

Customizing large-scale hydrological models: Harnessing the open data realm for impactful local applications

Ilias G. Pechlivanidis^{*}, Jude Lubega Musuuza

Swedish Meteorological and Hydrological Institute, Norrköping, Sweden

ARTICLE INFO

Keywords:

Large-scale hydrological modelling
Hydrological model customization
Local applications
Water resources management
Local data and knowledge

ABSTRACT

Study region: Lake Hume in Australia and Harsha Lake in USA.

Study focus: Large-scale hydrological models (LSHMs), though important for both scientific and societal reasons, require the representation of many unknown features that influence river system response. However, current model identification practices in catchment modelling cannot lead to robust LSHMs for local decision-making. To address this, it is necessary to customise the models by integrating local data and knowledge from various sources (e.g. in-situ and earth observations) and fluxes. We present a framework to customize LSHMs for impactful local applications and showcase this using the global WWH hydrological model as the reference LSHM.

New hydrological insights: We present significant improvements in modelling streamflow and actual and potential evapotranspiration, following WWH refinements to include local lakes and reservoir management. Local streamflow measurements and earth observations from NASA MODIS evapotranspiration products were used to re-calibrate the locally adapted model, leading to spatial consistency in performance. Combining multiple variables and metrics during model identification improved streamflow performance and robustness, with combination sets reducing variability and enhancing representation of diverse hydrological processes, highlighting the need for tailored metric and variable selection. This underpins the importance of including informative data in customized multi-objective modelling chains. Finally, incorporating reservoir management improved simulation of a regulated system, with local insights informing reservoir parameterization in LSHMs and bridging the gap to global-scale applications.

1. Introduction

Hydrological models have played a key role in filling in information gaps in space and time to better address water resources challenges, predict hydrological hazards, and inform decisions related to water management and infrastructure planning (Brachowitz et al., 2013; Pechlivanidis et al., 2011). Model outputs are yet subject to uncertainty in the input data, model structure and parameterization, affecting further the performance and process understanding. The limitations in the model setups and inherited and/or propagated model uncertainties must be carefully considered and addressed in order to ensure robust support tools for decision-making (Clark et al., 2017; Seibert and Van Meerveld, 2016). The difficulty in identifying catchment-scale models (structure and parameters), which are “good for the right reasons” depends on the complexity of the local hydrological processes, the information content of the available data to tune the model parameters, and the number of model parameters that proportionally increases their

^{*} Corresponding author.

E-mail address: ilias.pechlivanidis@smhi.se (I.G. Pechlivanidis).

<https://doi.org/10.1016/j.ejrh.2025.102390>

Received 20 January 2025; Received in revised form 8 April 2025; Accepted 11 April 2025

Available online 14 April 2025

2214-5818/© 2025 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY-NC license (<http://creativecommons.org/licenses/by-nc/4.0/>).

identification vulnerability to equifinality (Wagener and Gupta, 2005). Best practices aiming at improving the accuracy and robustness in catchment-scale modelling incorporate multiple variables and sources of information (Hundecha et al., 2020; Musuuza et al., 2023).

Hydrological modelling at the large-scale (national, continental and global) is scientifically attractive since, for instance, the large samples generated, as model outputs, can lead to process understanding along the Earth's strong hydroclimatic gradient (Dasgupta et al., 2025; Pechlivanidis and Arheimer, 2015). Such large-scale modelling is also important for society with the hydro-meteorological information being provided at poorly gauged and even ungauged locations addressing the various water-related challenges (Beck et al., 2020; Hundecha et al., 2016). Large-scale hydrological models (LSHMs) are consequently expected to provide valuable insights into complex river systems that are difficult to study through direct monitoring and describe the river system functions and responses to different inputs and environmental factors (Gupta et al., 2014; Kumar et al., 2013). Particularly in transboundary river systems, where occasionally local data are confidential to enhance downstream process understanding, LSHMs can simulate the system's response, including human impacts, and provide insights into the underlying processes that drive its behaviour (Bierkens et al., 2015; Du et al., 2022). However, the models often depend on open global datasets which are commonly erroneous or inconsistent, whilst the high complexity of large models can lead to overfitting, where parameters are tuned to specific data and/or regions, and therefore may not be applicable to other datasets and hydro-climatic conditions (Kauffeldt et al., 2013). Importantly, many large-scale catchments are significantly affected by localized water regulation, irrigation, abstractions, channel modifications, etc., which are not properly described in the global databases at the relevant spatiotemporal resolutions.

The usability of new data types, such as Earth observations (EOs), media and crowdsourced data, overcome limitations of traditional local in-situ observations and can add information in space and time when this is not adequate due to ungauged conditions or sparse in-situ networks (Alfieri et al., 2022). EOs in particular provide a wealth of information for different hydrologically useful variables, e.g., precipitation, soil moisture, snow cover, evapotranspiration etc. which can be used for: (i) improved understanding on how and against which observations the model estimates can be most appropriately compared, (ii) identifying model inconsistencies (i. e. systematic errors) that can subsequently lead to improvements in model structure and/or parameterization, (iii) driving the data assimilation schemes and adding value to the development of the techniques, and (iv) achieving a “good for the right reasons” model. The corresponding refinements to model structure and parameterization can lead to improved process understanding, while the incorporation of such new data into LSHMs can result in improved accuracy and robustness of model predictions (Andersson et al., 2022; Musuuza et al., 2020). However, the reliability of such data varies both geographically and seasonally, and therefore their direct integration into LSHMs needs to be done with caution (Nijzink et al., 2018).

The common lack of detail and LSHMs dependence on global open datasets often points towards the necessity to integrate local data and knowledge for enhancing the understanding of local hydrological processes. Local data such as land use, topography and soil properties can be used to refine model setups, while local knowledge on hydrological processes such as groundwater recharge, abstraction and flow paths can be incorporated into LSHMs to improve the representation of local conditions. In the cases that local streamflow measurements are available, the information content for hydrological modelling is often high and the data can be used to diagnose LSHM limitations at the local scale. The integration of knowledge into LSHMs can finally help to build user engagement and

