



Modelling water quantity and quality for integrated water cycle management with the Water Systems Integrated Modelling framework (WSIMOD) software

Barnaby Dobson, Leyang Liu, and Ana Mijic

Department of Civil and Environmental Engineering, Imperial College London, UK

Correspondence: Barnaby Dobson (b.dobson@imperial.ac.uk)

Received: 19 July 2023 – Discussion started: 13 October 2023

Revised: 10 April 2024 – Accepted: 11 April 2024 – Published: 30 May 2024

Abstract. Problems of water system integration occur when a model’s boundaries are too narrow to capture interactions and feedbacks across the terrestrial water cycle. We propose that integrated water systems models are required to overcome them and are necessary to understand emergent system behaviour, to expand model boundaries, to evaluate interventions, and to ensure simulations reflect stakeholder goals. We present the Water Systems Integrated Modelling framework (WSIMOD) software as one such approach and describe its theoretical basis, covering the node and arc nature of simulations, the integration framework that enables communication between model elements, and the model orchestration to customise interactions. We highlight data requirements for creating such a model and the potential for future development and refinement. WSIMOD offers a flexible and powerful approach to represent water systems, and we hope it will encourage further research and application into using model integration towards achieving sustainable and resilient water management.

1 Introduction

Water fluxes and their pollution concentration are influenced by the interactions and management of all components that make up the human-altered terrestrial water cycle, including but not limited to hydrological and groundwater catchments, agriculture, upstream rivers, reservoirs, freshwater treatment, distribution networks, residential and non-residential water consumers, foul sewers, urban drainage, storm sewers, and wastewater treatment works (and the many physical and op-

erational processes within each component). The importance of this interconnectedness is most evident in rivers, which reflect the overall condition of the catchment system since they aggregate behaviour over such large areas (Dobson and Mijic, 2020; Kirchner, 2009). Hydrological catchments are rarely dominated by the behaviour of any specific component or any individual stakeholder’s decisions. For example, 60 % of English hydrological catchments that do not achieve a “good” status in the Water Framework Directive (WFD) do so because of multiple different pollution sources, with wastewater infrastructure and agriculture being the most prevalent drivers, each affecting 50 % of hydrological catchments (Environment Agency, 2020a). The implication is that a modelled representation focussing on any individual component is unlikely to give accurate estimates of impacts beyond that component (Dobson et al., 2019; Beven, 2007; Schmitt and Huber, 2006; Blair et al., 2019). Furthermore, estimates within a subsystem representation may be inaccurate if sensible boundary conditions cannot be defined, something that water managers are highly sensitive to (Höller-mann and Evers, 2017).

We briefly pose a variety of questions to further motivate these challenges, some of which are developed throughout the paper. Can changes in agriculture fertilising behaviour offset the increased sewage due to population changes downstream (Liu et al., 2022)? Will water-efficient appliances concentrate household sewage to the extent that river water quality is worsened (Mott MacDonald, 2023)? Will wastewater reuse worsen low-flow conditions in rivers downstream of the wastewater treatment plant where it is implemented (Mott MacDonald, 2023)? Can upstream water supply river abstractions be strategically reduced on sewer overflow days

to dilute the spill without compromising supply reliability (Dobson and Mijic, 2020)? Will increasing prevalence of “working from home” change where sewage is produced to have knock-on impacts on wastewater infrastructure and rivers (Dobson et al., 2021)? In a multi-polluter hydrological catchment, can combinations of pollution reduction achieve a target water quality at a downstream checkpoint (Liu et al., 2023b)? For each of these questions, there will be water systems where these questions are true and cases where they are not. We term these problems of water systems integration, and it follows that understanding the terrestrial water cycle (“water cycle” hereafter) as a whole is needed to address them. Models that take such an approach will better capture component boundaries and the wider impacts of stakeholder decisions, ultimately enabling more accurate representations of water quality in rivers, which is essential to effectively manage, for example, water supply (Mortazavi-Naeini et al., 2019) and biodiversity (Dobson et al., 2022a).

In this paper we introduce the theoretical underpinning behind a novel method for modelling integrated water systems to address these challenges. Firstly, the need for integrated water cycle simulation models is explained, including their current coverage. The importance of parsimonious representations within an integrated model is then discussed, along with the methods used to achieve integration.

The environmental modelling research community has responded to problems of water systems integration primarily through computer simulation models (Bach et al., 2014; Rauch et al., 2017; Whitehead et al., 1998; Best et al., 2011; Douglas-Mankin et al., 2010; Tscheikner-Gratl et al., 2019). We distinguish an integrated water system modelling approach from a system dynamics approach (see Zomorodian et al., 2018) by further specifying that component representations must have a physical basis, which is needed to link observational data to model behaviour and to capture interventions (e.g. new infrastructure or changes to operations). We define integrated water system models as those which link component representations to capture and understand the complex interactions and feedbacks that occur between components. We categorise the four key goals of these models: (1) to understand which fundamental processes drive emergent behaviour at a whole-water system scale; (2) to avoid simulation inaccuracies caused by narrow boundary conditions; (3) to test interventions to the physical system or operational behaviour in order to understand their water-cycle-wide impacts or interactions; and (4) to capture impacts that align more closely with desired water system outcomes, in addition to performance indicators of individual components. For example, in-river pollutant concentration is a better indicator of wastewater system performance than the more typically monitored number of sewer spills (Giakoumis and Voulvoulis, 2023).

In addressing problems of water systems integration, existing modelling approaches have made significant progress. Bach et al. (2014) set out a comprehensive typology for in-

tegrated urban water systems modelling. However, among the reviewed models, only CityDrain3 (Burger et al., 2016), WEST (Vanhooren et al., 2003), and SIMBA (IFAK, 2007) can represent receiving water bodies (i.e. rivers), which is where the importance of an integrated representation is most pronounced. Furthermore, due to the urban focus of these models, the ability to simulate pollution concentrations in receiving waters impacted by upstream hydrological catchments is highly limited yet is central to quantifying in-river impacts (Liu et al., 2022). A more recent effort to characterise integrated water systems modelling places importance on in-river conditions (Tscheikner-Gratl et al., 2019) and presents a comprehensive review of urban and rural water cycles and their impacts on rivers. However, the reviewed modelling approaches omit some key factors: the importance of water resources infrastructure, which plays a significant role in concentrating pollution during low flows if abstractions take place; the relevance of groundwater, which provides baseflow to dilute pollution during low flow periods; and consideration of agricultural processes and associated pollution that results from them, which is a critical source of water pollution worldwide (Tang et al., 2021; Mateo-Sagasta et al., 2017) and the second most common hydrological catchment pollution source in England (Environment Agency, 2020a). Integrated models that capture groundwater and agricultural processes are present in the modelling literature, such as INCA (Whitehead et al., 1998) and HYPE (Lindström et al., 2010); however, in contrast, these are limited by their ability to capture urban systems. Thus, while water systems integration is well-served from a rural or urban modelling perspective, we identify that there is not yet an approach that offers a self-contained representation to capture all key processes required to model in-river water quality at a whole-water-cycle scale.

A further critical factor in creating an integrated water systems model is how components are represented. In general, current approaches have favoured identifying pre-existing detailed component representations which are then integrated (Schmitt and Huber, 2006). For example, DANCE4Water (Rauch et al., 2017), which is the most comprehensive application to date, includes high-resolution and sophisticated models for a wide range of urban components. However, as more and more components are captured by integrated modelling, it becomes increasingly difficult to parameterise such detailed models. Simply combining separately calibrated models provides no guarantee of performance as a whole (Lee, 1973). Meanwhile, integrated models typically have many parameters that may compensate each other, thus making calibration a challenging and risky process (Voinov and Shugart, 2013). An alternative approach is to forego calibration altogether by adopting parsimonious models with fewer parameters and ideally deriving those parameters from best available data (Dobson et al., 2021). Although complicated modelling approaches are needed for specific tasks such as design, these approaches are also more difficult to ap-

ply widely and thus may hinder the goals of integrated water systems modelling. For example, building scientific understanding requires repeated testing of an approach in various locations, and customising the model to match local conditions is essential when representing interventions. Therefore, a modelling approach that can be easily deployed on a wider scale is of significant benefit to problems of water systems integration.

To ensure that a range of water system configurations can be accommodated, flexibility or customisability must be incorporated into the approach for integration. Integration approaches vary broadly between tightly coupled and loosely coupled. In a tightly coupled approach, equations and interactions are pre-defined to create a self-contained integrated representation, such as with JULES (Best et al., 2011) or INCA (Whitehead et al., 1998). In a loosely coupled approach, component representations are self-contained, and the integration occurs by facilitating their interactions, filling the role as a message-passing interface, such as with OpenMI (Harpham et al., 2019) and DAnCE4Water (Rauch et al., 2017). Belete et al. (2017) describe the arrangement of components and their interactions as integrated model orchestration, highlighting that different orchestrations are suitable for different applications. While looser coupling provides greater control over orchestration, and thus greater ability to customise and capture a wide variety of systems, it also creates a higher user burden to set up and understand many subsystems, considered to be a key barrier to the uptake of such approaches (Zomorodian et al., 2018). Conversely, a tightly coupled model that represents the same components as a loosely coupled one may be easier to set up but typically offers less control over orchestration. In the middle ground is an integrated representation that gives flexibility around orchestration but comes with self-contained components that do not need to be onerously set up by a user, such as the CityDrain3 software for modelling urban drainage systems (Burger et al., 2016). We propose that this middle ground is the most beneficial for a modeller and believe that such an approach to integration is the most productive avenue towards creating highly flexible, user-friendly models of the integrated water cycle.

The concepts introduced above suggest that, for many problems of water systems integration, capturing a broad representation of the water cycle and interactions between its components is equally important as detailed component representations. We have created a tool to implement this modelling philosophy, the Water Systems Integrated Modelling framework (WSIMOD), which is an open-source Python package for flexible and customisable simulations of the water cycle that treats the physical components of the water cycle as nodes connected by arcs that convey water and pollutant fluxes between them. The software source code and online tutorials are published by Dobson et al. (2023a); in contrast, this paper presents WSIMOD's theoretical underpinning with a discussion on model setup and of integrated wa-

ter system modelling in general. To address the difficulties in application associated with integrated modelling mentioned above, WSIMOD contains a library of built-in component representations covering a more complete water cycle coverage than any identified integrated models and a default but customisable orchestration deemed to be suitable for many catchments and regional water systems coordination. Where possible these representations are based on parsimonious and peer-reviewed models. Extensive model documentation with worked examples is provided online (WSIMOD documentation, 2024), enabling users to gain confidence and become familiar with using WSIMOD.

2 WSIMOD

WSIMOD is an integrated modelling framework that provides ready-to-use objects (nodes, arcs, water stores, and model orchestration) that are suitable for a wide range of water systems and described in greater detail in the following sections. However, WSIMOD is not intended to be a one-size-fits-all solution; indeed, the ubiquity of non-textbook water systems led us to create a more customisable modelling approach in the first place. This paper describes the theory behind WSIMOD in general and user-friendly terms, avoiding the use of equations and technical details, while further documentation can be found online (WSIMOD documentation, 2024). The WSIMOD framework is implemented in Python 3, which is widely practised in the environmental modelling community and facilitates quick setup and easy customisation. WSIMOD is the combined effort of many studies, conducted as part of the CAMELLIA (Community Management for a Liveable London) project (<https://www.camelliawater.org/>, last access: 24 May 2024), which are linked to relevant sections of the model description to highlight the range of possible applications.

