Unidades Económicas (DENUE) 2017. Available from: http://www.beta.inegi.org.mx/app/descarga/?ti=6.

INEGI, 2020. Censo de Población y Vivienda 2020. *Censos y conteos*. [Online]. [Accessed 7 October 2022]. Available from: https://www.inegi.org.mx/programas/ccpv/2020/.

Jia, Y., Zheng, F., Maier, H.R., Ostfeld, A., Creaco, E., Savic, D., Langeveld, J., Kapelan, Z., 2021. Water quality modeling in sewer networks: review and future research directions. Water Res. 202, 117419.

De Keyser, W., Gevaert, V., Verdonck, F., De Baets, B., Benedetti, L., 2010. An emission time series generator for pollutant release modelling in urban areas. Environ. Model. Software 25 (4), 554–561.

Lovelace, R., Dumont, M., 2018. Spatial Microsimulation with R [Online] Website Ed. CRC Press. Available from: https://spatial-microsim-book.robinlovelace.net/.

McLane, A.J., Semeniuk, C., McDermid, G.J., Marceau, D.J., 2011. The role of agent-based models in wildlife ecology and management. Ecol. Model. 222 (8), 1544–1556.

Mesdaghinia, A., Nasseri, S., Mahvi, A.H., Tashauoei, H.R., Hadi, M., 2015. The estimation of per capita loadings of domestic wastewater in Tehran. J Environ. Health Sci. Eng. 13 (1).

OECD, 2010. Financing Water Resources Management in Mexico [Online] [Accessed 24 January 2023]. Available from: http://www.conagua.gob.mx/CONAGUA07/Contenido/Documentos/OECD.pdf.

Penn, R., Schütze, M., Gorfine, M., Friedler, E., 2017. Simulation method for stochastic generation of domestic wastewater discharges and the effect of greywater reuse on gross solid transport. Urban Water J. 14 (8), 846–852.

Rodríguez, J.P., McIntyre, N., Díaz-Granados, M., Achleitner, S., Hochedlinger, M., Maksimović, Č., 2013a. Generating time-series of dry weather loads to sewers. Environ. Model. Software 43, 133–143.

Rodríguez, J.P., McIntyre, N., Díaz-Granados, M., Quijano, J.P., Maksimović, Č., 2013b. Monitoring and modelling to support wastewater system management in developing mega-cities. Sci. Total Environ. 445 (446), 79–93.

Rose, C., Parker, A., Jefferson, B., Cartmell, E., 2015. The characterization of feces and urine: a review of the literature to inform advanced treatment technology. Crit. Rev. Environ. Sci. Technol. 45 (17), 1827–1879 s::can 2022. spectro::lyser UV • s::can. [Accessed 2 January 2023]. Available from: https://www.s-can.at/en/product/spectrolyser-uv/.

Sato, T., Qadir, M., Yamamoto, S., Endo, T., Zahoor, A., 2013. Global, regional, and country level need for data on wastewater generation, treatment, and use. Agric. Water Manag. 130, 1–13.

Schober, P., Schwarte, L.A., 2018. Correlation coefficients: appropriate use and interpretation. Anesth. Analg. 126 (5), 1763–1768.

SEGOB, 2022a. NORMA Oficial Mexicana NOM-001-SEMARNAT-2021, 11 March 2022]. Available from:. DIARIO OFICIAL DE LA FEDERACIÓN, pp. 1–20 https://www.dof.gob.mx/nota_detalle_popup.php?codigo=5645374, 1/20.

SEGOB, 2022b. REGLAS de Operación para el Programa de Agua Potable, Drenaje y Tratamiento a cargo de la Comisión Nacional del Agua, aplicables a partir de 2022. DIARIO OFICIAL DE LA FEDERACIÓN, pp. 1–48. Available from: https://www.dof.gob.mx/nota_detalle.php?codigo=5639902&fecha=31/12/2021#gsc.tab=0.

Sheppard, C., Railsback, S., Kelter, J., 2022. NetLogo 6.3.0 User Manual: Time Extension [Accessed 30 January 2023]. Available from: https://ccl.northwestern.edu/netlogo/docs/time.html.

Von Sperling, M., 2007. Wastewater characteristics, treatment and disposal. In: Biological Wastewater Treatment Series, ume 1. IWA Publishing, London iwapublishing.com.

Tabla-Vázquez, Cynthia, G., Chávez-Mejía, A.C., Ledesma, M.T.O., Ramírez-Zamora, R.M., Tabla-Vázquez, C.G., Chávez-Mejía, A.C., Ledesma, M T Orta, Ramírez-Zamora, R.M., 2020. Wastewater treatment in Mexico. World Water Resour. 6, 133–155.

Teerlink, J., Hering, A.S., Higgins, C.P., Drewes, J.E., 2012. Variability of trace organic chemical concentrations in raw wastewater at three distinct sewershed scales. Water Res. 3261–3271.

The World Bank, 2020. People Using Safely Managed Sanitation Services (% of Population) | Data. WHO/UNICEF Joint Monitoring Programme (JMP) for Water Supply. *Sanitation and Hygiene (washdata.org)*. [Online]. [Accessed 19 December 2022]. Available from: https://data.worldbank.org/indicator/SH.STA.SMSS.ZS?end=2020&start=2020&type=shaded&view=map.

Ujang, Z., Mogens, H., 2006. Municipal Wastewater Management in Developing Countries [Online]. IWA Publishing. Available from: https://iwaponline.com/ebooks/book/20/Municipal-Wastewater-Management-in-Developing?redirectedFrom=PDF.

UN-Water, 2020. Indicator | SDG 6 data. In: Progress on Wastewater Treatment (SDG Target 6.3). [Online], 25 July 2022]. Available from: https://sdg6data.org/indicator/6.3.1.

UN Statistics Division, 2020. Infographics: the Sustainable Development Goals Report 2022. SDGs Infographics.

UNSTATS, 2022. SDG Indicators — Metadata Repository. *Metadata: Indicator 6.3.1: Proportion of domestic and industrial wastewater flows safely treated. United Nations Statistics Division*. [Online]. [Accessed 23 November 2022]. Available from: https://unstats.un.org/sdgs/metadata/?Text=&Goal=6&Target=6.3.

WHO, W.H.O., UN-Habitat, U.N.H.S.P. and UNSD, U.N.S.D., 2020. Indicator 6.3.1: Proportion of Domestic and Industrial Wastewater Flows Safely Treated. *SDG indicator metadata (Harmonized metadata template-format version 1.0)* [Online]. [Accessed 6 October 2022]. Available from: https://unstats.un.org/sdgs/metadata/files/Metadata-06-03-01.pdf.

Zechman, E.M., 2011. Agent-based modeling to simulate contamination events and evaluate threat management strategies in water distribution systems. Risk Anal. 31 (5), 758–772.

Zhou, P., Li, Z., Snowling, S., Baetz, B.W., Na, D., Boyd, G., 2019. A random forest model for inflow prediction at wastewater treatment plants. Stoch. Environ. Res. Risk Assess. 33 (10), 1781–1792.