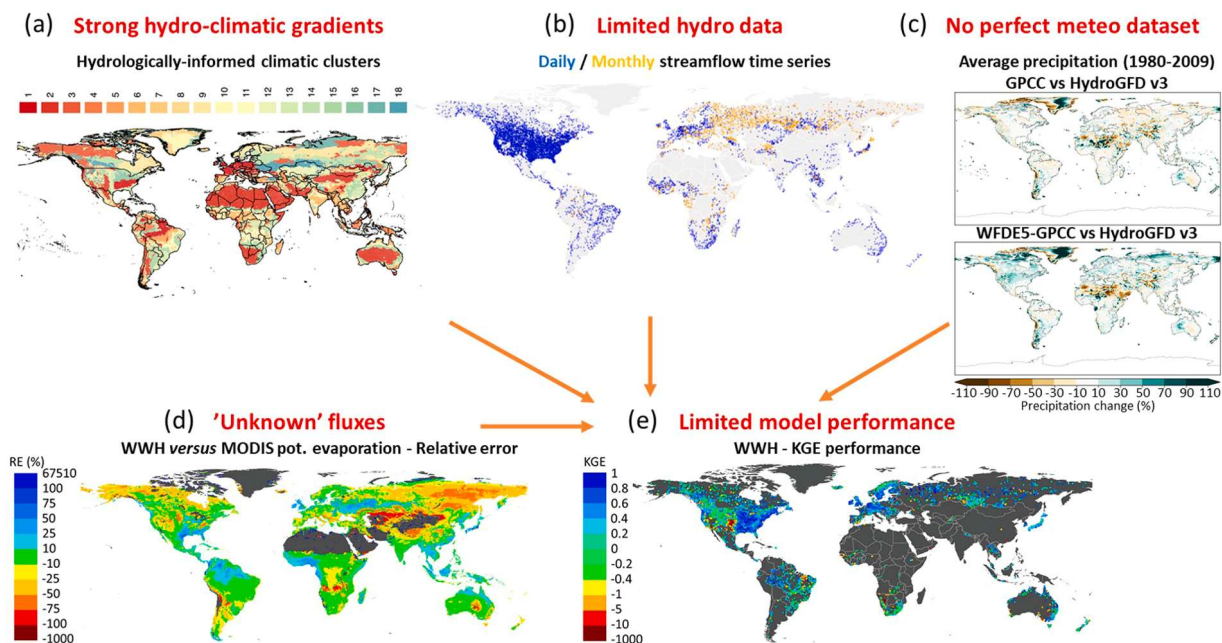


Fig. 1. Exemplifying the challenges in large-scale hydrological modelling resulting to knowledge gaps and uncertain decision-making at the local scale: (a) strong hydro-climatic gradients, (b) ungauged river systems, (c) lack of perfect meteorological dataset, (d) unknown hydro-meteorological fluxes, and (e) limited model performance.

improve the relevance and usability of model results. The incorporation of local knowledge and input from local users ensures that the LSHMs are tailored to local needs and priorities, further adding confidence in decision-making. Nonetheless, it is important to acknowledge that, like data, local knowledge can be subjective, uncertain and spatially variable; factors that can influence the model's adaptability and predictive accuracy.

Despite the large efforts, the information content in the “traditional” data used in catchment-scale modelling is not sufficient to diagnostically guide the parameter identification and improve process understanding in LSHM-based applications. The large spatial scales require many “unknown” features, e.g., lake extents, snowpack cover, reservoir regulations, dam capacities, irrigation practices, etc., that directly influence the hydrological response, while fluxes in the catchment can still be left unresolved. Due to the large physical heterogeneity, hydrological model complexity and the limitations in the traditionally used data, the current model identification practices in large-scale modelling cannot accurately represent the physical processes and lead to robust models for local applications. Even when more sophisticated identification practices are applied, LSHMs are still over-tuned to capture the dominant processes and their spatiotemporal dynamics, and little attention is paid to fluxes at regional scales. To make LSHMs useful at regional and smaller scales, effort is needed to customize the models using local data for various fluxes, e.g. in-situ and EOs.

Here, we present a framework to customize LSHMs for impactful local applications and showcase this for Lake Hume in Australia and Harsha Lake in the United States using the global World-Wide HYPE (WWH) hydrological model as the reference LSHM. We gain insights by posing the following questions that not only add scientific value but also guide developments in large-scale hydrological services: (1) *To what extent are LSHMs relevant for estimating hydrological fluxes at the regional and finer scales?* and (2) *How can the integration of local knowledge and data into LSHMs improve the local process understanding?* To address these questions, we: (i) present the information missed by WWH at the local scale, including mis-delineation and omission of important features, e.g., reservoirs, (ii) integrate local knowledge, e.g., reservoir operation into the model re-delineation, and (iii) quantify the model performance improvement (hence process understanding) by fine-tuning the processes and parameters using local in-situ data and EOs.

2. Challenges in large-scale hydrological modelling and customization for local applications

Firstly, we acknowledge the most important challenges in the implementation and assessment of LSHMs, which are also exemplified in Fig. 1. These sources of uncertainty and error during the setup and identification of LSHMs can lead to poor model performance (see Fig. 1e, reproduced from Arheimer et al. (2020)) and the ultimate incomplete process understanding. This knowledge gap limits the usability of existing LSHMs for local applications, highlighting further the need for proposing frameworks and developing methodologies that can efficiently add local value to LSHM data and information.

2.1. Strong hydro-climatic gradients

The high variability of the dominant hydrological processes across different geographic regions leads to variations in the movement, distribution and quality of the water, and is subject to the climatic and land-soil spatiotemporal conditions. Even regions within close proximity can experience different hydro-climatic conditions particularly where the elevation relief is high (i.e. the Norwegian river systems), which consequently affects the runoff generation mechanisms (Girons Lopez et al., 2025, 2021; Kuentz et al., 2017). For instance, mountainous areas can experience different climatic zones depending on their altitude, with snow and ice at higher elevations and temperate or subtropical conditions at lower elevations (Pechlivanidis et al., 2020). Moreover, regions near the equator tend to have high rainfall and warm temperatures, while arid regions such as deserts have very little precipitation and extreme temperature fluctuations. The challenge is also amplified at the national, continental and even global scales, where human impacts (irrigation, reservoir regulation and water transfer) can strongly disturb the natural flows and affect the water balance at the catchment scale. To exemplify the global hydro-climatic gradient, Fig. 1a presents the hydrologically-informed climatic clusters (a global classification of natural systems to structure the climatic forcing that underlies hydrologic response) generated based on the method of Knoben et al. (2018). The main outcome here, regardless of the strong hydro-climatic gradient, is that the single (process-based) model structure integrated in the LSHMs is expected to describe a highly heterogeneous set of dominant processes.