An example WSIMOD model schematic is shown in Fig. 1. The arrangement is based on the city of Oxford, UK, and is used in a demonstration in Sect. 3. Oxford was selected because its water cycle is highly integrated but also spatially contained. We first use this schematic to explain the generic water cycle implemented in WSIMOD; however, we note that this is customisable – see Sects. 2.2.3 and 2.3. A given time step (see “Orchestration” in Fig. 1) begins at freshwater treatment works (“FWTW”), which pump water from various sources (groundwater and a reservoir in this example), treat that water to fill their service reservoirs, and send sludge to the “Foul sewer” node. Note that capitals are used hereafter to denote nodes. Demand nodes (i.e. population and other water users) then generate water demand, retrieve water from the FWTW service reservoirs via a Distribution node, satisfy any garden irrigation demands (the link between Demand and Land), and discharge wastewater to foul sewers. The hydrological and agricultural processes are then run within the Land node; these are substantial and have a tutorial

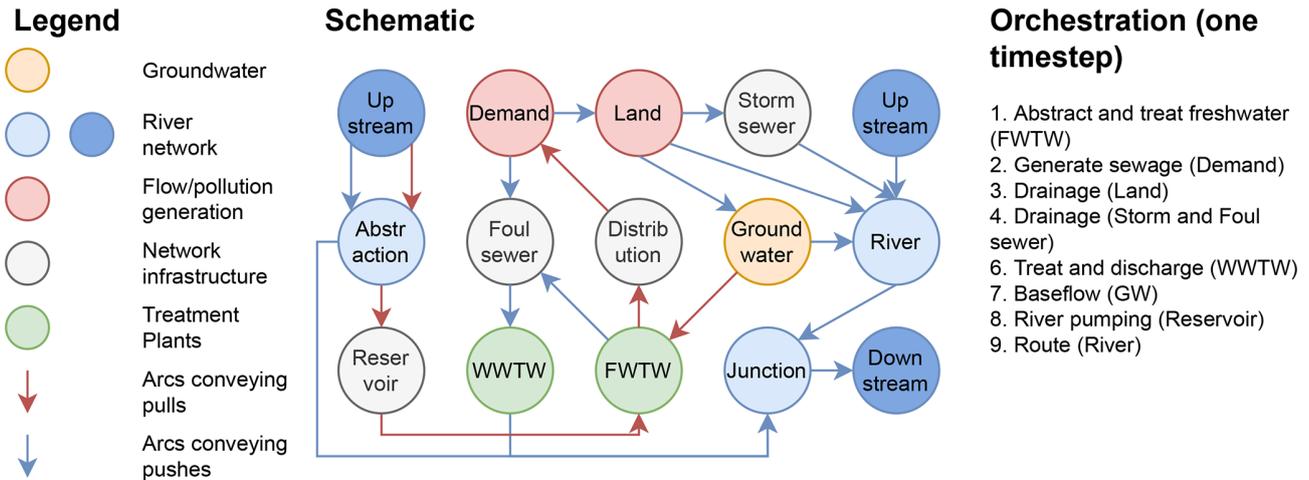


Figure 1. An example WSIMOD model for Oxford, UK. Orchestration is shown demonstrating the high-level functions called for each time step. Nodes are shown as circles and arcs as arrows. WWTW stands for wastewater treatment works, and FWTW stands for freshwater treatment works.

in the online documentation (https://imperialcollegelondon.github.io/wsi/demo/scripts/land_demo/, last access: 8 April 2024). In short, they track crop growth and calendars; apply fertilisers, nutrient cycling, and erosion; and generate impervious runoff (to Storm sewers), percolation (to Groundwater), and surface/sub-surface flows to the River. Foul sewer and Storm sewer node processes then run, generating flows to the wastewater treatment works (WWTW) and into the River, respectively. Baseflows are then calculated based on conditions in the Groundwater node, and reservoir abstractions are made. The time step completes by routing upstream to downstream along the river network.

The aggregated nature of the components depicted, e.g. an individual node represents the entire storm sewer network, is a common feature of many WSIMOD models (but not always; see Dobson et al., 2022b). A general philosophy that drives WSIMOD, as anticipated in the Introduction, is that it is easier to introduce complexity into an already integrated model rather than to integrate separate complex models. To achieve such customisation and introduction of new behaviours, WSIMOD uses object-oriented programming (OOP), which classifies components by common attributes and behaviours (classes). All objects in WSIMOD are a subclass of `WSIObj`, which predefines efficient arithmetic operations for water quality and volume; however, users will typically instead interact with the subclasses described in the following sections. Additionally, users may customise a model's high-level control over how interactions take place within a time step or the model's orchestration (Belete et al., 2017), which is a unique feature of WSIMOD and further developed in Sect. 2.3. Thus, while Fig. 1 depicts one possible arrangement and selection of nodes and arcs, a wide variety of water systems can be represented. Section 3 presents this Oxford WSIMOD model

in its simplest integrated form and demonstrates how to introduce complexity into such a model (specifically, in this example, to introduce more complex water resources behaviour). If readers wish to be involved with the development of WSIMOD, e.g. identify bugs, or require further clarification in terms of documentation, we recommend viewing the Contributing section of the main repository page (<https://github.com/ImperialCollegeLondon/wsi>, last access: 8 April 2024).

2.1 Nodes represent water cycle components

Physical representations of the different components in the water cycle are typically implemented as WSIMOD nodes, Fig. 1. In the software implementation, all nodes are instances of the “Node” class or its subclasses. Formulation of components as nodes using OOP draws heavily on the `CityDrain3` software (Burger et al., 2016). Our generic definition allows nodes to represent diverse entities, e.g. a collection of manholes representing a region of sewer network or individual manholes that can be connected to represent a sewer network, as demonstrated in Dobson et al. (2022b). In this section we describe the Node class, summarise the existing node subclasses currently implemented in WSIMOD, and describe how to customise them.

2.1.1 The Node class

The Node class in WSIMOD is a generic class that is expected to be the parent of any component represented. The class captures three key behaviours. Firstly, it predefines defaults for much of the interaction functionality later described in Sect. 2.2. Secondly, it contains a variety of useful functions to interact with other nodes via arcs (see Sect. 2.2.1). Thirdly, it enables mass balance checking (see Sect. 2.3.2).

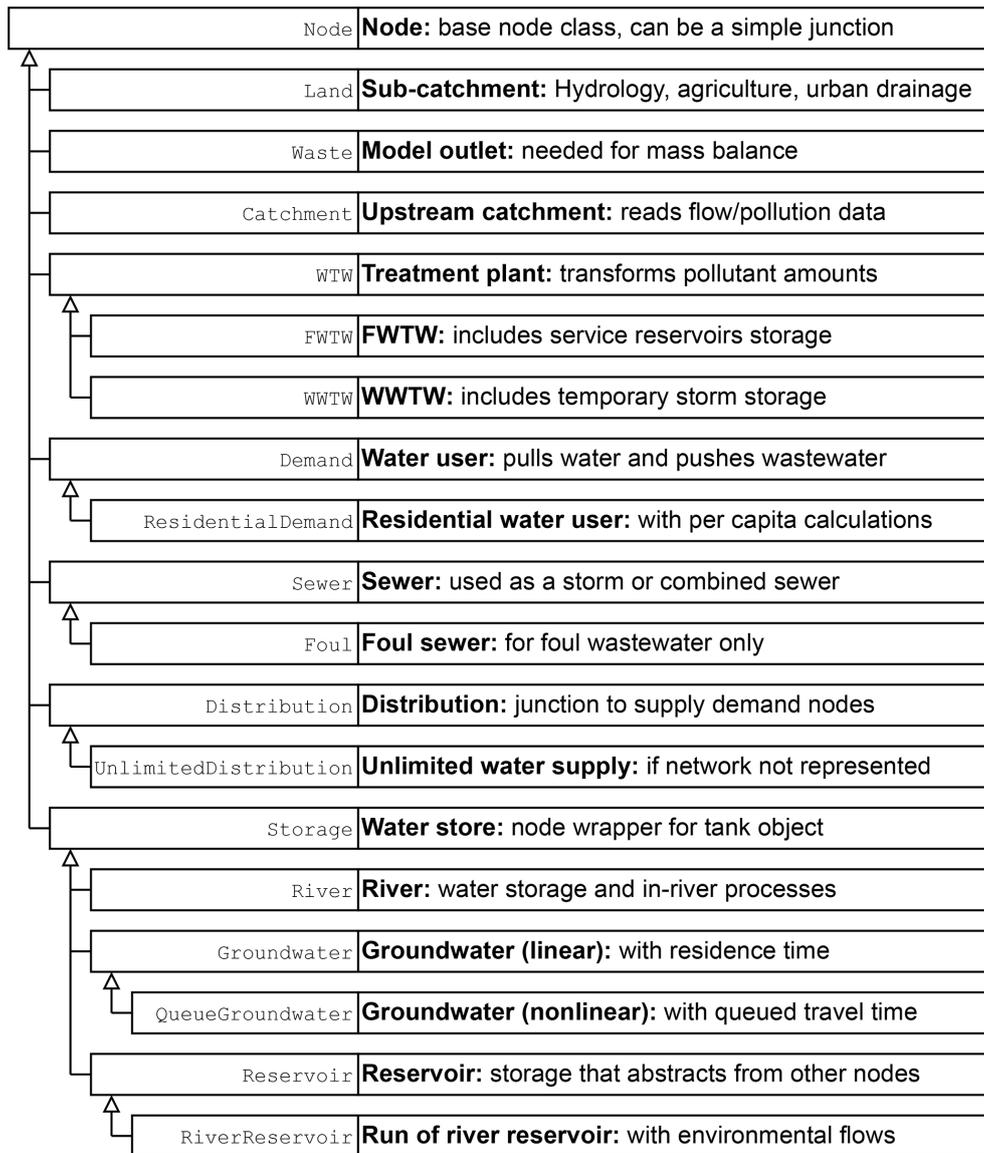


Figure 2. An inheritance diagram of components implemented as nodes in WSIMOD; arrows indicate that a node is a subclass of another node. The courier text is the name of the node in WSIMOD.

For these reasons, physical components should generally be implemented in WSIMOD as a subclass or variation of a Node, even if the computational implementation is simply a wrapper for another, pre-existing, model. Although the base Node class does not implement any physical processes, it may serve as a junction for branches or convergences in the water system, such as river bifurcations and confluences.

