Proceedings of the 2025 Winter Simulation Conference
E. Azar, A. Djanatliev, A. Harper, C. Kogler, V. Ramamohan, A. Anagnostou, and S. J. E. Taylor, eds.

A DIGITAL TWIN OF WATER NETWORK FOR EXPLORING SUSTAINABLE WATER MANAGEMENT STRATEGIES

Souvik Barat[1], Abhishek Yadav[1], Vinay Kulkarni[1], Gurudas Nulkar[2], Soomrit Chattopadhyay,[2] and Ashwini Keskar[3]

[1]Tata Consultancy Services Research, Pune, INDIA
[2]Gokhale Institute of Politics and Economics, Pune, INDIA
[3]Pune Knowledge Cluster, Pune, INDIA

ABSTRACT

Efficient water management is an increasingly critical challenge for policymakers tasked with ensuring reliable water availability for agriculture, industry and domestic use while mitigating flood risks during monsoon seasons. This challenge is especially pronounced in regions where water networks rely primarily on rain-fed systems. Managing such water ecosystem is complex due to inherent constraints in water source, storage and flow, environmental uncertainties like variable rainfall and evaporation, and increasing need for urbanization, industrial expansion and equity on interstate water sharing. In this study, we present a stock-and-flow based simulatable digital twin designed to accurately represent the dynamics of a rain-dependent water network comprising dams, rivers and associated environmental and usage factors. The model supports scenario-based simulation and the evaluation of mitigation policies to enable evidence-based decision-making. We demonstrate the usefulness of our approach using a real water body network from western India that covers more than 300 km heterogeneous landscape.

1 INTRODUCTION

Water management is a critical and increasingly complex challenge faced by policymakers around the world as they need to ensure sustainable water distribution for agriculture, industry and household use throughout the year under varying and evolving weather and geopolitical situations (Mishra et al. 2021) . They must also adapt control measures to prevent flooding and water crisis during adverse weather events and maintain the health of water ecosystems over the years. This challenge is particularly critical in regions that rely on river ecosystems with rain-dependent water sources. Water situations in these regions are extremely sensitive to factors such as environmental variability (e.g., rainfall and heat), differing water consumption patterns (e.g., agricultural, industrial and domestic) and the intricate policies governing water release and distribution. With the rising water demand due to urbanization and industrialization, and the prominence of climate-induced fluctuations in natural water resources like rainfall, the need for efficient water use and distribution is becoming critical for ensuring long-term sustainability. If water is not managed properly, vulnerable regions may experience seasonal water shortages that disrupt domestic and industrial supplies while damaging agricultural production. Conversely, inadequate or overly conservative water release strategies can lead to upstream flooding which typically results in forced release in bursts thus culminating in downstream flooding.

A range of modeling approaches with unique offerings and limitations has been applied to model and simulate river ecosystems. For example, coarse-grained System Dynamics (SD) models (Kotir et al. 2016; Ahmad and Simonovic 2000) are used to capture dynamics and feedback loops in water systems. They mostly focus on specific phenomena like flooding but are inadequate for representing spatiotemporal climate and geographic heterogeneities and capturing environmental uncertainties. Agent-Based Models (ABM) (Ghoreishi et al. 2021), in contrast, enable fine-grained modeling and are capable of considering heterogeneous elements. However, they require excessive data points, which are difficult to gather from real systems, and are also computationally challenging to scale for large water ecosystems. Moreover, they are often inadequate for representing hydrological characteristics as compared to traditional hydrological models. Traditional equation-based hydrological models (Devia et al. 2015)) offer sophisticated ways to model water storage and flow using mass-balance equations, but they often focus on limited environmental and geographical aspects, and are inadequate for representing inherent heterogeneity and uncertainties. Hybrid modeling techniques like WEAP21 (Yates et al. 2005) attempt an integrated approach, but their capabilities are constrained by the underlying formal technique, such as linear programming, especially when scaling to a large number of parameters and capturing uncertainties. On the other hand, early reports on applying digital twins in this context are promising. However, at this stage, most applications are either conceptual (Pal et al. 2025) or limited to specific aspects and contexts, such as dams (Yang et al. 2024). Decision-makers need a pragmatic and effective technology to rigorously comprehend complex river ecosystems by making use of realistically available data in a scalable manner.

We propose a stock-and-flow (SnF) based simulatable digital twin for modeling water networks, where the digital twin serves as a computational replica of a river ecosystem for in-silico what-if analysis leading to informed decision-making and strategy evaluation. The key contribution of our work is a set of modeling abstractions that represent primary building blocks of a river ecosystem, namely: catchment area, reservoir, river and confluences. These abstractions are capable of capturing environmental factors (e.g., rainfall, humidity and temperature) and geographical heterogeneities (e.g., vegetation and soil characteristics) in addition to flow dynamics resulting from inflows, outflows, spillage policy-driven controlled releases and water extractions for agriculture, industry, and household use. We further present an approach to combine these abstractions to first form a digital replica (i.e., digital model) of the river ecosystem, and propose a method to configure, validate and synchronize it with the real system to form a functional digital twin.

The simulatable nature of the digital twin aids in understanding non-linear effects and emergent behaviors resulting from the cascading impacts of diverse factors, such as climate variability, geographical features, water demand and fulfillment strategies, and existing policies. Our digital twin simulation also provides quantitative insights into the effects of inherent uncertainties including unpredictable environmental changes like sudden rainfall shortages or unexpected heavy rainfall events. Furthermore, it supports the assessment of both tactical and strategic water management strategies aimed at ensuring sustainable water supply while reducing risks from environmental disruptions such as flooding or drought-like conditions. It can also help mitigate the negative impacts of broader strategic decisions, such as urbanization, or arriving at better planning for vegetation.

The remainder of this paper is organized as follows: Section 2 discusses the inherent complexities that make water ecosystems difficult to comprehend and decision-making a challenging endeavor. Section 3 presents our digital twin approach along with the modeling abstractions used to construct a water ecosystem digital twin. Section 4 provides an illustrative case study and demonstrates the efficacy of the proposed approach by simulating potential disruptions and mitigation strategies. This section uses real-world scenarios from the western part of India. Finally, Section 5 concludes the paper with future possibilities.

2 SYSTEM COMPLEXITIES

Understanding the dynamics of river ecosystems, evaluating inherent risks like flooding and ensuring water demands in an uncertain environment are perpetual challenges for decision-makers. The core challenge lies in the fact that a river ecosystem is a complex system of systems operating in a constrained geological and physical settings with limited storage and flow capacities, span across a large heterogeneous landscape and also subject to spatiotemporal environmental variability and uncertainty.