2.2. Ungauged river systems

LSHMs rely heavily on accurate, standardized and quality assured streamflow time series in order to guide the calibration and validation process and to identify the most robust model structure and parameter values (Pechlivanidis et al., 2011). However, many river systems around the world are considered poorly gauged or even ungauged, which consequently makes hydrological modelling very challenging (Wagener and Gupta, 2005). The information content in the gauged river systems is also subject to the length of the available records and the data observation frequency, while many LSHM applications deal with incomplete time series and are subject to significant uncertainty (e.g. propagated uncertainty from the water level - streamflow rating curves). To exemplify the data-sparse streamflow networks, Fig. 1b presents the streamflow data collected during the setup of the WWH hydrological model, referenced in Arheimer et al. (2020). Traditional data collection practices, for instance stream gauging stations, are often challenging in remote or inaccessible areas, over which hydrological processes are unique, i.e. Tibetan Plateau, affecting the water cycle at the large-scale. Although regionalization practices have been developed for identifying catchment-scale hydrological models in ungauged conditions, these methods often have limitations that result in amplified propagated uncertainties in the model predictions at the local scale, particularly in a changing environment (Guo et al., 2021; Pagliero et al., 2019).

2.3. No “perfect” meteorological dataset

It is widely accepted that there is no perfect continental or global product (gridded observations, reanalysis, satellite-based, citizen science based, or hybrid) that can accurately capture meteorological variables (precipitation, temperature, humidity, etc.) at the regional scale across the globe (Haylock et al., 2008; Thiemiig et al., 2022). Such meteorological variables are complex and affected by a range of factors including topography, land use and atmospheric dynamics that vary across different regions. However, hydrological monitoring and predictions are subject to the meteorological input and its inherent uncertainty. The accuracy of such products relies on the network of in-situ observations used to correct/condition the products, while reanalysis and satellite-based products are driven by different retrieval algorithms and hypotheses, which lead to discrepancies in the meteorological estimates (Beck et al., 2017; Sun et al., 2018). For instance, Fig. 1c (reproduced from Berg et al. 2021) highlights the regional discrepancies in the global spatial patterns of average precipitation from the GPCC (Schneider et al., 2018) and WFDE5-GPCC (Copernicus Climate Change Service, 2020) products to the HydroGFD v3 for 1980–2009. Given that it is unlikely that a perfect global meteorological product will ever be achieved, using an ensemble of multiple data sources can compensate for the meteorological uncertainty in hydrological assessments. However, large ensembles can also largely increase the computational costs (run times and data storage) in LSHMs (Tang et al., 2022).

2.4. Unknown hydro-meteorological fluxes

In addition to the traditional hydro-meteorological networks, there has been an evolution in EO technologies and an expansion of monitoring stations using low cost instruments (Kouki et al., 2022; Mao et al., 2020; Pandeya et al., 2021). However, there are still hydro-meteorological state variables and fluxes that are critical components of the water cycle and yet are not well understood and/or quantified (i.e. evapotranspiration, groundwater recharge and soil moisture). For instance, although EOs can provide spatially and temporally continuous data, their reliability varies and hence the understanding of the corresponding fluxes can be limited. An example of the error between the NASA MODIS potential evapotranspiration (PET) product and the WWH hydrological model is presented in Fig. 1d highlighting the incomplete knowledge of the PET flux (Pimentel et al., 2023). New approaches, such as machine learning and data assimilation have allowed the incorporation of such data into LSHMs and assisted in filling the knowledge gaps (De Lannoy et al., 2022; Khaki et al., 2019). Despite their potential in extracting information from the data, these methods can mislead hydrological modelling if the data are highly uncertain or erroneous. Therefore, other methods to account for uncertain fluxes aim to condition the LSHM parameters (instead of optimising them) during the model identification process (Nijzink et al., 2018; Pechlivanidis and Arheimer, 2015).

2.5. LSHM customization framework for local applications

To address the limitations of LSHMs in accurately representing local basin characteristics, we propose a structured framework to customize LSHMs for impactful local applications supporting the co-evolution of knowledge and bridging the gap between research, operational services and decision-making. This framework comprises four core components: (1) Benchmark LSHM, (2) Explore Local Information, (3) LSHM Local Customization and (4) Customized LSHM Operationalization. This approach tailors the LSHM setup to local conditions, enhancing model accuracy and supporting more informed local decision-making, as outlined in Fig. 2.

In the initial phase, *Benchmark LSHM*, the objective is to establish a shared understanding of the baseline LSHM capabilities and

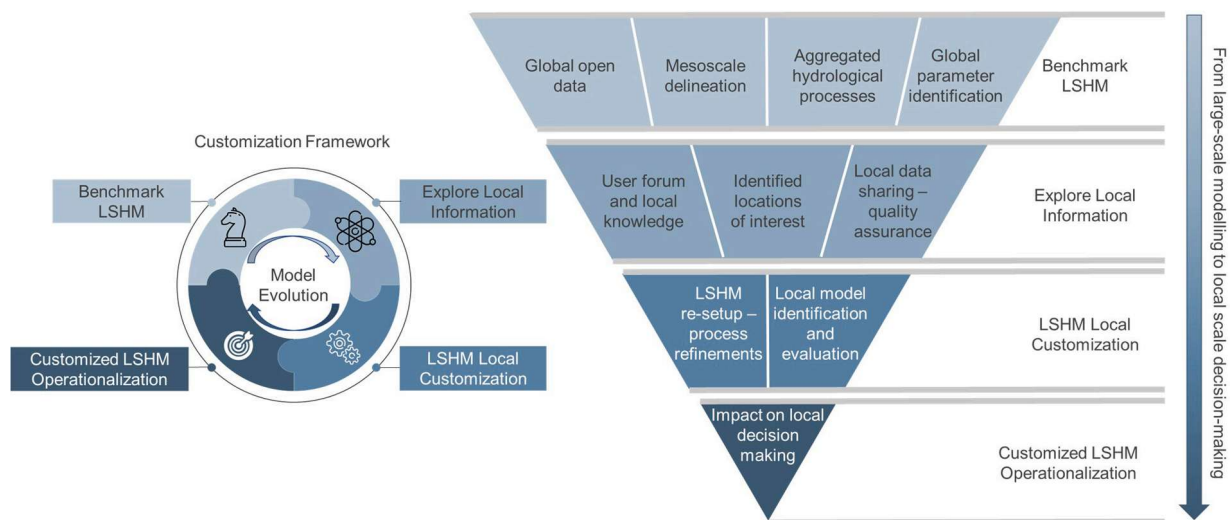


Fig. 2. The proposed framework for tailoring LSHMs for impactful local applications. The pillars of the framework are represented as layers on the reversed pyramid with each layer including the main tasks involved.

limitations. This phase involves interaction with local users to ensure the model accurately reflects the general domain characteristics while identifying systematic data errors. Given the inherent difficulty in capturing detailed local features across extensive domains, baseline delineations in LSHM setups are typically coarse. This can impact both model performance and applicability at finer scales. Therefore, *Explore Local Information* is a key phase where users are engaged to identify locally available data, knowledge, and their specific needs that inform the customization process. Emphasis is placed on balancing scientific advancements, technological feasibility, and the integration of quality-assured local data.