2.1.2 Node subclasses

To ensure WSIMOD is as easy to be implemented as possible, a variety of water cycle components have been developed in Python as Node subclasses. We summarise these components in Fig. 2 and the “Component Library” section

in the online documentation; for full details, we recommend viewing the API reference and “Key assumptions” section of Node subclasses in the online documentation, as well as various tutorials on their use (WSIMOD documentation, 2024). As anticipated in the introduction, these components are designed for parsimony and can be instantiated with as few parameters as possible, thus minimising data requirements and maximising utility. We note that this summary is up to date as of time of writing; however, the Imperial College London Water Systems Integration research group will be continually upgrading and adding new functionality to WSIMOD.

Many node types include water stores, which are sufficiently prevalent in water systems to warrant their own class in WSIMOD, referred to as a Tank, with further details pro-

vided in Sect. A1, “Tank object to generalise water stores”. It is important to note that a tank is not a node, but a node may have a tank. The number of tanks a node subclass may have varies depending on what it represents, with some having none (e.g. Demand), one (e.g. Reservoir), or multiple tanks (e.g. Land). If a user requires a tank that can act as a node, they can utilise the Storage class to achieve this functionality.

2.1.3 Customising nodes

We do not expect that the provided WSIMOD components will be sufficient to cover every water system or be suitable in cases where detailed representations are necessary. Where possible, users should consider customising model orchestration (Sect. 2.3) or component interactions (Sect. 2.2.3) to represent these cases. However, new physical subsystems may also be represented based on those defined in Fig. 2. Describing the variety of techniques to customise subclass behaviour in OOP is outside the scope of this paper, but they have been discussed extensively elsewhere (Gamma et al., 1995). In general, these techniques aim to avoid duplication of effort, because most functionality will be predefined in an existing class that can maintain specific interactions between existing subclasses. We enthusiastically encourage users to contribute to the open-source package to create a better software for the wider environmental modelling community, though we request that contributors read our online guidelines before doing so to avoid highly duplicative and bloated components that are likely to confuse others. If a pre-existing model aims to be integrated and it would be too significant a programming effort to re-implement under this philosophy, then we would request that wrappers are used to treat the model as a node and interface with the existing software.

To highlight the flexibility offered by WSIMOD, we briefly discuss some examples of complex behaviour being captured through node customisation. In Dobson et al. (2021), demand nodes were assigned time-varying population, water use behaviour, and pollutant generation to capture changing commuter patterns in London that resulted from the COVID-19 pandemic. In Liu et al. (2023a), the hydrological processes in land nodes were customised in a variety of ways to represent nature-based solutions. We note that these examples are provided with open-source code but are not included in the WSIMOD repository; they have significant data requirements to set up that are context specific and thus are unlikely to be generalisable to a wide range of cases. In the software documentation, we also provide a how-to guide for node customisation (WSIMOD documentation, 2024).

2.2 Integration framework

The WSIMOD integration framework facilitates interactions between nodes by serving as a message-passing interface that transfers information relating to water quality and quantity. It was developed in an application to London’s wa-

ter cycle at a wastewater catchment scale (Dobson et al., 2021). It draws significantly on the OpenMI (Harpham et al., 2019) and Open MPI (Graham et al., 2006) interfaces but has been tailored to water systems by providing a variety of built-in behaviours. The integration framework consists of three key concepts: arcs, pushes and pulls, and requests and checks (demonstrated in Fig. 3 and described in the following sections). Arcs are a class that facilitate interactions between nodes but can also represent physical entities (e.g. pipes). Arcs convey both water quality and quantity fluxes, which are discretised and packaged together, based on concepts from CityDrain3 (Burger et al., 2016). Pushes and pulls differentiate between the directionality of an interaction. Requests and checks differentiate between information passing that simulates the movement of water (requests) and that which does not (checks). While we do not recommend changing the integration framework itself, we provide a generic method to accommodate a wider variety of interactions in Sect. 2.2.3.

2.2.1 The Arc class and fluxes

Arcs are a class to establish connections between nodes. They transmit all passing messages and track fluxes when requests are made. Arcs can have a capacity property that limits the flow in a given time step, which can also be customised to be dynamically calculated, e.g. to implement Manning’s equation along pipes (Dobson et al., 2022b). Additionally, if multiple arcs are linked to a single node, they can be assigned a preference attribute. This enables the node to prioritise certain arcs over others. For example, a sewer node may connect to a wastewater treatment plant (WWTW) and to a river via a sewer spill. In this case, the spill arc could be assigned a low preference to ensure that it is only utilised if the WWTW cannot accept any water.

Fluxes in WSIMOD are described in discrete packages called volume-quality information packages (VQIPs). A VQIP is a dictionary that contains entries for volume and all simulated pollutants. Calculations in WSIMOD are typically performed on VQIPs rather than simply flows, thus ensuring simulation of both water quantity and quality. The core parent class (WSIObj) of all WSIMOD classes provides functions to perform basic operations with VQIPs in place of the normal arithmetical operations that would typically be only performed on flows or volumes. VQIPs track water quality as mass rather than as concentration to accommodate cases where pollutants are being moved with no associated water quantity, except for non-additive variables such as temperature or pH. WSIMOD refers to anything tracked in a VQIP that is not water volume as a pollutant; however, this should be interpreted as a water quality constituent and does not imply that everything simulated (e.g. temperature or dissolved oxygen) is a pollutant from an environmental/management perspective.

Table 1. The default reply from each kind of component interaction with an example.

		Push		Pull	
	Reply	Example	Reply	Example	
Request	Amount not received	A sewer sends a push request to a WWTW. The WWTW calculates the available capacity, updates its state variables to represent the increased throughput, and replies with how much water from the request that could not be received.	Amount sent	A FWTW sends a pull request to a reservoir. The reservoir calculates how much of the request can be met, updates its state variables to decrease the current volume, and replies with the amount of water abstracted.	
Check	Maximum volume available to push	A sewer sends two push checks to two downstream sewers (required to calculate what proportion to discharge water to them). The two sewers reply with total amount of water that they can each receive.	Maximum volume available to pull	A FWTW sends two pull checks to two reservoirs (required to determine what proportion to pull water from them). The reservoirs reply with each of their current abstractable volumes.	

whether they represent flux or not. Directionality of intent refers to either cases when a node has water that must be sent somewhere, called a push, or when a node needs water from somewhere, called a pull. Interactions between nodes, whether pushes or pulls, may convey flux of water and pollutants and simulate the movement of water, or they may convey non-flux information necessary for achieving realistic simulations. An interaction conveying flux is referred to as a request, while a non-flux interaction is a check.

Pushes occur when a node needs to discharge water. For example, when a wastewater treatment works (WWTW) must discharge effluent to a river or a hydrological catchment must discharge runoff downstream. In general, push scenarios are more common in water systems because water travels from upstream to downstream. Meanwhile, pulls occur when a node requires water. For example, when a farmer pumps water from a borehole to fill their irrigation reservoir. Pulls typically represent human-related effort to move water in a non-natural way.

A request occurs when a node intends to push or pull a certain amount of water to or from another node, regardless of the connected node's current state. For example, in the pull case, a demand node will intend to satisfy its entire water needs by pulling water from the distribution network, and this intention is independent of the availability of water in the distribution network. For example, in the push case, a demand node will always intend to send the entire volume of its foul water to a sewer system, even if the sewer cannot accommodate the full amount. All water flux in WSIMOD is ultimately simulated by requests; however, most cases are not as straightforward as the above examples, and nodes often require additional information about the state of giving or receiving nodes to calculate their requests. Interactions passing such non-flux information are referred to as checks.

Generically a check is any kind of non-flux information passing between nodes; it enables a node to use the state of the nodes that it interacts with to calculate its requests. For example, when a freshwater treatment works (FWTW) can draw water from multiple viable reservoirs collectively containing more water than needs to be treated, a calculation is required before requesting water. The FWTW will send checks to the connected reservoirs to determine their available water capacity and calculate the appropriate ratio to satisfy its treatment demand. Only then will the FWTW send requests for the required water.

2.2.3 Default and customised interactions

During simulation, a node needs to make responses when it receives a push/pull request/check, which we term as a reply. We formulate four types of predefined replies that are widely observed in node interactions in the water systems. These default replies allow nodes to interpret the responses of requests/checks sent to the interacting nodes without the need to understand their detailed behaviour. These defaults are set out in Table 1 with examples.

While these default interactions are likely to accommodate much of the information passing required to simulate an integrated water cycle, further customisation may be necessary to allow specific nodes to respond differently to others. For example, a sewer node might respond differently to a push request from another sewer than from a land node because sewer-to-sewer travel time may be calculated differently than land-runoff-to-sewer travel time. Furthermore, the default check behaviour which transmits capacity information may not be sufficient for a component to calculate where to push/pull water. For example, due to the importance of head in determining flow, the amount of water that a

floodplain can discharge would not be based on the receiving river's capacity but instead its head (Liu et al., 2023a).

To define what a node does in reply to an interaction, all WSIMOD interactions pass through “handlers” that are associated with a “tag”. A handler is a Python dictionary belonging to a node, and the tag is a key to the handler that determines which function is called during a given interaction (see Fig. 3). All nodes must have handlers which contain a “default” tag, thus enabling any node to interact with any other node. However, additional tags and replies can be added to enable different behaviours based on the type of node that is interacting with it. An example interaction for a Demand node draining to a Foul sewer node is given in Fig. 3. The generation of wastewater in the Demand node is triggered by orchestration and sent as a push request, via an arc, to a sewer node, with a tag “Demand” to indicate how the sewer node should reply. The sewer node has predefined behaviours for tags: “default”, “Sewer”, “Demand”, and “Land”, with different functions associated with these tags. The sewer node uses the handler to identify that it should evaluate the “push_set_land” function (which calculates the travel time through the sewer using the same equation as is used for calculating travel time of runoff arriving from a land node) and returns the reply via the arc. The base node predefines simple default handlers that enable it to convey interactions to connected nodes, e.g. a push request to a node will trigger push requests to connected nodes; in this way, the node can behave as a simple junction. A how-to guide to explain this behaviour in greater detail, with examples, is available in the documentation (WSIMOD documentation, 2024).

2.3 Orchestration to manage and enable simulation

As defined by Belete et al. (2017), orchestration is how nodes and their interactions are managed and enabled. The default steps that we consider to be important to include in orchestration for many water systems are shown in Fig. 1. In integration frameworks such as OpenMI, all interactions in a simulation occur by nodes triggering pulls to other nodes, orchestrated by a single external pull each time step (Harpham et al., 2019). Meanwhile, in WSIMOD, a finer level of control of orchestration is given to a user. We argue that the WSIMOD approach is better at capturing complex within-time-step behaviours.

In Fig. 4, building on the system shown in Fig. 1, we provide a motivating example of how it is more efficient to customise the behaviour of water systems by customising orchestration, rather than customising/adding nodes or arcs, as would be required in other integration frameworks. We consider a common case for an integrated water system of requiring the simulation to perform downstream re-abstractation of wastewater effluent, visualised by the red arc in Fig. 4. Wastewater re-abstractation (water from W reaching A) requires available wastewater at W to be matched against de-

mand for re-abstracted wastewater at A, with both calculations involving the interconnected node, J.