From a systems theory perspective, we visualize a river ecosystem along four interoperable dimensions: the water body network, environmental factors, water usage, and policies and interventions. Figure 1 presents a Causal Loop Diagram (CLD) that highlights key factors from all four dimensions and their interactions. In this context, the water body network represents the physical components and hydrological

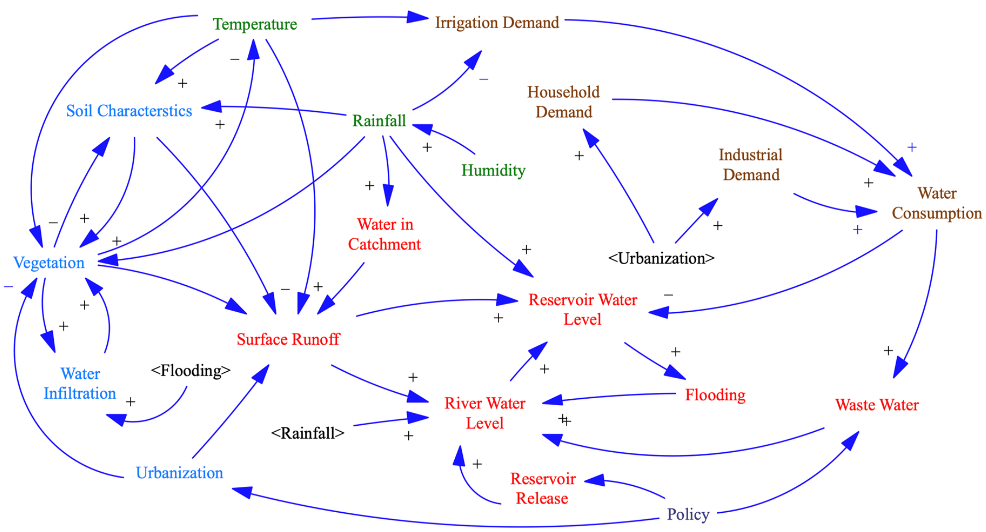


Figure 1: A Causal Loop Diagram describing influencing factors and key dependencies. **Red**: water body network, **Blue**: environmental spatial factors, **Green**: climatic factors, **Brown**: human usage and demand.

flow dynamics; environmental factors introduce climatic and geographical characteristics along with their associated uncertainties and temporal evolution; the usage dimension captures demand–fulfillment dynamics across domestic, agricultural and industrial sectors; and interventions serve as control mechanisms to achieve system-wide goals, such as meeting water demands while minimizing stressors like floods and droughts. The remainder of this section examines each dimension, their systemic complexities, and their inter-dependencies.

2.1 Water Body Network

This dimension, highlighted using red-colored text in the CLD depicted in Figure 1, represents physical aspects and elements, which are chiefly governed by hydraulic laws. The key complexity around this dimension is how physical elements with known dynamics get influenced by factors from other dimensions causing non-linear, emergent and unanticipated behavior. For example, catchment areas play a crucial role in accumulating water from rainfall and generating river flow in ecosystems that rely heavily on rain-fed rivers. While water accumulation is a physical phenomenon, the conversion of rainfall into runoff and subsequent river flows is largely influenced by multiple factors, such as geographical characteristics (e.g., soil composition, vegetation, urbanization and water infiltration capacity) and climate conditions (e.g., temperature and humidity) in the area, as shown in the CLD. Similarly, the natural replenishment and depletion of water from rivers to reservoirs (or dams), and subsequently from dams to rivers, are governed by hydraulic properties (and can be represented using mass-balance equations). However, these flows can be heavily influenced by the quantity of water fetched from reservoirs and dams for human consumption, climate conditions and the properties of the local land.

2.2 Environmental Factors

We categorize environmental factors into two broad subcategories: area-specific geographical characteristics and climatic factors. The geographical characteristics, which include soil composition, vegetation cover, water infiltration capacity and the degree of urbanization (shown using blue-colored text in the CLD Figure 1) describe the spatial properties of the land while exhibiting area-specific heterogeneity. In contrast, climatic factors, such as rainfall, temperature and humidity (shown in green text in the CLD Figure 1), describe seasonality and temporal uncertainty. Each subcategory of environmental factors has its own internal dynamics, and these factors also influence each other within and across dimensions. For example, soil characteristics are affected by vegetation cover, subsurface dynamics (e.g., water infiltration) and urbanization. Water infiltration is cyclically influenced by vegetation and the occurrence of flooding events in the area. Urbanization, in particular, can significantly alter the natural behavior of catchment areas by increasing the amount of impervious surfaces (e.g., roads, pavements, buildings). This reduces infiltration and increases surface runoff leading to sudden elevated water levels in reservoirs and rivers. Additionally, changes in soil composition and reductions in vegetation cover diminish the ability to retain water in an area. On one hand, vegetation helps water absorption and reduces runoff; on the other hand, urban expansion typically limits groundwater recharge.

While climate factors, such as temperature, rainfall intensity and humidity, have their own dynamics and natural influences with each other, they interact with other elements in both positive and negative ways within environmental factors. For instance, higher rainfall may increase surface water, thus contribute to excessive runoff. It can cause flooding if the land is not capable of adequate infiltration. The temperature and humidity of an area influence evaporation rates, which in turn alter soil moisture levels and can impact the local climate further. Across these two domains, we observe strong feedback loops. For example, reduced vegetation due to drought (a climatic event) can degrade soil quality (a geographical factor), further reducing the ability to absorb rainfall.

2.3 Usage

Water consumption in terms of household, irrigation and industrial usage plays a crucial role in depleting water stocks from reservoirs. While the demands for different types of water consumption have their own ecosystems and are governed by their respective dynamics, they, at a broader level, are influenced by climate, urbanization and policies. The elements are shown using brown text in CLD (Figure 1). Broadly, water demand tends to follow seasonal patterns, which are chiefly influenced by weather conditions. For example, temperature levels directly affect household consumption and temperature along with rainfall affect irrigation needs. During summer months, both domestic and agricultural water demands increase to compensate for reduced natural water availability. Irrigation, in particular, places significant pressure on river water levels, contributing to water scarcity and often resulting in reduced river flow. Beyond seasonal variation, water demand is also steadily increasing over time due to climate change, shifts in vegetation patterns and continued urbanization. In the industrial sector, urban expansion frequently leads to higher water consumption, which impacts both water stock and flow. When water levels drop to critically low thresholds, these changes may even trigger a significant change in the hydraulic dynamics of the ecosystem resulting in emerging behavior.