The *LSHM Local Customization* phase focuses on incorporating this local information into the locally adapted model, enhancing its representation of dominant hydrological processes within targeted sub-domains. Customization may include re-delineation of LSHM sub-basins in areas of specific interest and fine-tuning model parameters to align with local hydrological dynamics. This re-delineation is informed by specific monitoring needs, availability of local streamflow observations and data that support parameter estimation. Model parameter identification is conducted to refine the optimal parameter set, thereby improving the model's overall performance by aligning it with observed local behaviour. In the *Customized LSHM Operationalization* phase, the framework aims to facilitate service evolution by emphasizing the added value of localized LSHM outputs for local decision-making, while also addressing inherent uncertainties. This step is essential to foster user engagement and illustrate how the customized LSHM enhances decision-making accuracy.

3. Case studies and data

3.1. Case studies

3.1.1. The Lake Hume (Australia)

The Murray-Mitta Mitta catchment, encompassing a 15,300 km² area with its outlet at Lake Hume, is predominantly forested (80 %) and includes a National Park. This catchment is largely undisturbed by cultivation and features minor urbanised areas in its lowlands. Annual precipitation ranges from 350 mm in the lowlands to 850 mm in the mountainous areas, with significant temperature variations between summer (29–32°C) and winter (11–14°C) (Davy and Koen, 2013). The predominant clay soils exhibit marked seasonal swelling and cracking. The 20 km² Lake Hume is a large impoundment formed by the construction of the Hume Dam below the junction of the Murray and Mitta Mitta Rivers, and has a 3005 million m³ capacity (Fig. 3a).

The Upper Murray catchment comprises only 2 % of the Murray-Darling Basin but contributes 17 % of the entire basin water and is therefore one of the major freshwater sources of South Western Australia. The area receives most of the precipitation during June to August, yet still a significant amount of rainfall occurs during the other months. Temperature also has a clear seasonal pattern, with the peak in December to February, and the mean temperature dropping to about 5 °C in July. Streamflow can be close to zero from January till May and then increase during the rainy season (June–November). Inflows and outflows of Lake Hume are highly regulated with several dams upstream of the lake comprising the Snowy Mountains hydro-electric scheme in the Murray catchment and the Dartmouth Dam along the Mitta Mitta River (Fig. 3a). This is also reflected in a strong seasonal, as well as interannual change in the water level of Lake Hume.

3.1.2. The Harsha Lake (USA)

The Harsha Lake, located in the East Fork catchment of the Little Miami River near Cincinnati, Ohio, serves a key role in flood control, recreation and water supply. The 1,300 km² catchment area is predominantly agricultural (42.5 %) with significant urban residential (32.3 %) and forested areas (25.2 %), especially around the lake (Fig. 3b). The soil consists mainly of moderately thick loess over loamy glacial till. The region experiences maximum precipitation in summer (May–July) with notable rainfall all year round.

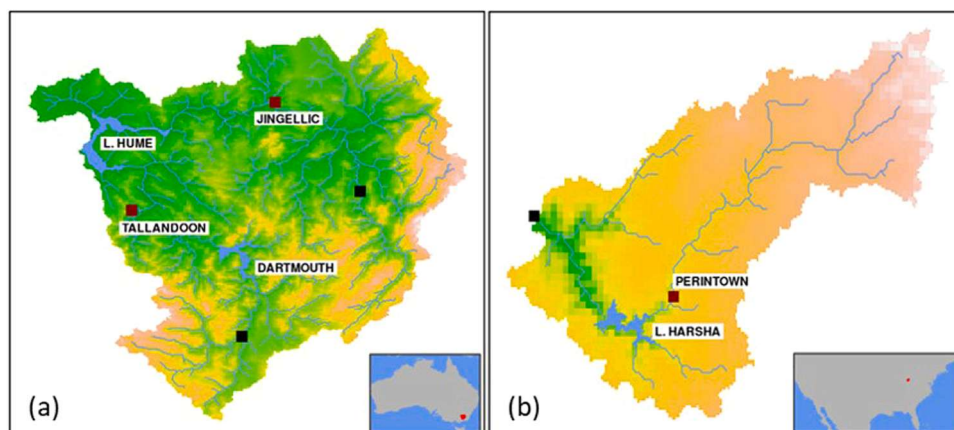


Fig. 3. The location of the two case studies: (a) the Lake Hume in Australia, and (b) the Harsha Lake in the USA. The river networks are overlaid over terrain maps with the black and red markers indicating the original open data and data provided by local users, respectively.