To capture such an interaction without customising orchestration and using only pulls, as would be required with an OpenMI approach (Fig. 4b), customisation of J would need to specify whether the pull is originating from A or O, and this information would need to be conveyed to W. If the pull is intended to route the system for that time step (i.e. it is originating from O via J), then W should release all its effluent, while if it is intending to satisfy the demand (i.e. it is originating from A via J), it should only release enough to meet A's requirements. Additionally, the external pull that initially triggers all interactions must be customised to pull from A before O so that W does not fully route before abstractions can be made.

Instead, a user with a broad overview of the water cycle may more easily accommodate such behaviour if given high-level control over the model orchestration (Fig. 4c). During a time step, the orchestration can specify that W first calculates its treated effluent without discharging it into a receiving river (abstractions are triggered at A), which draws water from W via J, and then W discharges any remaining effluent into its receiving water. This flexibility in orchestration enables representations of a wide variety of water systems while minimising changes to the behaviour of underlying components.

2.3.1 The model class

A “Model” class contains all nodes, arcs, and forcing data and provides a default orchestration that we expect can represent a wide variety of water systems. Whether the WSIMOD default orchestration is used or whether an entirely new orchestration is defined, we recommend the use of a model class for a variety of reasons. Firstly, built-in load and save functionality enables easy sharing and editing of a specific model. Secondly, it contains a run function that can carry out orchestration, perform mass balance checking (see following section), store simulation results, and trigger end-of-time-step functions. Thirdly, it enables easy collection and grouping of nodes to facilitate orchestration, e.g. all WWTWs are referenced by dictionary belonging to the model class so that they can easily be triggered during a time step. A tutorial on the use of the model class is provided in the online documentation (WSIMOD documentation, 2024).

2.3.2 Software quality control

Due to the integrated nature of the water systems that WSIMOD simulates, it can be easy to introduce errors that are difficult to spot. We provide extensive unit testing in line with best software development practice, which enables ensuring that changes to code do not introduce unintended behaviour changes. However, a further safeguard against this is a unified method for mass balance error checking. Both nodes and

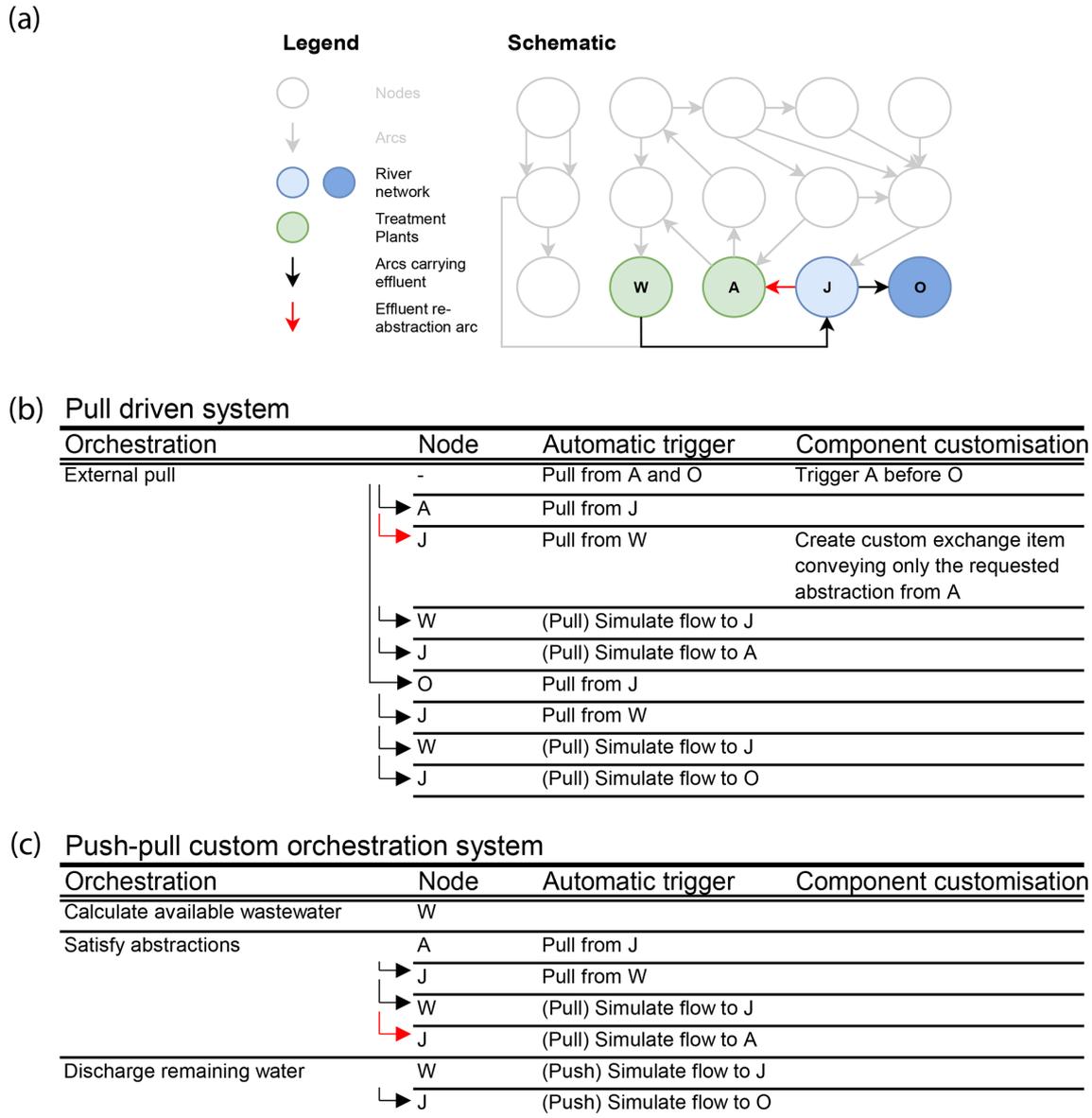


Figure 4. (a) An example water system based on the Oxford example from Fig. 1 that includes re-abstraction of treated effluent (red arc). (b) The model steps to achieve re-abstraction of effluent when using a pull-driven integration framework, of the kind used in OpenMI, without customisable orchestration and (c) when using a push-driven integration framework, of the kind used in WSIMOD, with customisable orchestration.

arcs have predefined lists of functions to calculate the total inflows, total outflows, and change in any storage. Any newly defined behaviour for nodes and arcs must then also consider how mass balance checks will be impacted and thus update the lists of functions associated with inflows/outflows/change in storage.

Because the core functionality of WSIMOD that performs mass balance checking is indifferent to units (see Sect. 2.3.3) and because some pollutants exist in far smaller quantities than others, it is possible that both incredibly small and large numbers may be present in VQIPs in the model. Therefore,

mass balance checking compares at the magnitude of the largest value of inflows/outflows/storage for a given pollutant or water volume for a given model element in each time step. Any discrepancy that is larger than a user-specified value is reported, although because of floating-point accuracy some discrepancies will be unavoidable. We encourage users to exercise common sense to not chase down incredibly small discrepancies while seeking to understand larger discrepancies, which are usually indicative of some implementation error. The user control over orchestration typically makes debugging WSIMOD models easier than most integrated models,

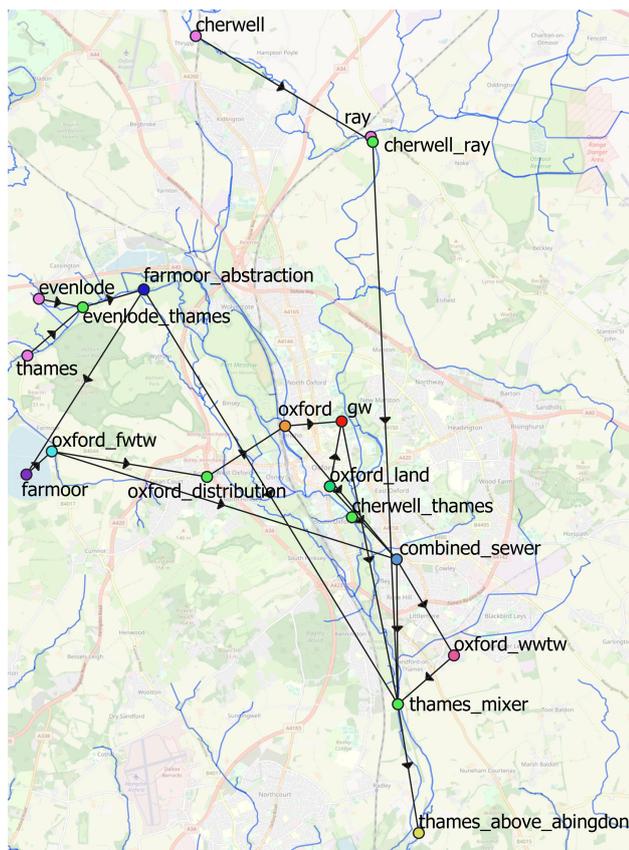


Figure 5. Map of the model nodes and arcs over the city of Oxford; map base layer ©OpenStreetMap contributors 2024. Distributed under the Open Data Commons Open Database License (ODbL) v1.0.

as a user may step through each set of triggers in the orchestration and recheck mass balances until the error occurs.

2.3.3 Units and time steps

While the core of WSIMOD is agnostic of units, many nodes are parameterised and assume input data are in SI units; thus, we recommend use of SI units throughout. If extensive work would be required to re-implement equations in new units, we recommend converting before and after the calculations from/to SI units.

Because the time step size will vary depending on the application of WSIMOD, nodes do not necessarily make assumptions about time step, instead requiring the user to define the time step that is consistent with their parameters and input data. This enables significant flexibility in representations that can enable studies investigating the impact of time step size on simulations (Dobson et al., 2022b). We note that the detailed biochemical processes used in agricultural surfaces and rivers assume a daily time step. In addition, these processes are developed to focus on pollutants associated with crop nutrient cycles.

3 Demonstration

Throughout this paper, numerous WSIMOD case studies are referenced. However, to illustrate how WSIMOD works, we present a demonstration case study. We have chosen a model featured as a tutorial in the online documentation so that readers will be able to easily reproduce, run, and edit it (https://imperialcollegelondon.github.io/wsi/demo/scripts/oxford_demo/, last access: 8 April 2024). The demonstration covers the area of Oxford, UK, depicted in Fig. 5 as a map and Fig. 1 as a schematic. We highlight that this simple demonstration and online tutorial is included because of its usefulness in explaining the underlying functionality and flexibility of WSIMOD, rather than for motivating the complexity and importance of integrated modelling in general.

The model includes four upstream hydrological catchments, a reservoir-based water supply system, and a combined sewerage system, all of which ultimately combine and drain into the river Thames. Because the model is primarily demonstrative, the parameters given in the model are estimated based on local knowledge. However, high-resolution (weekly sampling) water quality data are available in the area (Bowes et al., 2018), thus enabling accurate boundary conditions at upstream hydrological catchments and water quality representations in the model. In Fig. 6, we plot simulations against the weekly sampling; the water quality indicators to plot were selected based on those that had data for all locations for the simulation duration.