2.4 Interventions

This dimension captures human-driven actions or interventions that represent both policy-related and infrastructural changes within the water body network, interactions between various elements and usage factors. We broadly categorize these actions into two types: operational controls and strategic changes. **Operational controls**, such as modifying water release schedules, are considered to address sudden changes in water demand or weather conditions. **Strategic changes** include the construction of new reservoirs or the reconfiguration of existing water networks using canals. Such decisions are inherently complex and require a comprehensive understanding of physical constraints, interactions with the environment and long-term uncertainties. As shown in the CLD, here we consider a limited number of interventions in this context for illustration. They include water release policies, infrastructure development (as part of urbanization), and wastewater management strategies. These interventions primarily aim to manage water availability, mitigate risks of flooding and water scarcity, and balance competing demands across different usages.

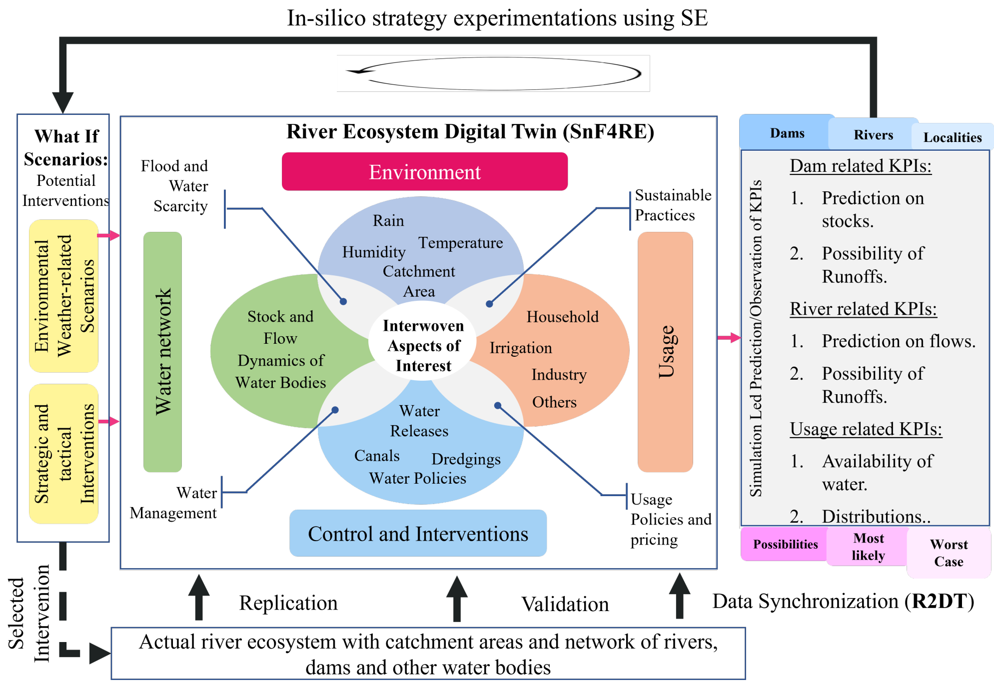


Figure 2: A river ecosystem digital twin framework.

3 APPROACH

We propose a simulatable stock-and-flow-based digital twin to facilitate experimental analysis, enabling stakeholders to comprehend inherent complexities, assess the impacts of environmental disturbances and evaluate the efficacy of various interventions within a controlled virtual environment, as depicted in Figure 2. The methodology follows a structured four-step process comprising model construction, data synchronization, validation and iterative evaluation of hypothetical scenarios. These steps are supported by three key components:

- A stock and flow model ($SnF4RE$): a core component that captures the structural and behavioral aspects of the water body network while incorporating all influencing factors.
- An interface ($R2DT$) that facilitates the initialization and synchronization of real-world data with $SnF4DT$.
- A simulation engine (SE) for performing what-if analyses.

While the stock-and-flow model based digital twin construction and validation methods that extends modelling, validation and simulation methodology presented by (Sargent 2020) are described by (Barat et al. 2024), the key contribution of this research is a set of modelling abstractions for constructing a pragmatic digital twin of river ecosystem, i.e., $SnF4RE$, and its use in a real context. This section discusses the modelling abstractions in detail and provides an overview of the extended methodology.