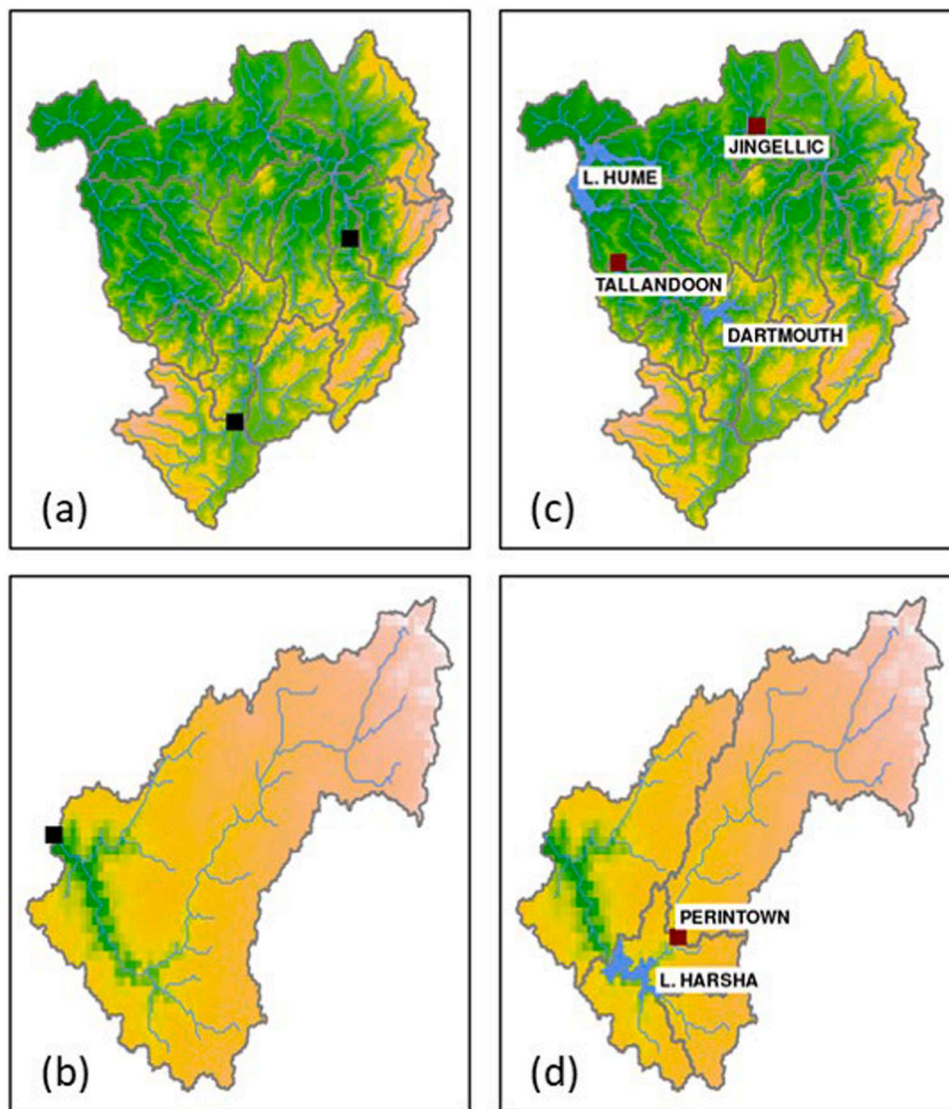


Fig. 4. The default WWH (a and b) and locally-refined (c and d) representation of the Lake Hume (top) and Harsha Lake (bottom) drainage areas. The square markers are the streamflow gauges, the blue shapes are the resolved local water bodies, and the grey lines are the borders of the sub-basins.

Temperature peaks occur from June to August. Streamflow is minimal in late summer to early autumn (June–October), increasing in winter and spring. Harsha Lake provides water to 30 % of the 206,000 residents in Clermont County and covers about 69 % of the East Fork catchment, approximately 890 km², highlighting its importance in regional hydrology and water resource management.

3.2. Historical meteorological data

We use the HydroGFD v3.0 (Hydrological Global Forcing Data version 3.0) product, an observation-corrected reanalysis dataset providing historical daily precipitation and minimum, mean and maximum temperatures (Berg et al., 2021). The dataset is global at a spatial resolution of about 25 km and is available operationally for the period 1979 to near real-time. HydroGFD is adjusted against global gridded observational datasets, while to acquire real-time data, a collection of several reference datasets is used. Temporal consistency in the data is achieved by relying on a background climatology and only making use of anomalies from the different datasets. Precipitation is adjusted for mean bias, as well as the number of wet days in a month. The latter relies on a calibrated statistical method with input of the monthly precipitation anomaly, such that no additional input data about the number of wet days are necessary. The daily mean temperature is adjusted toward the monthly mean of the observations and applied to 1-hour time steps of the ERA5 reanalysis (Hersbach et al., 2020). Daily mean, minimum and maximum temperature are then calculated.

3.3. Earth observations and in-situ hydrological data

The global EO-based actual and potential evapotranspiration (AET and PET respectively) were available from the National Aeronautics and Space Administration (NASA) at a 500 m spatial resolution and 8-day aggregates (Mu et al., 2011). The data were retrieved from the Moderate Resolution Imaging Spectroradiometer (MODIS) on the Terra and Aqua satellites in 36 spectral bands, and cover the period 2001 to near real-time. To obtain daily estimates, the 8-day aggregates were uniformly distributed across each day of the interval.

Daily streamflow data (QOB) from the Global Runoff Data Centre (GRDC) had been available, while local QOB were also provided by the Commonwealth Scientific and Industrial Research Organisation (CSIRO) and the United States Geological Survey (USGS) for the Lake Hume and Harsha Lake catchments, respectively. The data for Lake Hume were available at the Tallandoon and Jingellic stations with the former being subject to upstream regulations and the latter relatively pristine, with the records starting in the early 20th century to 2020 (the year when the data were accessed); in particular for Tallandoon and Jingellic, the records start in 1934 and 1890, respectively. QOB for the Harsha Lake was available at the Perintown station that is upstream of the lake, with the records covering the period 2000–2018. See Table A1 in the Appendix for a summary of all datasets and their characteristics.

4. Methodology

4.1. The global WWH hydrological model - benchmark LSHM

The HYPE hydrological model structure was applied at the global scale covering a domain of about 135 million km², named World-Wide HYPE (WWH). The model has a spatial resolution of about 131300 sub-basins with an average size of 1000 km² and runs at a daily time step. The model setup is based on open data using 36 land use classes and 5 elevation zones, and based on their combination, a maximum of 169 hydrological response units (HRUs) were identified. Of the 5338 gauging stations that have at least 10 years of data, 2475 stations were used for calibration while 2863 independent stations were used for model validation. The multi-basin calibration followed a stepwise approach, where parameters for eleven hydrological processes were calibrated individually using representative gauged catchments selected based on key physiographic characteristics (e.g., climate, land use, elevation), with each gauge contributing to a limited subset of parameters to ensure a robust and spatially transferable parameterization across the global domain. This approach helps prevent overfitting despite the large number of parameters and enables a robust and spatially transferable parameterization across the global domain. See Arheimer et al. (2020) for details about the station selection and calibration procedure.

At the calibration stations, the model performance is estimated to a median monthly Kling-Gupta Efficiency (KGE; Gupta et al. 2009) of 0.40 (see Fig. 1e). When decomposing the KGE, a median correlation coefficient of 0.76 and a median relative error of −15 % is found. This means that the model captures the temporal dynamics of the hydrographs reasonably well in many sites, while it generally underestimates streamflow volume. Overall, WWH shows, independently of physiography, good potential to capture the spatial variability of high streamflow, duration of low streamflow, monthly high streamflow, and constancy of daily streamflow. The results imply that the overall process understanding behind the WWH model setup may be relevant at the global scale; however, the estimation of parameter values or the quality of forcing data are not always optimal for capturing the streamflow dynamics at the local scale.

4.2. Experimental setup

4.2.1. Parameter identification approach

LSHMs traditionally undergo a multi-basin approach accounting for calibration and validation stations over the domain and identifying the parameter sets based on the mean/median performance across the selected stations (Hundecha et al., 2016; Kumar et al., 2013). Here, we argue that the LSHM identification needs to additionally be based on a multi-objective, multi-period and multi-variable approach that integrates as much local data and knowledge as possible, thus allowing robust data information extraction through parameter fitting (Hundecha et al., 2020; Nijzink et al., 2018; Rakovec et al., 2019; Samaniego et al., 2013). This approach will allow accounting for the spatial heterogeneity of the river systems (multi-basin), capturing the variability and changes in the temporal dynamics (multi-period), representing the attributes of the hydro-meteorological signals (multi-objective) and ensuring a “good model for the right reasons” by adequately simulating the different fluxes (multi-variable).