In addition to WSIMOD's capabilities as a general-purpose simulator of hydrological and water quality, it is also a valuable tool for management and interventions. The software's design prioritises customisation, making it easy to incorporate specific operational preferences which are not typical to capture due to the limited modelling scope of most water simulators. For example, in the online customisation guides (<https://imperialcollegelondon.github.io/wsi/how-to/>, last access: 8 April 2024), we show how the behaviour of nodes and arcs can be altered to accommodate changing abstraction licensing and environmental flow requirements for the Oxford case study, with results shown in Fig. 7. Liu et al. (2023b) demonstrate how a user can implement highly sophisticated water quality management strategies in WSIMOD based on pollution load allocation.

4 Discussion

4.1 Why are integrated water systems models necessary?

We defined four key goals for integrated models of water systems. In this section we will discuss how WSIMOD helps to meet these goals.

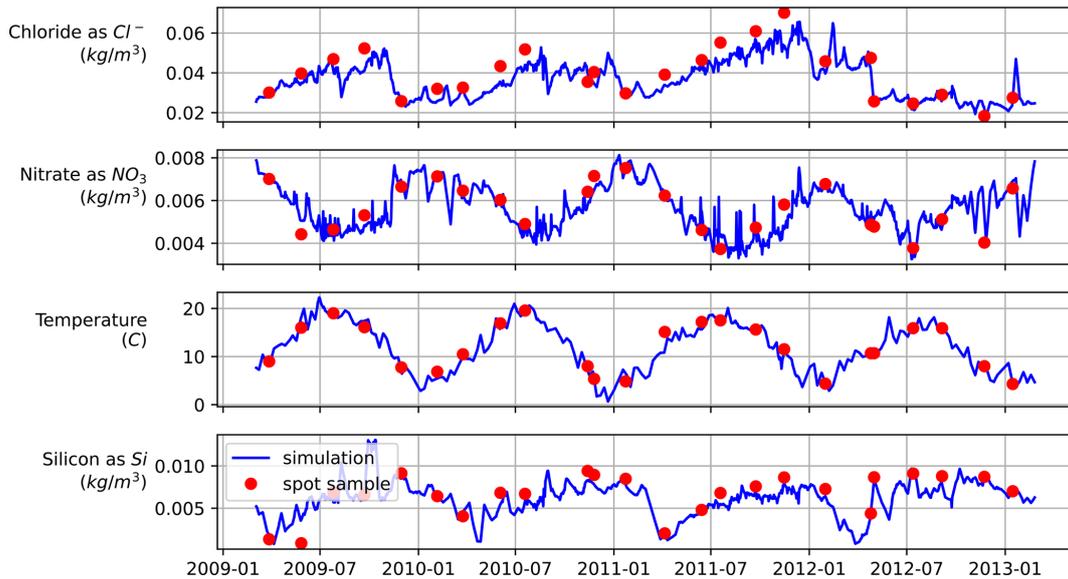


Figure 6. Demonstration of simulated levels of chemical water quality indicators against spot samples.

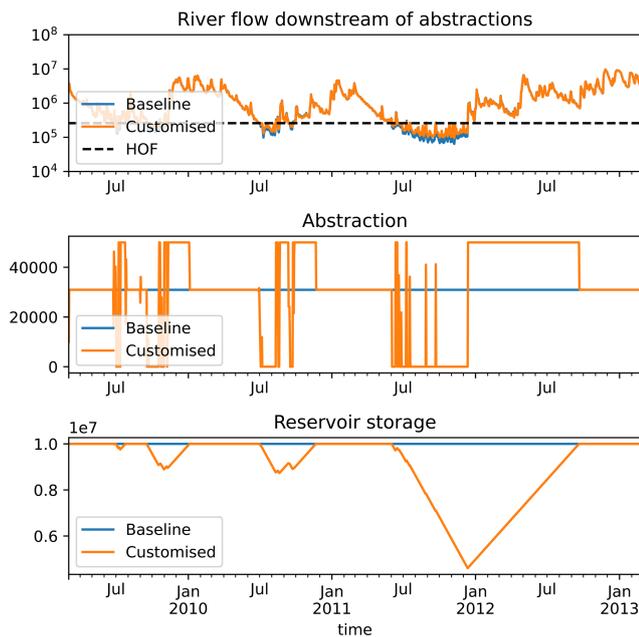


Figure 7. Example of implementing a “hands-off flow” (HOF), which abstractions cannot draw water below, for the reservoir abstraction node in the Oxford case study model. The “customised” simulations are with HOF implemented, and “baseline” simulations are without.

4.1.1 Integration to understand emergent behaviour in water systems

Emergent behaviours in human and environmental systems arise from the interactions between multiple components and are difficult to predict and understand without addressing

the complexity that occurs from this interconnectedness (Liu et al., 2007). The human-altered water cycle covers a diverse set of components and so any attempt to understand the fundamental (physical and operational) drivers behind emergent behaviours must acknowledge the interactions and feedbacks between humans and hydrological processes (Wada et al., 2017). WSIMOD represents many components within the water cycle needed to capture these interactions to reveal and quantify fundamental processes, e.g the fact that in-river water quality during wetter periods is driven by agricultural processes and during drier periods by urban processes (Liu et al., 2022). Without integrated modelling, we risk oversimplifying the system and omitting feedbacks that could have significant implications for water management decisions (Dobson and Mijic, 2020).

Understanding and reducing the uncertainty behind behaviour in hydrological systems require the intercomparison of hydrological process representations and thus a flexible modelling framework (Knoben et al., 2019). We suggest the same is true for the wider water cycle, and so we anticipate that the flexible approach to integrated modelling in WSIMOD is well suited for these purposes, e.g by comparing different assumptions around water consuming behaviour (Dobson et al., 2021). Furthermore, customisability enables accommodating unconventional water systems that may stress the assumptions of underlying process representations. This approach can help identify any weaknesses or gaps in the model and refine our understanding of both the process in question and the behaviour of the wider system. For example, the poor simulations of hydraulic structures identified in Dobson et al. (2022a) are a result of reduced accuracy in capturing water head, in turn caused by the discrete modelling scheme inherent to a non-tightly coupled integrated

framework. Identifying this weakness has implications beyond simulating hydraulic structures and provides guidance for future work.

Finally, research that contributes to knowledge must be reproducible; if it is not, then it can hardly be considered science (Hutton et al., 2016). To ensure that WSIMOD applications are reproducible, the software is open source with a permissive licence, and it is provided with significant documentation and worked examples to ensure that it is used as intended. This documentation transparently lists the assumptions made for each model component, both in the source code of that component and in a self-contained library page (WSIMOD documentation, 2024). Furthermore, the ability to save the model class in a self-contained and human-readable file enables publishing WSIMOD applications in an easy-to-reproduce way.

4.1.2 Integration to expand model boundaries

The WSIMOD modelling framework offers a versatile approach to representing various processes that are commonly treated as boundary conditions in water cycle models. One example is the agricultural–hydrological Land node, which utilises the CatchWat model to estimate the pollution and flow in rivers upstream of the model region (Liu et al., 2022, 2023a). Accounting for upstream pollution and dilution is crucial for accurately assessing the impact of urban processes on in-river conditions. By incorporating the Land node, WSIMOD provides a means to capture the typically neglected upstream agricultural system in urban wastewater studies, which is often viewed as a boundary. WSIMOD also facilitates the representation of boundary conditions within urban systems. For instance, it allows for the inclusion of abstractions, which are critical for understanding in-river impacts of wastewater effluent because of their influence on dilution during low flows (Dobson and Mijic, 2020).

Besides achieving accurate simulations, we further propose that a modelled representation of boundary conditions is essential towards assessment of the system under future scenarios. For example, by treating the upstream river as a model rather than as a fixed boundary condition, the in-river impact of the wastewater system for future scenarios can be quantified, which would not be possible if the changes to upstream river behaviour under the future scenario were fixed (Mijic et al., 2022).

4.1.3 Integration to evaluate intervention impacts at whole-water-system scale

The OOP deployed in WSIMOD enables flexibility to incorporate a variety of different physical and management interventions to the water cycle. As demonstrated in Liu et al. (2023a), the included WSIMOD nodes may be customised both in terms of their parameters and in terms of their physical processes to represent a range of nature-based solutions

such as flood plains, runoff attenuation features, regenerative farming, urban green space, and urban wetlands. Because all WSIMOD nodes can communicate with all other nodes, this is a reasonably straightforward exercise as new rules for interactions with other model components do not need to be re-defined when additional components are added to the model.

Due to the flexibility and coverage provided by WSIMOD models, it is often straightforward to include decision-making, policy constraints, and operational rules (Dobson and Mijic, 2020). The high-level control over orchestration provides an ideal place to represent these stakeholder actions that require information on a multitude of states across the wider system. For example, the amount available to be abstracted from the river Thames, UK, is based off of both reservoir levels and river flow, meaning no specific node has access to all of the information required to determine abstraction amount (Lower Thames Operating Agreement, 28/39/M/2). Rather than creating a complicated interaction between London reservoirs and the river Thames, a modeller may more simply inspect the state of these two systems in the orchestration to dynamically set the abstraction capacity, filling the role of an operator (Dobson and Mijic, 2020). Thus, the nationally important question of modelling changes to abstraction policies may easily be implemented in WSIMOD, and the impacts of these, both on water supply and in-river water quality, may be quantified (Mijic et al., 2022).

The parsimonious methods selected to represent physical components in WSIMOD ensure that simulations are computationally efficient. Further to this, if computational speed is of critical interest, then we have demonstrated that varying time step and spatial resolution is possible while still achieving accurate simulations (Dobson et al., 2022a). The result is an integrated water cycle model that can be used for purposes that are not typically computationally tractable. Dobson et al. (2022a) demonstrate that exploration of uncertainty in sewer model parameters (roughness and runoff coefficient) is possible with WSIMOD; this study is the first of its kind because pipe network models are typically too computationally expensive to perform the numerous simulations required for uncertainty analysis. Meanwhile, Liu et al. (2023a) demonstrate that regional portfolios of 5-year catchment-scale nature-based solutions can be created by applying optimisation to a WSIMOD model spanning 32 hydrological catchments.

4.1.4 Integration to align simulations with systems level outcomes

A key benefit we have found in applications of WSIMOD is that, due to the breadth of systems that are represented, performance metrics that are relevant to different stakeholders can usually be included in the model in a physically based way. The most obvious example, which we have drawn on throughout this paper, is the ability to place in-river impacts central to decision-making. This enables better alignment with policy goals, such as the Water Framework Direc-

tive chemical water quality classifications, which are defined based on in-river average pollutant concentration (Environment Agency, 2020a). However, a variety of other metrics can be conjunctively included such as the reliability of water supply. For example, Dobson and Mijic (2020) examine both the water quality and water supply benefits of a variety of water cycle interventions (leakage reduction, rainwater harvesting, etc.). Such an approach is particularly beneficial to stakeholder engagement, because non-water-facing stakeholders can still be represented and understand their interactions with the water cycle and system-wide goals. For example, Puchol-Salort et al. (2022) demonstrate how developers can measure the impacts of new developments on flooding, water quality, and water supply. Furthermore, this study demonstrates how integrated modelling can enable them to quantify how to ensure developments achieve a net-zero impact on these metrics, interestingly revealing that retrofitting households outside of the development area is typically required to offset changes.