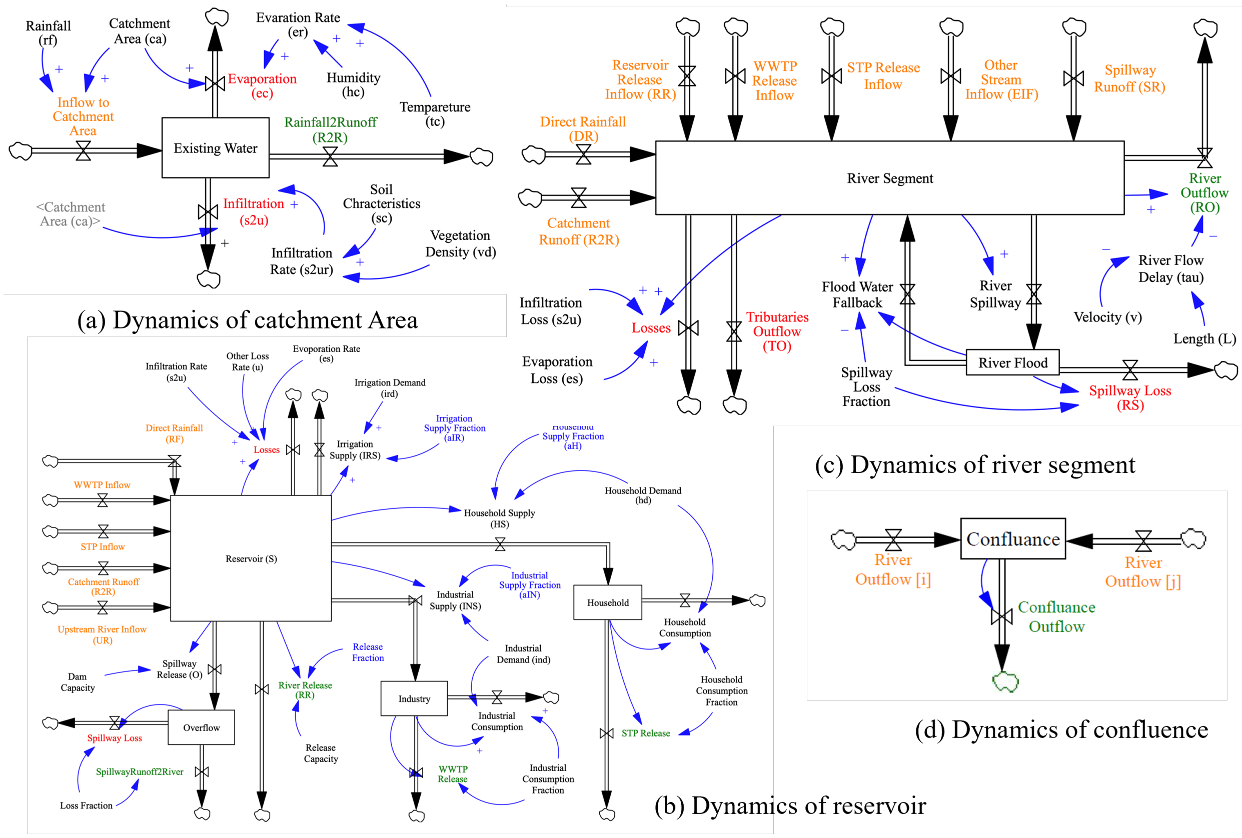
3.1 Models

In ecosystems that depend on rain-fed water, the network originates from catchment areas runoff and flows through large and diverse geographical regions before merging with other rivers through confluence or terminating in various sinks. This network comprises different types of water storage (e.g., lakes, dams, and other storage systems) as well as connecting water bodies, such as rivers, tributaries, and canals. We model this type of network using a set of abstractions, namely: catchment area, reservoir, river, and confluence. These abstractions primarily represent physical aspects and are designed to capture all relevant influencing factors from various perspectives as discussed in Section 2.

Catchment area (*CA*): A catchment area collects rainfall water (it could be from other sources, such as glacial meltwater, for other types of river ecosystem), temporarily stores accumulated water and contributes to river discharge once the shallow storage capacity is exceeded. While the cumulative water collected over a period t can be represented by a simple rainfall computation based on area, the transformation of rainfall (rf) into runoff, referred to as Rainfall2Runoff ($R2R$), is governed by more complex dynamics involving already accumulated water in the catchment area at time t represented as st_t, temporary storage capacity (st) and cumulative losses (cl) of the catchment area as discussed in (Wagener et al. 2007). The cumulative losses are primarily due to evaporation (ec) and infiltration, specifically seepage to the underground ($s2u$). Both volume loss due to evaporation and infiltration are functions of the catchment area under consideration (ca) along with average evaporation rate (er) and infiltration rate ($s2ur$) respectively. The evaporation rate is further influenced by the average temperature (tc) and average humidity (hc) within the catchment. In contrast, infiltration is dependent on a coefficient (f_i) related to soil characteristics (sc) and vegetation density (vd) of the area. We represent this complex dynamics of transformation, i.e., Rainfall2Runoff ($R2R$), using a Stock and Flow model shown in Figure 3(a) and associated equation is described below.

$$\begin{aligned} \mathrm{R2R}_{t+1} &= \max\big(0,\ \mathrm{st}_t + ca\cdot \mathrm{rf}_{\delta t} - ec_{\delta t} - s2u_{\delta t} - \mathrm{st}\big) \\ \text{where } ec_{\delta t} &= ca\cdot f_e(tc_{\delta t}, hc_{\delta t}), \\ s2u_{\delta t} &= ca\cdot f_i(sc, vd). \end{aligned} \tag{1}$$

In the context, the area under cover (ca) is derived from geographical surveys, while weather parameters such as temperature (tc_t), humidity (hc_t) and rainfall (rf_t) are time series data obtained from historical records, real-time sensor data and predictive weather models (for simulation). Infiltration loss ($s2u$) is modeled as a complex function (f_i) dependent on the soil characteristics (sc) and vegetation density and type, represented using (vd). Here, we estimate this function heuristically using historical data using similar method described in (Ghashghaei and Morid 2013). **Reservoir (*S*)**: We use a reservoir abstraction to represent all forms of water storage. This abstraction is modeled using a Stock and Flow (SnF) model as illustrated in Figure 3(b). The model includes multiple inflows, outflows, water gains and losses. Potential inflows to the reservoir include water from upstream rivers (UR) and runoff from its own catchment area of the reservoir S (i.e., $R2R^s$). Additionally, the reservoir accumulates various forms of water releases, such as treated wastewater from households via sewage treatment plants (R_{STP}) and industrial effluents released from wastewater treatment plants (R_{WWTP}). Direct rainfall (RF) over the reservoir also contributes to water gain. Water losses primarily occur due to evaporation (es) and infiltration ($s2u$) as previously discussed. Additional losses may arise from unaccounted withdrawals and water theft (U). During flooding or overflow events, the reservoir may also experience spillway losses. A portion of this overflow continues downstream into subsequent rivers and tributaries, whereas some of it is lost on land. Outflows from the reservoir consist of river releases (RR) and water distribution for various types of usage. In this model, we consider three primary distribution categories: Household Supply (HS), Irrigation Supply (IRS), Industrial Supply (INS). These distributions are dependent on their corresponding demands: household demand (hd), irrigation demand (ird), and industrial demand (ind). These demands fluctuate over time based on seasonal and situational factors. We incorporate three intervention levers representing decision variables (i.e., α_h, α_{ir} and α_{in}) that determine the fraction of each demand to be fulfilled. The actual volume of water distributed from the reservoir is governed by these fractions, as illustrated in the Figure 3(b). Furthermore, the SnF model shown in the Figure 3(b) depicts how consumed water is transformed into STP and $WWTP$ releases, which contribute to inflows in subsequent rivers and reservoirs. As shown below, the governing stock equation is dependent on the inflows and outflows of the reservoir stock.



(a) Dynamics of catchment Area

(b) Dynamics of reservoir

(c) Dynamics of river segment

(d) Dynamics of confluence

Figure 3: Modeling abstractions and their dynamics.

$$
\begin{aligned}
S_{t+1} = S_t + (UR_{\delta t} + R2R^S_{\delta t} + STP_{\delta t} + WWTP_{\delta t} + RF_{\delta t}) \\
- (RR_{\delta t} + \alpha_{hs} \cdot hd + \alpha_{irs} \cdot ird + \alpha_{ins} \cdot ind + ES_{\delta t} + s2u_{\delta t} + \Delta O_{\delta t} + U_{\delta t})
\end{aligned}
\tag{2}
$$

where, $UR = \sum_{R1}^{Rn} RiverOutflow_{Ri}$, and $\Delta O_{\delta t} = \max\left(0,\ S_t + \sum \text{Inflows}_{\delta t} - \sum \text{Outflows}_{\delta t} - S_{\text{capacity}}\right)$.