Here, the multi-period approach moved beyond the traditional distinct splitting of time-series into calibration and validation. Instead a split-sample cross-calibration-validation approach was followed enabling the use of short historical records, commonly found in EOs, by including alternate years in the calibration and validation periods (Clemenzi et al., 2023; Klemes, 1986). Moreover, we assessed the multi-objective model performance with respect to the Kling-Gupta Efficiency (k ; accounting for the overall performance), the absolute relative error (α ; accounting for the volumetric biases), the correlation coefficient (c ; accounting for timing errors) and the ratio between standard deviations in simulations and observations (v ; accounting for the biases in variability). Finally, the multi-variable approach accounted for the model performance with regard to the in-situ streamflow and EO-based potential and actual evapotranspiration.

Expert knowledge based on previous parameter identification investigations has allowed setting a priori ranges on WWH sensitive parameters that control dominant processes (Brendel et al., 2023; Clemenzi et al., 2023; Pechlivanidis and Arheimer, 2015; see Table A2 in the Appendix). This constraint allows tackling parameter overfitting to some extent in this processed-based model. Here, a Monte Carlo ensemble with 100,000 parameter sets was generated for each case study using the efficient Latin hypercube sampling

algorithm that adequately covers the parameter space, available in the SAFE toolbox (Pianosi et al., 2015). To account for parameter equifinality, the behavioural models were identified through a multi-step identification process based on a modified Generalised Likelihood Uncertainty Estimation (GLUE) technique (Beven and Binley, 1992). The process has the following steps: (i) a Monte Carlo sampling from a parameter space is used to generate a prior distribution of model realizations (Beven and Freer, 2001), (ii) prior likelihood transformations of the performance, for a selected metric, are introduced to obtain the posterior distributions using the pseudo-likelihood values as weights to rescale the posterior distribution, and (iii) the rescaling is furnished by resampling the prior distribution with replacement with the estimated pseudo-likelihood values as probability.

The performance values are further scaled on the range [0,1], subject to cut-off thresholds which designates non-behavioural models deemed to be unacceptable. Here, these thresholds were the 50th and 90th percentiles of the performance distribution, and were the only subjective choice made. Additionally, the pseudo-likelihood transformations enable a straightforward (multiplicative) combination of different performance metrics and variables to rescale the prior distribution into the posterior distribution without the need for subjective weights. In particular, the WWH parameters for the evapotranspiration, recession, water storages, and macropore flow processes (see Table A2 in the Appendix for the ranges) were optimised in the dominant land use and soil classes upstream of the unregulated gauges. Finally, the optimum parameter set derived with the likelihood approach was used to fine-tune the reservoir and lake parameters so that the regulated streamflow response is matched as closely as possible.

4.2.2. Evaluating customized model's credibility

Specifically, this study evaluates: (i) the capability of the default WWH model (Arheimer et al. 2020; with parameters derived from multi-basin calibration across the global domain) to accurately estimate hydrological fluxes at the local scale, and (ii) the improvement in model performance and local process understanding gained through the integration of local data and knowledge. In line with the proposed customization framework, task (i) involves identifying limitations in the WWH model at local scales, such as mis-delineation and the omission of certain features, and subsequently improving model delineation by incorporating local knowledge. Task (ii) assesses the performance enhancement achieved by adjusting model parameters using local in-situ data and EOs. In the task (ii) experiments, streamflow serves as the target variable.

Model performance is evaluated at a daily timestep based on behavioural models identified by the Kling-Gupta Efficiency, using various combinations of metrics and variables. The first set of experiments examines models calibrated with single variables across all performance metrics, resulting in 24 distinct variable-metric-period (VMP) setups (3 variables $\times$ 4 metrics $\times$ 2 periods). This initial analysis aims to identify model configurations achieving high streamflow performance. Subsequently, we explore multiple (up to 6) VMP combinations involving different metrics, periods and variables to further refine model performance. The VMP-n naming convention is further used to denote setups that include n (n ranging from 1 to 6) distinct combinations of variables, metrics and/or periods (e.g., VMP-1, VMP-2 etc.), allowing for a structured evaluation of the role of increasing information content in model identification. Overall, we accounted for more than 75 additional VMP setups. This analysis provides insight into the effects of incorporating additional data sources on model accuracy and helps identify an optimal combination set for accurately representing streamflow. Additionally, we conduct a visual assessment of model performance using streamflow signatures, including the flow duration curve and intra-annual variability. Finally, we evaluate the spatial consistency of model performance, with respect to EO information for evapotranspiration.

5. Results

5.1. Explore local information - integrating local knowledge in model re-delineation

Fine-tuning the delineation of catchments based on local knowledge has been a crucial step in improving the local model performance. Here we present these refinements to the Lake Hume and Harsha Lake catchments, addressing critical limitations in the default WWH model setup (benchmark LSHM). In the re-delineated setup, the Lake Hume was modelled as a single, autonomous sub-basin, which can allow for a more precise representation of its storage and regulation schemes. Similarly, the Dartmouth Reservoir, originally unresolved in the global WWH model, was also defined as an autonomous sub-basin. This adjustment enables a more explicit modelling of the regulation impacts downstream. The re-delineation process also involved updating the gauges used for model evaluation. The two GRDC gauges previously utilized in WWH were replaced with gauges at Tallandoon and Jingellic, which were provided by local stakeholders (CSIRO) as more informative in terms of data quality and record length. The Tallandoon gauge, located downstream of the Dartmouth Reservoir, can particularly provide data for calibrating the reservoir's outflow. These changes ensure that the re-delineated model benefits from higher-quality data inputs.

In the WWH model, the entire Harsha Lake catchment was represented as a single sub-basin. Through the re-delineation process, this catchment was split into five small and hydrologically meaningful sub-basins allowing for a more accurate representation of the spatial variability in hydrological processes across the study area. The WWH model previously relied on a gauge located downstream of Harsha Lake, which introduced challenges on understanding upstream streamflow variability due to the damping effects of the lake. In the re-delineated model, the gauge at Perintown, located upstream of the lake, was used instead. This adjustment provides two distinct advantages: (a) streamflow data are now free from the effects of lake storage and regulation, ensuring that the upstream hydrological dynamics are better captured, and (b) if information on the lake's management scheme was available, it could facilitate the accurate representation of the lake's outflow, further enhancing model accuracy. Additionally, Harsha Lake itself, previously unresolved in the WWH model, was resolved as an autonomous sub-basin in the re-delineated setup. This inclusion ensures the lake's hydrological influence is explicitly modelled, improving the overall representation of the system.