4.2 Data to support water systems integration

A key challenge to the application of an integrated approach to water systems modelling is the difficulty of setting up models and associated availability of data. In WSIMOD, we provide parsimonious but physically based representations to ensure parameterisation is possible with widely available data, and we propose model evaluation with observed in-river flow and water quality. The catalogue of documented nodes in WSIMOD (WSIMOD documentation, 2024) presents data requirements to help users understand how feasible it will be to apply the approach in their study area. The data requirements will be entirely dependent on what nodes and at what resolution a model user chooses to represent. However, we provide a broad overview of data requirements for a generic catchment-scale WSIMOD case study in Table 2, which may help to develop future automatic model setup at national and global scales to facilitate applicability.

4.2.1 Future work and research direction

One of the main concerns that other modellers have expressed regarding a WSIMOD-like approach is the level of detail with which components are represented. While hydrologists and agricultural modellers are often comfortable with parsimony and aggregation in their catchment modelling, most other parts of the water cycle tend towards more complexity in their representations. This desire for complexity is likely due to the detailed application context in which different models have been developed. For instance, designing a new process in a wastewater treatment plant inevitably requires a highly detailed model (Hreiz et al., 2015). However, questions are being raised in many fields with a tradition of complex modelling about the need for such complexity. Models of in-river phosphorus (Jackson-Blake et al.,

2017), urban flooding (Li and Willems, 2020), and sewer flow (Thryssøe et al., 2019; Dobson et al., 2022a) have shown that good results can be achieved with simpler approaches. While practical modellers typically question what level of complexity is necessary to answer their questions, scientific modellers examine the impacts of assumptions to build evidence around whether they are suitable and under which circumstances. WSIMOD prioritises integration of the whole water cycle, which is enabled by reduced-complexity modelling of the system components. In the examples provided, we demonstrate that sacrificing complexity in terms of detail should be viewed as an opportunity to better accommodate and contextualise components in the wider water cycle, as well as highlighting the importance of interactions between components.

We see the WSIMOD platform as an ideal opportunity for the environmental modelling community to implement and compare (or benchmark) different modelling assumptions and examine water cycle impacts. We plan to continue developing the representation of different components and continue testing whether more complex representations can improve simulations. Our focus will be on representations of a wider range of nature-based solutions, treatment plants, and urban sewer network hydraulic structures. Furthermore, we believe that complementing WSIMOD with machine learning representations of components that are too complex to be captured in a physical way can be a promising approach, thus implementing a “surrogate” strategy (Razavi et al., 2012, 2022).

A key opportunity for improving the accessibility of WSIMOD will be in the development of a graphical user interface (GUI). The current implementation as a Python package makes the software well suited to customisable and flexible simulations for programmers but inaccessible to a wide range of potential users. We see a variety of different approaches towards greater interactivity and visualisation that are not mutually exclusive. A “virtual decision room” approach may provide an ideal environment for non-technical stakeholders to explore simulation results and to highlight integrated system-wide impacts (Schouten et al., 2016). Meanwhile, incorporation into GIS-based frameworks such as 3DNet (Todorović et al., 2019) or Google Earth Engine (Gorelick et al., 2017) would enable more seamless incorporation of pre-processing and provide a suite of streamlined tools to help users create, edit, and run WSIMOD models.

5 Conclusion

We have presented the theoretical underpinning of WSIMOD, which is an open-source software for simulating a range of urban and rural processes and operations in the integrated water cycle. WSIMOD represents different components of the water cycle as nodes that are connected by arcs. The nodes that we have discussed throughout the pa-

Table 2. List of datasets typically required in catchment-scale WSIMOD applications with references for data sources or further information.

Category	WSIMOD input	Availability	Further information (scale)
Climate	Evapotranspiration	Global datasets available	(Khan et al., 2018) (global)
	Temperature	Global datasets available	(Morice et al., 2021) (global)
	Precipitation	Global datasets available	(Sun et al., 2018) (global)
Rural	Hydrological catchment outlines	Global datasets available	(Lin et al., 2019) (global)
	Hydrological catchment connectivity	Global datasets available	(Lin et al., 2019) (global)
	Hydrological catchment parameters	Global datasets available for some hydrological models	(Zhang and Schaap, 2018) (global)
	Crop surfaces	Global datasets available	(Thenkabail et al., 2016) (global)
	Crop properties	Lookup tables available	(Allen et al., 1998) (–)
	Pollutants/nutrients	National datasets may be available with high uncertainty	(Liu et al., 2022) (UK)
Urban	Population	Global datasets available	(Leyk et al., 2019) (global)
	Garden area	National datasets may be available, otherwise rule of thumb may be acceptable	Office for National Statistics (2023) (UK)
	Wastewater treatment plants	European dataset available through the Urban Wastewater Treatment Directive	European Commission (Environment Agency, 2016) (Europe)
	Foul catchments	National datasets may be available; otherwise, contacting wastewater companies is required	(Hoffmann et al., 2022) (UK)
Water use	Irrigation water use	Global datasets available	(Thenkabail et al., 2009) (global)
	Water resources system	Not typically available, contacting water supply companies is required	–
Evaluation	Flow observations	National datasets typically available	(https://nrfa.ceh.ac.uk/ , last access: 24 May 2024) (UK)
	Water quality observations (river and WWTW)	National datasets may be available	(Environment Agency, 2020b; WIMS, 2020) (England)
	Reservoir levels	Not typically available, contacting water supply companies is required	–

per are parsimonious implementations that are conducive towards easy parameterisation and setup. Arcs convey interactions between nodes that fall under four key categories: pushes (a node has water to go somewhere), pulls (a node needs water from somewhere), requests (interaction represents flux of water and/or pollutants), and checks (interaction does not represent flux). This integration framework allows all nodes to communicate with each other, thus facilitating a flexible method that can accommodate a wide variety of water systems. Because this approach uses object-oriented programming, WSIMOD enables customisation to capture unconventional behaviours and implementation of a wide variety of physical and management interventions.

In summary, our early case studies show WSIMOD to be a useful and versatile tool for water systems modelling. We hope to have persuaded other modellers of the importance of an integrated approach and believe the design philosophy behind WSIMOD can serve as a helpful starting point for understanding integration in their respective contexts.

Appendix A: Technical details in WSIMOD

A1 Tank object to generalise water stores

A particularly important concept in developing new and understanding existing nodes is the Tank object. The concept of a water store is so common in water systems that a generic Tank object is provided. Tanks streamline a variety of uses for stores and have a range of child objects that implement travel time (Sect. A2) and pollutant decay (Sect. A3). A node that represents a water store should be a subclass of the Storage class (see Fig. 2), which itself is a generic node wrapper for the Tank object. The simplest case of a tank would be a water supply reservoir, demonstrated for WSIMOD at a lumped London scale in Dobson and Mijic (2020). However, many nodes use stores but in an auxiliary fashion; for example, WWTWs have temporary storage tanks, while FWTWs have service reservoir tanks.

A2 Non-flux pollutant changes

By default, everything tracked in a VQIP follows mass balance. However, in the water cycle, many pollutants undergo transformations due to biological, physical, or chemical processes; thus, preservation of mass may be insufficient to simulate them. WSIMOD represents the nitrogen/phosphorus cycles in soil (see documentation of Land nodes) and denitrification/mineralisation/production/macrophyte uptake in rivers (see documentation of River nodes), based on the equations from Liu et al. (2022) and Lindström et al. (2010).

While the transformations that act on chemicals in soils and rivers are well studied in the literature, there are difficulties in conceptualising biochemical transformations that take place in groundwater and sewers, despite agreement that these are chemically active (Almeida et al., 1999). As a result, WSIMOD provides a generic two-parameter method to implement temperature-sensitive chemical decay, given by

$$M_t = M_{t-1} \left(1 - cd^{T_t - T_{\text{ref}}} \right), \quad (\text{A1})$$

where M is the mass of a chemical in each time step, c is a parameter that determines non-temperature sensitive decay, d is a parameter that determines temperature sensitive decay, and T is the temperature with a reference temperature (T_{ref}) assumed to be 20 °C. We do not intend that Eq. (A1) can be a substitute for well-researched and verified process representations; however, in our experience of using WSIMOD it is an easy and useful option to improve water quality representations.

Wastewater and freshwater treatment processes are well-studied fields. However, simulation models of these systems require detailed information describing the different treatment technologies and processes that are present in a specific plant. While we plan to include these types of models in WSIMOD in the future, we have opted to take a parsimonious approach to treatment modelling under the assumption that most users will not have detailed information about the plants they model. This approach assumes that the plant performs a single operation, based on Eq. (A1), to transform influent, which is then split into three streams of effluent, liquor, and solids. Depending on whether it is freshwater or wastewater treatment, these streams go to different places; see the documentation of WTW for further details.

A3 Travel time of water

Arcs are the key model element to implement travel time of water. Two arc subclasses that provide alternate methods to implement travel time are provided in WSIMOD. The first, more simple approach, formulates the travel time of the arc as a dictionary object where each key is the number of time steps remaining; when water is sent along the arc, it is combined with the any existing water for the key that matches the specified travel time. These travel times are

updated at the end of each time step. This method is computationally efficient, because the number of operations each time step is limited by the maximum number of time steps the arc takes to traverse. However, this approach cannot represent a dynamic flow capacity, as is the case in, for example, sewer networks, where hydraulic head governs flow. Thus WSIMOD also contains a less computationally efficient arc to accommodate this behaviour, described and demonstrated in Dobson et al. (2022b). Arcs can also implement pollutant changes associated with decay over this travel time, using Eq. (A1).

Code availability. WSIMOD is provided open source under the terms of the BSD 3-Clause license. The code can be accessed at <https://github.com/imperialcollegelondon/wsi> (last access: 26 March 2024) and <https://doi.org/10.5281/zenodo.7662569> (Dobson et al., 2023b), and documentation is at <https://imperialcollegelondon.github.io/wsi/> (last access: 26 March 2024), with further technical details in Appendix A. The code has been tested up to Python 3.10 and requires minimal dependencies (see website).

Author contributions. BD and LL created and tested all model code and documentation. All authors were involved in theoretical development and writing.

Competing interests. The contact author has declared that none of the authors has any competing interests.

Disclaimer. Publisher's note: Copernicus Publications remains neutral with regard to jurisdictional claims made in the text, published maps, institutional affiliations, or any other geographical representation in this paper. While Copernicus Publications makes every effort to include appropriate place names, the final responsibility lies with the authors.

Acknowledgements. This work was funded by the NERC CAMEL-LIA project (Community Water Management for a Liveable London), grant NE/S003495/1. Leyang Liu is funded by the president's PhD scholarships provided by Imperial College London. The views expressed in this paper are those of the authors alone and not the organisations for which they work. We are grateful to Liliane Manny and Liu Bo for their insightful comments on the manuscript that have improved the paper.