River Segment: A river segment is a dynamic flow element that receives water from multiple upstream sources and transports it to the downstream while being influenced by climate related factors, geographical heterogeneity and human interventions. Traditionally, the flow of river segment is represented using mass balance equation that considers transmission delay (τ) – the time takes for water to travel from the start (upstream) to the end (downstream) of the segment. We model the river segment as a stock that temporarily holds water while transporting it from the start to the end of the segment of length L. We consider the river storage at time t, denoted as St, and all inflows, outflows, and losses during the travel time τ, where inflows include upstream reservoir release (RR), discharges from sewage and wastewater treatment plants (STP and $WWTP$ respectively), external inflows from other rivers, canals or tributaries (EIF), possible spillway flows from nearby rivers (SR), catchment runoff (Rainfall2Runoff or $R2R$) and direct rainfall (DR) over the river surface. Losses during this period include evaporation (es) and infiltration or seepage to underground ($s2u$). The outflows are of two types: tributary outflows (TO), which represent water forking out of the ecosystem, and river outflow (RO), which continues to the next river segment, reservoir or confluence. Additionally, we consider spillway losses (RS) during overflow conditions, similar to the reservoir spillway dynamics described earlier.

In our model, the initial water storage in a river segment is approximated as $S_t = L \cdot w_{avg} \cdot d_{avg}$, where w_{avg} is the average width and d_{avg} is the average depth. The transportation delay is computed based on the length L of the segment and average flow velocity (v), where the velocity primarily depends on the depth of the water, width and resistances due to riverbed characteristics (e.g., roughness), geometrical properties (e.g., curvature), and water quality (e.g., sediment concentration). While this is a complex phenomenon, the travel time of a segment (τ) can be estimated using empirical formulas as discussed in (Jobson 2001), specifically the variation of Kirpitch Equation, or derived from sensing data. We approximate travel time τ as $\tau = L/v$, where average velocity v is derived from observed flow data and approximated using Float Method, i.e., a basic field technique using a floating object is timed for a known distance to estimate surface velocity. The stock dynamics over the interval τ is modeled using a mass balance formulation, where the volume of water in the segment at time $t+\tau$, denoted as $S_{t+\tau}$, accounts for inflows, losses, and outflows during the interval. This is captured in (3) and (4) below.

$$S_{t+\tau} = S_t + (RR_t + \sum_i (STP^i_\tau + WWTP^i_\tau + EIF^i_\tau + SR^i_\tau) + R2R_\tau + DR_\tau) - (es_\tau + s2u_\tau + TO_\tau + RO_\tau + RS_\tau) \tag{3}$$

The outflow of the river RO considering the transportation delay τ, therefore, can be derived as below.

$$RO = \frac{RR_t + \sum_i (STP^i_\tau + WWTP^i_\tau + EIF^i_\tau + SR^i_\tau) + R2R_\tau + DR_\tau}{\tau} - (es_\tau + s2u_\tau + TO_\tau + RS_\tau) \tag{4}$$

A representative stock and flow model of the river is shown in component (c) within Figure 3.

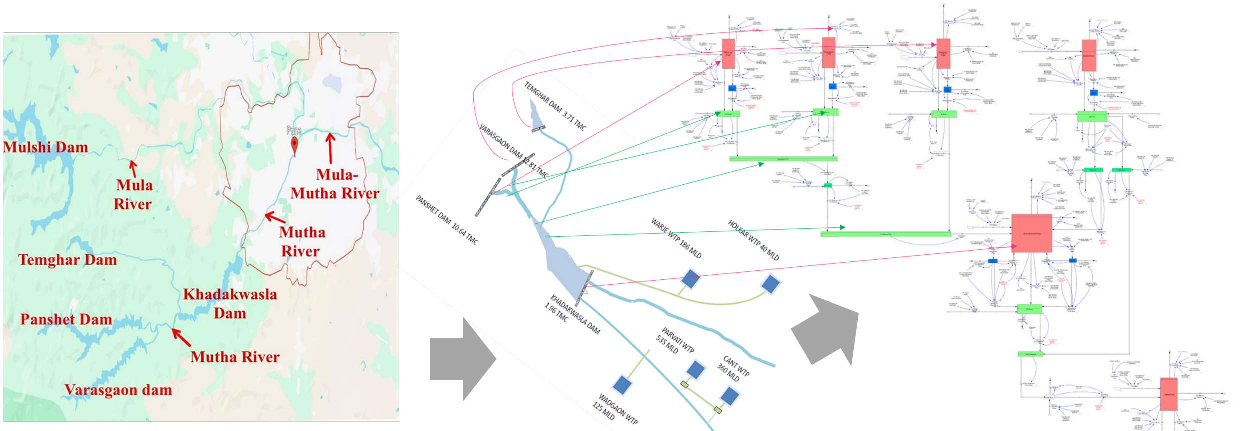
Confluence: We use an abstraction to represent a confluence, which primarily accumulates river outflows and transmits them to the subsequent segment. The stock-and-flow model for this abstraction is illustrated in component (d) within Figure 3.

3.2 Method

Our digital twin construction and use process evolve around four key steps that enable comprehensive modeling and simulation of river ecosystems. The first step, replication, involves constructing a virtual river ecosystem network ($SnF4RE$) using proposed modelling abstractions to mirror all elements of the real ecosystem. This is followed by synchronization, where the model is contextualized through the integration of real-world data by configuring auxiliary variables to ensure the digital representation remains dynamically aligned with evolving conditions (using $R2DT$ as shown in Figure 2). The third step, validation, considers traditional operational validity as recommended by (Sargent 2020), where the constructed model, configured and synchronized using historical scenarios, is simulated and simulation observations are compared with real-life observations. The differences are correlated with the real observations and iteratively refined until the simulation results satisfactorily match them. This step is context-specific and needs to be evaluated for each river under consideration. Finally, simulation-based exploration involves what-if simulations to comprehend system complexities, evaluate the effects of potential environmental disruptions, and assess the effectiveness of various intervention strategies as illustrated in Figure 2. While adopting standard techniques for synchronization, validation and simulation-based exploration, we introduce a specialized replication step that leverages the proposed abstractions and considers a set of assumptions to pragmatically construct and configure a digital twin of the river ecosystem as described below.