5.2. LSHM local customization

5.2.1. Model performance across variables, metrics and periods

We next focus on the re-delineated local hydrological models and assess the streamflow performance enhancement (in terms of KGE) achieved by adjusting model parameters using local in-situ data and EOs. Performance is evaluated for all behavioural models using various combinations of variables, metrics and periods, as explained in section 4.3. We aim to identify optimal model configurations achieving high performance and hence provide insight into the effects of incorporating additional data sources on model accuracy. The performance analysis highlights distinct trends across the calibration and validation periods for the re-delineated case study model setups.

Results from the “VMP-1 setups” highlight the caution needed on the implementation of the model identification framework (Fig. 5). Poor streamflow performances can easily be achieved when different metrics and/or variables are applied, and hence it is important that the selection of metrics/variables is fit-for-purpose. As expected, the models for both case studies achieve high KGE values when the model identification is done for KGE and streamflow as a target variable (see QOB.k1 and QOB.k2 for calibration and validation periods, respectively). Overall, for both case studies, the median performances increase with the number of combined metrics (see Fig. 5 for up to 3 VMP combinations and Fig. A1 in the Appendix for 4–6 combinations). In particular, for all “VMP-3 setups”, the behavioural models outperform the default WWH for both calibration and validation periods. It is important also to note that the variability in the KGE values (represented by the box extents) is significantly reduced with increased combinations. This indicates the models’ robustness, hence the capability to represent different fluxes and signatures, due to the information content extracted in these data and metric combinations. In the case that the VMP setups consist of QOB and other metrics, the overall performance is still similar to QOB.k which implies that the additional metrics do not lead to higher information extraction and hence increased performance. Although this indicates an effective local customization, it is not necessary that a “good model for the right reasons” has been achieved which results from accounting multiple variables, for instance in the “VMP-3 setups” sets consisting of QOB and PET or AET. Although the KGE performance declined, the models seem to be robust (in terms of performance variability) and still performing well in terms of streamflow.

5.2.2. Model performance across signatures

We next focus on investigating specific attributes of the streamflow signal and visually assess the model’s ability to represent the

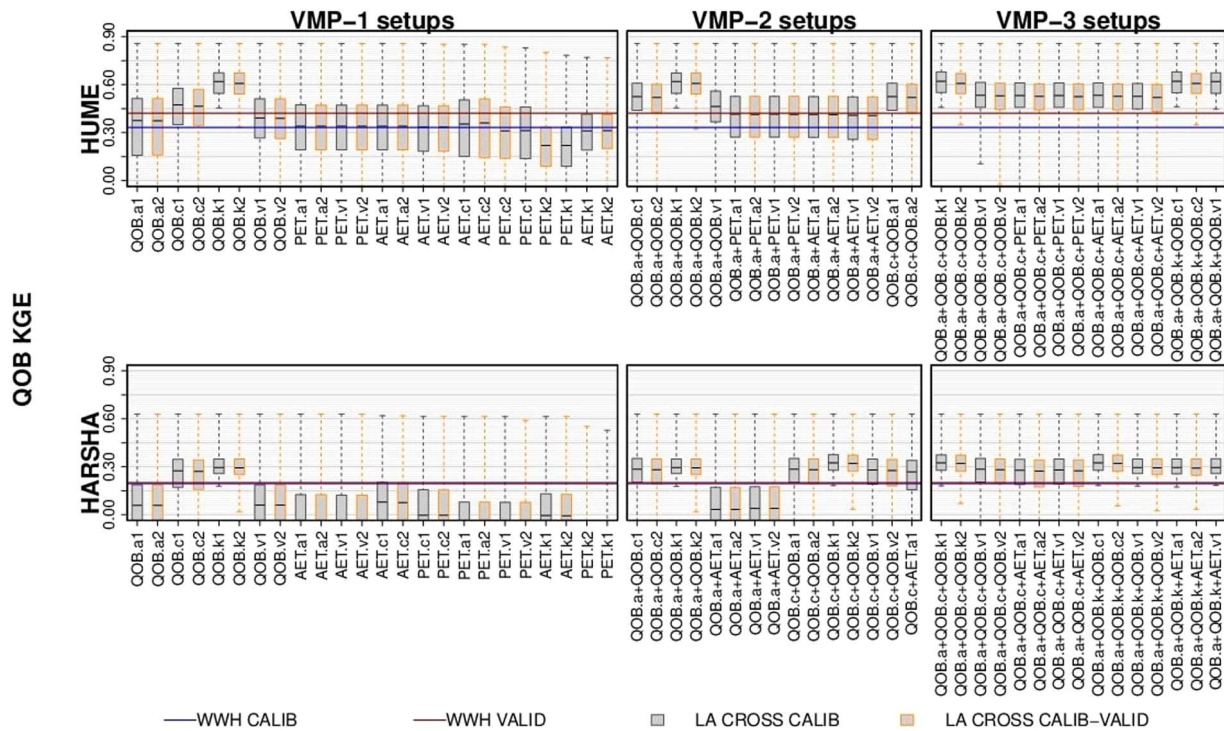


Fig. 5. Model performance for up to 3 variable-metric-period (VMP) combinations based on the identified locally adaptive (LA) model for Lake Hume (top; unregulated Jingellic station) and Harsha Lake (bottom; unregulated Perintown station). The horizontal lines represent the performance of the default WWH model in the calibration (blue) and validation (red) periods. The data sets QOB, AET and PET are the streamflow and the actual and potential evapotranspiration, respectively, while the metrics a, c, k and v are the absolute relative error, correlation coefficient, KGE and ratio of standard deviations in the simulations and observations, respectively. The values 1 and 2 at the end of the combination name indicate the calibration and validation periods, respectively.

observed flow duration curve, intra-annual variability, and biases at different streamflow ranges (extracted from the 1:1 plots). Fig. 6 shows notable differences between WWH and the locally adapted multi-objective models LA-MO (the calibration period combinations QOB.a+QOB.c+QOB.k + AET.a+PET.a and QOB.a+QOB.c+QOB.k + AET.a+AET.v for Lake Hume and Harsha Lake, respectively) across various streamflow percentiles. Those LA-MO models were used here due to both their high median streamflow KGE and capability of a “good model for the right reasons”, as explained above. Overall, the shape of the flow duration curve for LA-MO for both study areas aligned more closely with observations than for the WWH model, underscoring the improvements achieved through local customizations. In particular, WWH overestimates the high extremes in Lake Hume, yet a systematic underestimation of streamflow is observed throughout the other percentiles including the low extremes. Conversely, LA-MO can much better represent the flow duration curve, with only a small underestimation of high extremes and an underestimation of streamflow less than the 20th percentile (>80 % of time streamflow is exceeded). In Harsha Lake, both models adequately captured high extremes, but the default WWH model overestimated low extremes, whereas LA-MO underestimated them. With regard to the slope in the flow duration curve (defined as the slope between the 20 % and 70 % streamflow exceedance probability), LA-MO outperformed WWH in both study areas. For Harsha Lake, WWH underestimated high extremes but overestimated streamflow below the 20th percentile. The local customization resulted in a correction of the flat slope in the flow duration curve, although it slightly underestimated the high and low extremes.