Financial support. This work was funded by the NERC CAMEL-LIA project (Community Water Management for a Liveable London) (grant no. NE/S003495/1). Leyang Liu is funded by the president's PhD scholarships provided by Imperial College London.

Review statement. This paper was edited by Taesam Lee and reviewed by Hongyi Li and one anonymous referee.

References

- Allen, R. G., Pereira, L. S., Raes, D., and Smith, M.: Crop evapotranspiration-Guidelines for computing crop water requirements-FAO Irrigation and drainage paper 56, Fao, Rome, 300, D05109, 1998.
- Almeida, M. C., Butler, D., and Friedler, E.: At-source domestic wastewater quality, *Urban Water*, 1, 49–55, [https://doi.org/10.1016/s1462-0758\(99\)00008-4](https://doi.org/10.1016/s1462-0758(99)00008-4), 1999.
- Bach, P. M., Rauch, W., Mikkelsen, P. S., McCarthy, D. T., and Deletic, A.: A critical review of integrated urban water modelling – Urban drainage and beyond, *Environ. Model. Softw.*, 54, 88–107, <https://doi.org/10.1016/j.envsoft.2013.12.018>, 2014.
- Belete, G. F., Voinov, A., and Laniak, G. F.: An overview of the model integration process: From pre-integration assessment to testing, *Environ. Model. Softw.*, 87, 49–63, <https://doi.org/10.1016/j.envsoft.2016.10.013>, 2017.
- Best, M. J., Pryor, M., Clark, D. B., Rooney, G. G., Essery, R. L. H., Ménard, C. B., Edwards, J. M., Hendry, M. A., Porson, A., Gedney, N., Mercado, L. M., Sitch, S., Blyth, E., Boucher, O., Cox, P. M., Grimmond, C. S. B., and Harding, R. J.: The Joint UK Land Environment Simulator (JULES), model description – Part 1: Energy and water fluxes, *Geosci. Model Dev.*, 4, 677–699, <https://doi.org/10.5194/gmd-4-677-2011>, 2011.
- Beven, K.: Towards integrated environmental models of everywhere: uncertainty, data and modelling as a learning process, *Hydrol. Earth Syst. Sci.*, 11, 460–467, <https://doi.org/10.5194/hess-11-460-2007>, 2007.
- Blair, G. S., Beven, K., Lamb, R., Bassett, R., Cauwenberghs, K., Hankin, B., Dean, G., Hunter, N., Edwards, L., Nundloll, V., Samreen, F., Simm, W., and Towe, R.: Models of everywhere revisited: A technological perspective, *Environ. Model. Softw.*, 122, 104521, <https://doi.org/10.1016/j.envsoft.2019.104521>, 2019.
- Bowes, M. J., Armstrong, L. K., Harman, S. A., Wickham, H. D., Nicholls, D. J. E., Scarlett, P. M., Roberts, C., Jarvie, H. P., Old, G. H., Gozzard, E., Bachiller-Jareno, N., and Read, D. S.: Weekly water quality monitoring data for the River Thames (UK) and its major tributaries (2009–2013): the Thames Initiative research platform, *Earth Syst. Sci. Data*, 10, 1637–1653, <https://doi.org/10.5194/essd-10-1637-2018>, 2018.
- Burger, G., Bach, P. M., Urich, C., Leonhardt, G., Kleidorfer, M., and Rauch, W.: Designing and implementing a multi-core capable integrated urban drainage modelling Toolkit: Lessons from CityDrain3, *Adv. Eng. Softw.*, 100, 277–289, <https://doi.org/10.1016/j.advengsoft.2016.08.004>, 2016.
- Dobson, B. and Mijic, A.: Protecting rivers by integrating supply-wastewater infrastructure planning and coordinating operational decisions, *Environ. Res. Lett.*, 15, 114025, <https://doi.org/10.1088/1748-9326/abb050>, 2020.
- Dobson, B., Wagener, T., and Pianosi, F.: How important are model structural and contextual uncertainties when estimating the optimized performance of water resource systems?, *Water Resour. Res.*, 55, 1–24, <https://doi.org/10.1029/2018WR024249>, 2019.
- Dobson, B., Jovanovic, T., Chen, Y., Paschalis, A., Butler, A., and Mijic, A.: Integrated modelling to support analysis of COVID-19 impacts on London’s water system and in-river water quality, *Front. Water*, 3, 26, <https://doi.org/10.3389/frwa.2021.641462>, 2021.
- Dobson, B., Barry, S., Maes-Prior, R., Mijic, A., Woodward, G., and Pearse, W. D.: Predicting catchment suitability for biodiversity at national scales, *Water Res.*, 221, 118764, <https://doi.org/10.1016/j.watres.2022.118764>, 2022a.
- Dobson, B., Watson-Hill, H., Muhandes, S., Borup, M., and Mijic, A.: A Reduced Complexity Model With Graph Partitioning for Rapid Hydraulic Assessment of Sewer Networks, *Water Resour. Res.*, 58, e2021WR030778, <https://doi.org/10.1029/2021WR030778>, 2022b.
- Dobson, B., Liu, L., and Mijic, A.: Water Systems Integrated Modelling framework, WSIMOD: A Python package for integrated modelling of water quality and quantity across the water cycle, *J. Open Source Softw.*, 8, 4996, <https://doi.org/10.21105/joss.04996>, 2023a.
- Dobson, B., Liu, L., and Mijic, A.: Water Systems Integrated Modelling framework, WSIMOD: A Python package for integrated modelling of water quality and quantity across the water cycle (0.3_joss_reviewed), Zenodo [code], <https://doi.org/10.5281/zenodo.7662569>, 2023b.
- Douglas-Mankin, K. R., Srinivasan, R., and Arnold, J. G.: Soil and Water Assessment Tool (SWAT) Model: Current Developments and Applications, *T. ASABE*, 53, 1423–1431, <https://doi.org/10.13031/2013.34915>, 2010.
- Environment Agency: Urban Wastewater Treatment Directive – Treatment Plants: Environment Agency, Reading, UK, <https://www.data.gov.uk/dataset/0f76a1c3-1368-476b-a4df-7ef32bfd9a8b/urban-waste-water-treatment-directive-treatment-plants> (last access: 24 May 2024), 2016.
- Environment Agency: WFD Classification Status Cycle 2, Environment Agency, Reading, UK, <https://www.data.gov.uk/dataset/41cb73a1-91b7-4a36-80f4-b4c6e102651a/wfd-classification-status-cycle-2> (last access: 24 May 2024), 2020a.
- Environment Agency: Open water quality archive datasets (WIMS), Environment Agency, Reading, UK, <https://environment.data.gov.uk/water-quality/view/download>, last access: 19 March 2020b.
- Gamma, E., Johnson, R., Helm, R., Johnson, R. E., and Vlissides, J.: Design patterns: elements of reusable object-oriented software, Pearson Education, Boston, USA, ISBN 978-0201633610, 1995.
- Giakoumis, T. and Voulvoulis, N.: Combined sewer overflows: relating event duration monitoring data to wastewater systems’ capacity in England, *Environ. Sci. Water Res. Technol.*, 9, 707–722, <https://doi.org/10.1039/D2EW00637E>, 2023.
- Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., and Moore, R.: Google Earth Engine: Planetary-scale geospatial analysis for everyone, *Remote Sens. Environ.*, 202, 18–27, <https://doi.org/10.1016/j.rse.2017.06.031>, 2017.
- Graham, R., Shipman, G., Barrett, B., Castain, R., Bosilca, G., and Lumsdaine, A.: Open MPI: A High-Performance, Heterogeneous MPI, in: 2006 IEEE International Conference on Cluster Computing, 25 September 2006 to 28 September 2006, Barcelona, 1–9, <https://doi.org/10.1109/CLUSTER.2006.311904>, 2006.