To replicate a river ecosystem, we construct a network model by composing four core abstractions: catchment areas, reservoirs, river segments and confluences, as a directed graph. These abstractions function as modular building blocks, each instantiated and configured to reflect the characteristics and behaviors of real-world hydrological entities. The replication process involves three key steps: identifying the appropriate abstraction for each hydrological component, configuring them with relevant parameters (e.g., inflows, outflows, capacities, and other characteristics), and interconnecting them through their inflow-outflow relationships to form a coherent flow network.

To balance fidelity with scalability, we apply a set of modeling assumptions. Each river segment is modeled with uniform properties, such as average width, depth and climate conditions, while reservoirs are treated as homogeneous units in terms of their internal structure and environmental characteristics. Spatial



(a) River ecosystem (b) Schematic representation of the network (c) Configured SnF model (SnF4RE)

Figure 4: A real river ecosystem with schematic representation of the ecosystem using SnF building blocks.

heterogeneity and environmental variations can be captured by decomposing large element into smaller and more granular segments where necessary.

While the proposed approach addresses heterogeneity at a pragmatic meso-level, it approximates several finer-grained variabilities, such as fluctuations in width, depth and river curvature of the river. Additionally, uncertainties related to climatic conditions, including rainfall and humidity, are also approximated into the model. These factors are treated as auxiliary variables having well-defined ranges and probability distributions. To better understand the potential emergent phenomena caused by such uncertainties and model assumptions, multiple simulation runs are conducted using varied points sampled from these distributions.

4 CASE STUDY

This case study considers a part of the water body network that originates from the Western Ghats of India and flows through the Deccan Plateau spanning approximately 350 km. A google-map view of the ecosystem under consideration and its schematic representation using the proposed modelling abstractions are shown in component (a) and (b) within Figure 4 respectively. The network begins in the high-rainfall catchments of the Sahyadri hills and includes four major dams (i.e., reservoirs): Panshet (Capacity: 294 MCM, Catchment area: 120 km²), Varasgaon (Capacity: 362 MCM, Catchment area: 130 km²), Temghar (Capacity: 104 MCM, Catchment area: 190 km²), and Mulshi (Capacity: 523 MCM, Catchment area: 318 km²). Panshet (PD), Varasgaon (VD), and Temghar Dam (TD) receive rainwater from their respective catchment areas and regulate the flow of the Mutha River through small tributaries (VD and PD River) of lengths approximately 2-3 km respectively. The Mutha River (PV River) flows for approximately 10 km and converges at Khadakwasla Dam (KD, Capacity: 55 MCM, Catchment area: 501 km², receives regulated upstream flow). Another river, known as the Mula (MD River), originates from the Mulshi Dam (MD) and flows around 60 kilometers before reaching its confluence with the Mutha River Next Segment (KD River), which travels approximately 15 km from Khadakwasla Dam to form the Mula-Mutha River (KM River). This river then continues eastward for approximately 160 kilometers before reaching Ujjani Dam (UD, Capacity: 2,016 MCM, Catchment area: 14,850 km²). While all dams supply irrigation water to their respective areas, the Khadakwasla Dam functions as a key balancing reservoir located near Pune city and fulfills most of the household and industrial water demands. This case study considers the water body network up to Ujjani Dam as the water from Ujjani is regulated by interstate policy and subsequently diverted to other state. Geographically and climatologically, the basin spans highly heterogeneous zones. The upper catchments areas experience tropical monsoon rainfall exceeding 2,000 millimeters annually and exhibit steep gradients with dense vegetation resulting in rapid runoff and high inflow rates to the reservoirs.

Table 1: Data Sources for configurations and synchronization.

Parameters	Sources
Dam capacity, catchment area and other configuration parameters	1. (Central Ground Water Board) 2. (Numerical The Universe in Number)
Rainfall, Humidity, Temperature	(Weather and Climate)
River Length, Width	(Google)
Other Sources	(Wikipedia)

In contrast, the downstream plains, particularly those within Pune city, lie in the rain-shadow region and receive less than 600 millimeters of rainfall annually. Due to these environmental variations, particularly in rainfall, the region experiences both urban flooding and water scarcity in most years. Decision- makers face increasing challenges in managing and preserving water resources to reduce flood risks and mitigate water shortages while ensuring that water demands are effectively met.

Construction, Validation, and Synchronization: We represent the water network up to Ujjani Dam using four abstractions and configure these elements using the data collected from different sources. Constructed network using SnF building blocks is shown in component (c) within Figure 4. In the figure, red blocks represent dams, and green blocks denote river segments. The diagram also illustrates illustrative mappings from real-world elements to their corresponding model abstractions (Although individual model labels are not readable, the objective here is to convey the overall network structure and the mappings).

Key data sources considered in our experiment are described in Table 1. The configured model synchronized with real data, i.e. digital twin, is then simulated using known parameters including historical rainfall, temperature profile and water release rates for two years (i.e., 2023–2024). A set of simulation results is presented in Figure 5. These simulation outcomes are compared with observed data from the real system to assess the accuracy of the digital twin and establish its fidelity.

What-if Scenario: Upon establishing the faithfulness of the constructed digital twin, we explored varying possibilities. In this section, we discuss five scenarios: a baseline scenario describing typical conditions experienced in the last two years, two weather-related disruptive situations capturing high and low rainfall scenarios, and two potential mitigation strategies as described in Table 2. This table also summarizes the key observations from our simulation experiments. Each experiment was simulated over a two-year period to capture seasonal variability and observed their impacts. For the baseline situation, all reservoir levels cycled between 60–100% capacity with peaks from July to August except Khadakwasla Dam which touched 30% mark around May. Despite no abnormally high rainfall, 3 out of 6 reservoirs experienced moderate overflow events lasting not more than a month between July and August. River depths ranged between 20%–80% of their limits, while three river segments (VD, PD, PV) crossed 100% mark and experienced mild overflows during peak monsoon months. The simulation also indicates that the household, industrial and irrigation demands are fully met throughout the year. These observations match with reality - this demonstrates the faithfulness of the model and simulation. We then constructed a heavy rainfall scenario by increasing total annual precipitation by 45 percent (shown in row 2). In this experiment, all reservoirs remained at or near full capacity for approximately 3 months compared to less than a month in the baseline. Monthly overflow volumes peaked at nearly 50% higher than the base case and lasted 2 months longer than baseline. River depths exceeded limits in five out of nine river segments. While all demands were fulfilled, the flooding situation worsened and impacted citizens.

To address this limitation, a mitigation scenario involving pre-monsoon buffer releases of 80-85% more water than base from all the Dams was introduced. This experiment aimed to explore the efficacy of a strategy to create an operational headroom during non-rainy seasons. As reported in row 3 of Table 2, this intervention reduced peak overflow volumes by 30–35 percent and halved the duration of flooding events. River depths were better controlled and only three segments (VD, PD and PV) exceeded their limits causing moderate flooding. Demand fulfillment remained unaffected, which indicates that anticipatory releases can

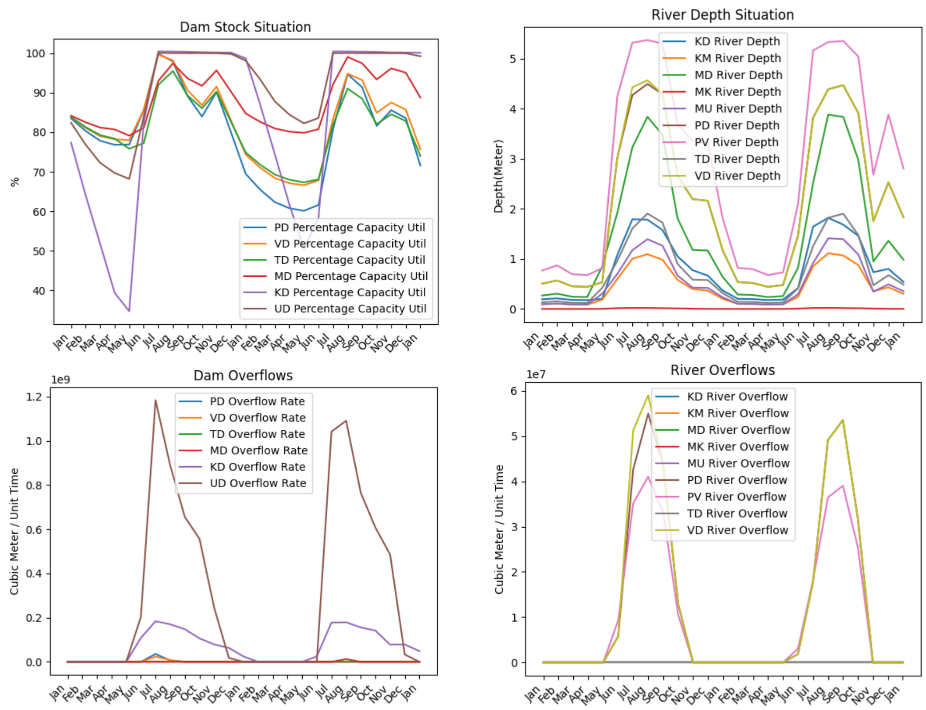


Figure 5: Base experiments.

simultaneously reduce flood stress and maintain supply security. However, determining the optimal volume of release to further reduce flood risk without compromising demand fulfillment remains a question that requires further exploration.

Following the flood-risk evaluation, the next objective was to assess system behaviour under less rainfall - a low rainfall scenario was created with a 70 percent reduction in annual rainfall to mimic a drought-like condition. The simulation results showed that most reservoirs fell below 40 percent storage during the critical supply window while Khadakwasla fell below 5 percent. River flows diminished substantially, and irrigation water was affected under existing distribution policy, as shown in Row 4 of Table 2. To counteract these deficits, an essential release strategy was evaluated in a fifth experiment, where water release was reduced by 20 percent throughout the year from all dams. The results were encouraging: five dams maintained storage above 50 percent and seasonal recovery was observed earlier than in the non-mitigated case. However, this strategy failed to maintain sufficient river depth in non-rainy seasons. The intervention demonstrated that in dry years, carefully timed releases are more effective than conservation alone.

While we experimented several scenarios and mitigation strategies, here we discussed a limited set of scenarios to illustrate the usefulness of our digital twin in supporting informed decision-making in the face of uncertainties.

5 CONCLUSION

In this paper, we presented a pragmatic digital twin of a river ecosystem. The key contributions include four stock-and-flow-based modeling abstractions, which are sufficiently expressive to capture the interconnected dynamics of catchment areas, reservoirs, river segments and their inherent environmental and geographical characteristics. We demonstrated how these modeling abstractions can be effectively used to configure a complex water body network, and how the model can be synchronized with historical and observed data to mimic the behavior of the real-world system. The applicability and usefulness of the proposed framework were illustrated through a case study involving a real river basin in western India along with the simulation

Table 2: An overview of simulation results.

#	Scenario	Observation on Reservoirs	Dam Overflow	Observation on Rivers	River Overflow
1	Base	All Reservoirs are **Nearly Full** between July–Aug and Oct–Nov; Normal in Jan–Apr; and Low in May. KD dropped to **Low** in May.	**Moderate** overflow (< 200 MCM) from 3 of 6 dams lasting 3–4 weeks.	River depths ranged from **Normal to Above Safe**. 3 river segments (VD, PD, PV) crossed **Danger** mark. Peaks lasted ∼6 weeks.	**Mild** overflows in 3 river segments (40–60 MCM), each lasting ∼6 weeks during monsoon.
2	Heavy Rainfall	All Reservoirs remained **Overflowed** for ∼2 months, ∼2x longer than base.	All dams **Overflowed** with **Peak overflow** volumes between (200–300 MCM), ∼50% higher than base. Duration doubled.	5 of the 9 river segments exceeded **Danger** levels which persisted for ∼3.5 months.	**Severe** overflow in 5 segments with ∼66% increase from base. Duration is significantly higher than base.
3	Heavy Rainfall Mitigation (Pre-emptive Release)	4 Reservoirs came down to **Near Full** while 2 reservoirs were still **Overflowed**. KD stock improved to **Normal** in contrast to base and heavy rainfall (**Low**).	Overflow volumes reduced by **30–40%** compared to Heavy rainfall, while peak duration reduced to ∼4–5 weeks.	Only 3 segments (VD, PD, PV) reached **Danger** level. Others remained between **Above Safe to Safe** zone.	**High** overflows (50–80 MCM), fewer affected rivers compared to Heavy rainfall; flooding duration halved compared to overflow case.
4	Low Rainfall (∼70% Rainfall)	All dams dropped to **Low – Almost Empty** during Apr–Jul. Khadakwasla was most affected.	**No overflow** observed at any dam throughout the year.	Persistently **Low** flow. All river levels dropped to **Low** most of the year. Several segments reached **Almost Dry**.	**No overflow** in any segment. **Severe** water stress observed, especially for irrigation and household.
5	Low Rainfall Mitigation (20% less Release)	5 reservoirs stayed **Normal** most of the year. Seasonal recovery started early (Jul vs. Oct in unmitigated case) and UD and KD reached **Near Full**.	**No overflow** observed at any dam throughout the year.	**Worsen** river flow: most segments stayed in **Almost Dry** in non-rainy days.	**No overflow** observed. Better than low rainfall case but river health remained suboptimal. All demand fulfilled.
Dam Condition: Overflow: >100%, Near Full: 90–100%, Above Average: 70–90%, Normal: 40–70%, Low: 20–40%, Almost Empty: 0–20%					
River water levels: Danger: >100%, Above Safe: 80–100%, Safe: 60–80%, Normal: 40–60%, Low: 20–40%, Almost Dry: 0–20%					

of a base scenario and four representative scenarios as interventions. These scenarios covered typical, extreme and mitigated conditions highlighting the ability of the proposed digital twin for supporting both strategic planning and operational decision-making.