- Harpham, Q. K., Hughes, A., and Moore, R. V.: Introductory overview: The OpenMI 2.0 standard for integrating numerical models, *Environ. Model. Softw.*, 122, 104549, <https://doi.org/10.1016/j.envsoft.2019.104549>, 2019.
- Hoffmann, T., Bunney, S., Kasprzyk-Hordern, B., and Singer, A.: Wastewater catchment areas in Great Britain, ESS Open Archive, <https://doi.org/10.1002/essoar.10510612.2>, 2022.
- Höllermann, B. and Evers, M.: Perception and handling of uncertainties in water management – A study of practitioners' and scientists' perspectives on uncertainty in their daily decision-making, *Environ. Sci. Policy*, 71, 9–18, <https://doi.org/10.1016/j.envsci.2017.02.003>, 2017.
- Hreiz, R., Latifi, M. A., and Roche, N.: Optimal design and operation of activated sludge processes: State-of-the-art, *Chem. Eng. J.*, 281, 900–920, <https://doi.org/10.1016/j.cej.2015.06.125>, 2015.
- Hutton, C., Wagener, T., Freer, J., Han, D., Duffy, C., and Arheimer, B.: Most computational hydrology is not reproducible, so is it really science?, *Water Resour. Res.*, 52, 7548–7555, <https://doi.org/10.1002/2016WR019285>, 2016.
- IFAK: SIMBA (Simulation of Biological Wastewater Systems): Manual and Reference, Tech. rep., Institut für Automation und Kommunikation e. V., Magdeburg, Germany, 2007.
- Jackson-Blake, L. A., Sample, J. E., Wade, A. J., Helliwell, R. C., and Skeffington, R. A.: Are our dynamic water quality models too complex? A comparison of a new parsimonious phosphorus model, SimplyP, and INCA-P, *Water Resour. Res.*, 53, 5382–5399, <https://doi.org/10.1002/2016WR020132>, 2017.
- Khan, M. S., Liaqat, U. W., Baik, J., and Choi, M.: Stand-alone uncertainty characterization of GLEAM, GLDAS and MOD16 evapotranspiration products using an extended triple collocation approach, *Agr. Forest Meteorol.*, 252, 256–268, <https://doi.org/10.1016/j.agrformet.2018.01.022>, 2018.
- Kirchner, J. W.: Catchments as simple dynamical systems: Catchment characterization, rainfall-runoff modeling, and doing hydrology backward, *Water Resour. Res.*, 45, W02429, <https://doi.org/10.1029/2008WR006912>, 2009.
- Knoben, W. J. M., Freer, J. E., Fowler, K. J. A., Peel, M. C., and Woods, R. A.: Modular Assessment of Rainfall–Runoff Models Toolbox (MARRMoT) v1.2: an open-source, extendable framework providing implementations of 46 conceptual hydrologic models as continuous state-space formulations, *Geosci. Model Dev.*, 12, 2463–2480, <https://doi.org/10.5194/gmd-12-2463-2019>, 2019.
- Lee, D. B.: Requiem for Large-Scale Models, *J. Am. I. Planners*, 39, 163–178, <https://doi.org/10.1080/01944367308977851>, 1973.
- Leyk, S., Gaughan, A. E., Adamo, S. B., de Sherbinin, A., Balk, D., Freire, S., Rose, A., Stevens, F. R., Blankespoor, B., Frye, C., Comenetz, J., Sorichetta, A., MacManus, K., Pistolesi, L., Levy, M., Tatem, A. J., and Pesaresi, M.: The spatial allocation of population: a review of large-scale gridded population data products and their fitness for use, *Earth Syst. Sci. Data*, 11, 1385–1409, <https://doi.org/10.5194/essd-11-1385-2019>, 2019.
- Li, X. and Willems, P.: A hybrid model for fast and probabilistic urban pluvial flood prediction, *Water Resour. Res.*, 56, e2019WR025128, <https://doi.org/10.1029/2019wr025128>, 2020.
- Lin, P., Pan, M., Beck, H. E., Yang, Y., Yamazaki, D., Frasson, R., David, C. H., Durand, M., Pavelsky, T. M., Allen, G. H., Gleason, C. J., and Wood, E. F.: Global Reconstruction of Naturalized River Flows at 2.94 Million Reaches, *Water Resour. Res.*, 55, 6499–6516, <https://doi.org/10.1029/2019WR025287>, 2019.
- Lindström, G., Pers, C., Rosberg, J., Strömqvist, J., and Arheimer, B.: Development and testing of the HYPE (Hydrological Predictions for the Environment) water quality model for different spatial scales, *Hydrol. Res.*, 41, 295–319, <https://doi.org/10.2166/nh.2010.007>, 2010.
- Liu, J., Dietz, T., Carpenter, S. R., Alberti, M., Folke, C., Moran, E., Pell, A. N., Deadman, P., Kratz, T., Lubchenco, J., Ostrom, E., Ouyang, Z., Provencher, W., Redman, C. L., Schneider, S. H., and Taylor, W. W.: Complexity of Coupled Human and Natural Systems, *Science (80-)*, 317, 1513–1516, <https://doi.org/10.1126/science.1144004>, 2007.
- Liu, L., Dobson, B., and Mijic, A.: Hierarchical systems integration for coordinated urban-rural water quality management at a catchment scale, *Sci. Total Environ.*, 806, 150642, <https://doi.org/10.1016/j.scitotenv.2021.150642>, 2022.
- Liu, L., Dobson, B., and Mijic, A.: Optimisation of urban-rural nature-based solutions for integrated catchment water management, *J. Environ. Manage.*, 329, 117045, <https://doi.org/10.1016/j.jenvman.2022.117045>, 2023a.
- Liu, L., Dobson, B., and Mijic, A.: Water quality management at a critical checkpoint by coordinated multi-catchment urban-rural load allocation, *J. Environ. Manage.*, 340, 117979, <https://doi.org/10.1016/j.jenvman.2023.117979>, 2023b.
- Lower Thames Operating Agreement, 28/39/M/2: Section 20 Water Resource Management Agreement, Environment Agency, Thames Water, <http://www.thames21.org.uk/wp-content/uploads/2013/08/LTOA-Description-and-Operation.pdf> (last access: 28 May 2024), 1989.
- Mateo-Sagasta, J., Zadeh, S. M., Turrall, H., and Burke, J.: Water pollution from agriculture: a global review, *Exec. Summ.*, 35, FAO, Rome, Italy, 2017.
- Mijic, A., Bromwich, B., Crilly, D., Dobson, B., Forrow, D., Giambona, J., Kirk, S., Liu, L., and MacDonald, R.: Systems Approach to Regional Water Planning, Imperial College London, London, UK, 1–15 pp., 2022.
- Morice, C. P., Kennedy, J. J., Rayner, N. A., Winn, J. P., Hogan, E., Killick, R. E., Dunn, R. J. H., Osborn, T. J., Jones, P. D., and Simpson, I. R.: An Updated Assessment of Near-Surface Temperature Change From 1850: The HadCRUT5 Data Set, *J. Geophys. Res.-Atmos.*, 126, e2019JD032361, <https://doi.org/10.1029/2019JD032361>, 2021.
- Mortazavi-Naeini, M., Bussi, G., Elliott, J. A., Hall, J. W., and Whitehead, P. G.: Assessment of risks to public water supply from low flows and harmful water quality in a changing climate, *Water Resour. Res.*, 55, 2018WR022865, <https://doi.org/10.1029/2018WR022865>, 2019.
- Mott MacDonald: Subregional integrated water management strategy, Mott MacDonald, Reading, UK, 2023.
- Office for National Statistics: Gardens and public green space in Great Britain, Office for National Statistics, Newport, Wales, UK, <https://www.ons.gov.uk/economy/environmentalaccounts/datasets/accesstogardensandpublicgreenspaceingreatbritain>, last access: 13 April 2023.
- Puchol-Salort, P., Boskovic, S., Dobson, B., van Reeuwijk, M., and Mijic, A.: Water neutrality framework for systemic design of new urban developments, *Water Res.*, 219, 118583, <https://doi.org/10.1016/j.watres.2022.118583>, 2022.

- Rauch, W., Urich, C., Bach, P. M., Rogers, B. C., de Haan, F. J., Brown, R. R., Mair, M., McCarthy, D. T., Kleidorfer, M., Sitzenfrei, R., and Deletic, A.: Modelling transitions in urban water systems, *Water Res.*, 126, 501–514, <https://doi.org/10.1016/j.watres.2017.09.039>, 2017.
- Razavi, S., Tolson, B. A., and Burn, D. H.: Review of surrogate modeling in water resources, *Water Resour. Res.*, 48, W07401, <https://doi.org/10.1029/2011WR011527>, 2012.
- Razavi, S., Hannah, D. M., Elshorbagy, A., Kumar, S., Marshall, L., Solomatine, D. P., Dezfouli, A., Sadegh, M., and Famiglietti, J.: Coevolution of machine learning and process-based modelling to revolutionize Earth and environmental sciences: A perspective, *Hydrol. Process.*, 36, e14596, <https://doi.org/10.1002/hyp.14596>, 2022.
- Schmitt, T. G. and Huber, W. C.: The scope of integrated modelling: system boundaries, sub-systems, scales and disciplines, *Water Sci. Technol.*, 54, 405–413, <https://doi.org/10.2166/wst.2006.595>, 2006.
- Schouten, A. P., van den Hooff, B., and Feldberg, F.: Virtual Team Work, *Commun. Res.*, 43, 180–210, <https://doi.org/10.1177/0093650213509667>, 2016.
- Sun, Q., Miao, C., Duan, Q., Ashouri, H., Sorooshian, S., and Hsu, K.: A Review of Global Precipitation Data Sets: Data Sources, Estimation, and Intercomparisons, *Rev. Geophys.*, 56, 79–107, <https://doi.org/10.1002/2017RG000574>, 2018.
- Tang, F. H. M., Lenzen, M., McBratney, A., and Maggi, F.: Risk of pesticide pollution at the global scale, *Nat. Geosci.*, 14, 206–210, <https://doi.org/10.1038/s41561-021-00712-5>, 2021.
- Thenkabail, P., Teluguntla, P., Xiong, J., Oliphant, A., and Massey, R.: NASA MEaSURES Global Food Security Support Analysis Data (GFSAD) Crop Mask 2010 Global 1 km V001, NASA EOSDIS L. Process. DAAC, <https://doi.org/10.5067/MEaSURES/GFSAD/GFSAD1KCD.001>, 2016.
- Thenkabail, P. S., Biradar, C. M., Noojipady, P., Dheeravath, V., Li, Y., Velpuri, M., Gumma, M., Gangalakunta, O. R. P., Turrall, H., Cai, X., Vithanage, J., Schull, M. A., and Dutta, R.: Global irrigated area map (GIAM), derived from remote sensing, for the end of the last millennium, *Int. J. Remote Sens.*, 30, 3679–3733, <https://doi.org/10.1080/01431160802698919>, 2009.
- Thrysoe, C., Arnbjerg-Nielsen, K., and Borup, M.: Identifying fit-for-purpose lumped surrogate models for large urban drainage systems using GLUE, *J. Hydrol.*, 568, 517–533, <https://doi.org/10.1016/j.jhydrol.2018.11.005>, 2019.
- Todorović, A., Stanić, M., Vasilčić, Ž., and Plavšić, J.: The 3DNet-Catch hydrologic model: Development and evaluation, *J. Hydrol.*, 568, 26–45, <https://doi.org/10.1016/j.jhydrol.2018.10.040>, 2019.
- Tscheikner-Gratl, F., Bellos, V., Schellart, A., Moreno-Rodenas, A., Muthusamy, M., Langeveld, J., Clemens, F., Benedetti, L., Rico-Ramirez, M. A., de Carvalho, R. F., Breuer, L., Shucksmith, J., Heuvelink, G. B. M., and Tait, S.: Recent insights on uncertainties present in integrated catchment water quality modelling, *Water Res.*, 150, 368–379, <https://doi.org/10.1016/j.watres.2018.11.079>, 2019.
- Vanhooen, H., Meirlaen, J., Amerlinck, Y., Claeys, F., Vangheluwe, H., and Vanrolleghem, P. A.: WEST: modelling biological wastewater treatment, *J. Hydroinform.*, 5, 27–50, <https://doi.org/10.2166/hydro.2003.0003>, 2003.
- Voinov, A. and Shugart, H. H.: “Integronsters”, integral and integrated modeling, *Environ. Model. Softw.*, 39, 149–158, <https://doi.org/10.1016/j.envsoft.2012.05.014>, 2013.
- Wada, Y., Bierkens, M. F. P., de Roo, A., Dirmeyer, P. A., Famiglietti, J. S., Hanasaki, N., Konar, M., Liu, J., Müller Schmied, H., Oki, T., Pokhrel, Y., Sivapalan, M., Troy, T. J., van Dijk, A. I. J. M., van Emmerik, T., Van Huijgevoort, M. H. J., Van Lanen, H. A. J., Vörösmarty, C. J., Wanders, N., and Wheeler, H.: Human–water interface in hydrological modelling: current status and future directions, *Hydrol. Earth Syst. Sci.*, 21, 4169–4193, <https://doi.org/10.5194/hess-21-4169-2017>, 2017.
- Whitehead, P. G., Wilson, E. J., and Butterfield, D.: A semi-distributed Integrated Nitrogen model for multiple source assessment in Catchments (INCA): Part I – model structure and process equations, *Sci. Total Environ.*, 210–211, 547–558, [https://doi.org/10.1016/S0048-9697\(98\)00037-0](https://doi.org/10.1016/S0048-9697(98)00037-0), 1998.
- WSIMOD documentation: WSIMOD Docs [code], <https://imperialcollegelondon.github.io/wsi> (last access: 26 March 2024).
- Zhang, Y., Schaap, M. G., and Zha, Y.: A high-resolution global map of soil hydraulic properties produced by a hierarchical parameterization of a physically based water retention model, *Water Resour. Res.*, 54, 9774–9790, <https://doi.org/10.1029/2018WR023539>, 2018.
- Zomorodian, M., Lai, S. H., Homayounfar, M., Ibrahim, S., Fatemi, S. E., and El-Shafie, A.: The state-of-the-art system dynamics application in integrated water resources modeling, *J. Environ. Manage.*, 227, 294–304, <https://doi.org/10.1016/j.jenvman.2018.08.097>, 2018.