We argue that such a digital twin can assist decision-makers in making proactive interventions to avoid disruptions caused by environmental variability, as well as reactive responses to manage emerging on-ground situations with greater certainty. Our early experience with real-life simulations has been encouraging; however, further work is needed to enhance model validation and optimize sensor placement for real-time synchronization and improved insights.

REFERENCES

Ahmad, S., and S. P. Simonovic. 2000. "System Dynamics Modeling of Reservoir Operations for Flood Management". *Journal of Computing in Civil Engineering* 14(3):190–198.

Barat, S., P. Deshpande, A. Yadav, A. A. Paranjape, V. Kulkarni, B. Rai, *et al*. 2024. "Digital Twin as an Aid for Optimizing Climacteric Fruit Supply Chains". In *2024 Annual Modeling and Simulation Conference (ANNSIM)*, 1–14.

Central Ground Water Board. "Aquifer Mapping and Management of Ground Water Resources". https://www.cgwb.gov.in/sites/default/files/2022-11/pune_f.pdf, accessed 19.08.2025.

Devia, G. K., B. P. Ganasri, and G. S. Dwarakish. 2015. "A Review on Hydrological Models". *Aquatic Procedia* 4:1001–1007.

Ghashghaei, B., and Morid. 2013. "Rainfall-Runoff Modeling in a Watershed Scale Using an Object Oriented Approach Based on the Concepts of System Dynamics". *Water Resources Management* 27(3):5119–5141.

Ghoreishi, M., S. Razavi, and A. Elshorbagy. 2021. "Understanding Human Adaptation to Drought: Agent-Based Agricultural Water Demand Modeling in the Bow River Basin, Canada". *Hydrological Sciences Journal* 66(3):389–407.

Google. "Google Earth". https://earth.google.com/, accessed 19.08.2025.

Jobson, H. E. 2001. "Predicting River Travel Time from Hydraulic Characteristics". *Journal of Hydraulic Engineering* 127(11):911–918.

Kotir, J. H., C. Smith, G. Brown, N. Marshall, and R. Johnstone. 2016. "A System Dynamics Simulation Model for Sustainable Water Resources Management and Agricultural Development in the Volta River Basin, Ghana". *Science of the Total Environment* 573:444–457.

Mishra, B. K., P. Kumar, C. Saraswat, S. Chakraborty, and A. Gautam. 2021. "Water Security in a Changing Environment: Concept, Challenges and Solutions". *Water* 13(4):490.

Numerical The Universe in Number. "Live Storage, Water Levels and Rainfall Data for Dams in Pune, Maharashtra". https://www.numerical.co.in/numerons/collection/5985e2501d6090dc136dc381, accessed 19.08.2025.

Pal, D., H. Marttila, P. Ala-Aho, E. Lotsari, A.-K. Ronkanen, C. Gonzales-Inca, *et al*. 2025. "Blueprint Conceptualization for a River Basin's Digital Twin". *Hydrology Research* 56(3):197–212.

Sargent, R. G. 2020. "Verification and Validation of Simulation Models: An Advanced Tutorial". In *2020 Winter Simulation Conference (WSC)*, 16–29 https://doi.org/10.1109/WSC48552.2020.9384052.

Wagener, T., M. Sivapalan, P. Troch, and R. Woods. 2007. "Catchment Classification and Hydrologic Similarity". *Geography Compass* 1(4):901–931.

Weather and Climate. "Pune, India - Weather and Climate". https://weather-and-climate.com/average-monthly-Rainfall-Temperature-Sunshine,pune,India, accessed 19.08.2025.

Wikipedia. "Dams in Pune District". https://en.wikipedia.org/wiki/Category:Dams_in_Pune_district, accessed 19.08.2025.

Yang, Y., C. Xie, Z. Fan, Z. Xu, B. W. Melville, G. Liu *et al*. 2024. "Digital Twinning of River Basins Towards Full-Scale, Sustainable and Equitable Water Management and Disaster Mitigation". *Npj Natural Hazards* 1(1):1–7.

Yates, D., J. Sieber, D. Purkey, and A. Huber-Lee. 2005. "WEAP21—A Demand-, Priority-, and Preference-Driven Water Planning model: Part 1: Model Characteristics". *Water International* 30(4):487–500.

AUTHOR BIOGRAPHIES

SOUVIK BARAT is a Principal Scientist at Tata Consultancy Services (TCS) Research, India and a visiting researcher at Middlesex University London. His research interests include Digital Twin technology, Modelling and simulation of complex systems, and Model Driven Engineering. His email address is souvik.barat@tcs.com.

ABHISHEK YADAV is a Researcher at TCS Research, Pune, India. His current research interests are Software Engineering, Digital Twin modeling and Simulation, and Generative AI. His email address is y.abhishek1@tcs.com.

VINAY KULKARNI is a TCS Fellow at TCS Research, and a Fellow of Indian Academy of Engineering. His current research interests are Digital Twins, Adaptive Enterprises, and Software Engineering. His email address is vinay.vkulkarni@tcs.com.

GURUDAS NULKAR is a Professor and Director at the Centre for Sustainable Development, Gokhale Institute of Politics and Economics (GIPE), Pune. His email address is gurudas.nulkar@gipe.ac.in.

SOOMRIT CHATTOPADHYAY is a Senior Research Associate at the Centre for Sustainable Development at GIPE Pune, leading initiatives in urban water management and circular economy. His email address is soomrit.chattopadhyay@gipe.ac.in.

ASHWINI KESKAR is the Program Manager for Sustainability and Environment at the Pune Knowledge Cluster (PKC), focusing on water circularity, climate resilience, and forest landscape restoration. Her email address is ashwini.keskar@pkc.org.in.