In addition, both WWH and LA-MO could follow the seasonal patterns of observed monthly mean streamflow in both study areas; however, biases are observed in the months of high and low streamflow conditions. Regarding magnitude, WWH consistently underestimates observations during all months in Lake Hume and during high conditions (January–May) in Harsha Lake. The local customization improved the estimation of high streamflow in Harsha Lake, but led to an overestimation of low streamflow. This most probably indicates structural limitations particularly with the models’ capability to evapotranspire during the low streamflow conditions, hence achieving a more accurate water balance. The scatter plot analysis further illustrates potential biases in these models. The local customization overall reduced the spread in simulated daily streamflow and better clustered them around observed values, particularly for Lake Hume. For Harsha Lake, WWH showed less spread at high streamflow magnitudes, yet LA-MO still exhibited improvements, especially at low magnitudes. In terms of consistent biases, LA-MO for Lake Hume closely matched observed values but tended to underestimate high extremes. For Harsha Lake, the calibrated model performed well at low magnitudes, but still underestimated high extremes, although it showed improvement compared to the default WWH model.

5.2.3. Improving the spatial consistency

We next address the potential of the local customization to improve the evapotranspiration fluxes across the entire study areas. Overall, Fig. 7 shows that significant improvements in terms of KGE were achieved through local customization for both AET and PET;

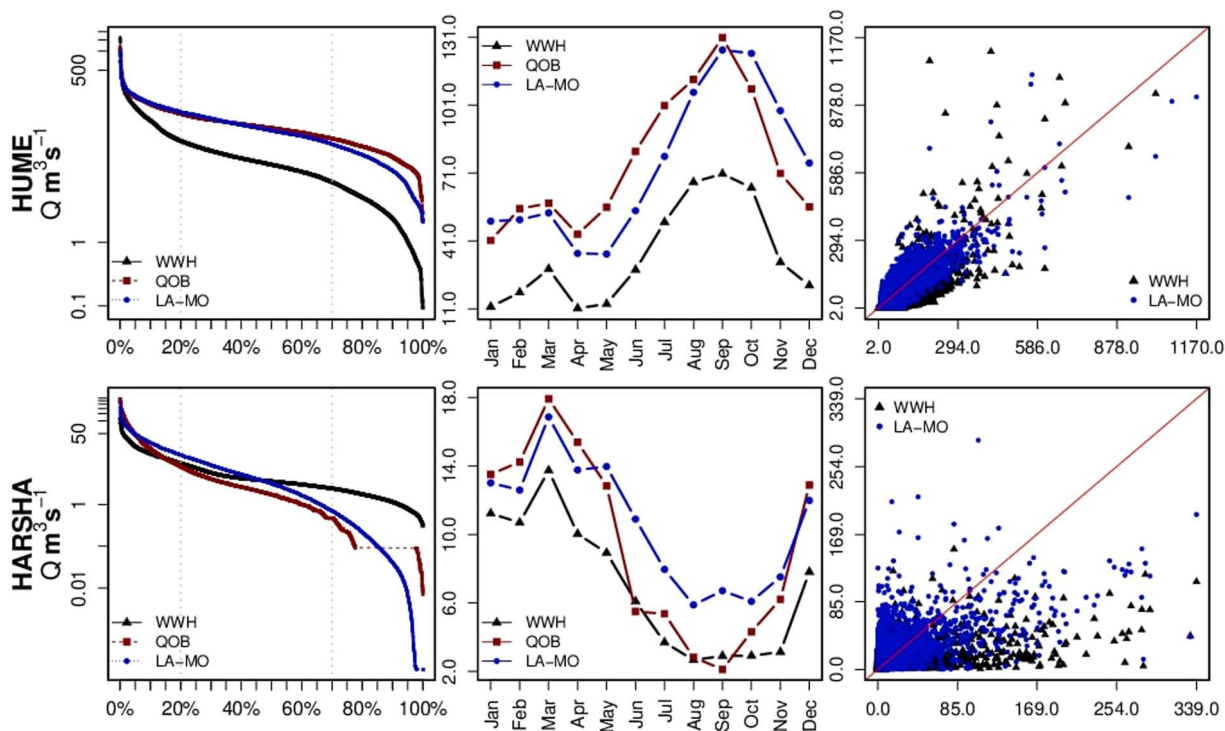


Fig. 6. The flow duration curve (first column), monthly mean streamflow (middle column) and scatter plot of daily streamflow (third column, with observations on the horizontal axis) for the unregulated Jingellic (top) and Perintown (bottom) stations in Lake Hume and Harsha Lake, respectively. The maroon colour represents the observed streamflow (QOB), black the default WWH model, and blue the locally adapted multi-objective model (LA-MO).

however, the extent of the improvements varied between the two study areas. In Lake Hume, LA-MO resulted in greater improvements in AET over the entire domain, while in Harsha Lake, PET showed the most significant gains in performance. These differences can be attributed to the distinct climatic zones of the two regions. Particularly, Lake Hume lies in the temperate climate zone and is characterized by distinct rainy seasons, perennial vegetation, and high temperatures. In this environment, transpiration is primarily constrained by the vegetation's ability to extract water from the soil, making accurate AET modelling critical for capturing hydrological dynamics. In contrast, Harsha Lake lies in a humid continental climate zone with relatively low temperatures, clear seasonal distinctions and all-year-round precipitation. Here, transpiration is limited during the summer months by high temperatures, leading plants to transpire at rates close to PET. Consequently, improving PET estimates had a more pronounced impact on model performance in this region. The spatial patterns of evapotranspiration fluxes also reveal key differences between the two catchments. In Lake Hume, AET estimates are generally higher in the headwaters, consistent with its climate, where vegetation actively transpires during the rainy season. Conversely, in Harsha Lake, AET values are lower in the headwater regions. This inadequate representation of AET likely contributes to the overestimation of monthly streamflow by the customized model seen in Fig. 6. Addressing these discrepancies requires a more detailed calibration approach that incorporates local soil-vegetation-atmosphere interactions to better capture the hydro-meteorological processes governing evapotranspiration fluxes.

5.2.4. Impact of reservoir management on hydrological response

Finally, we investigate the model performance for the Tallandoon station downstream of the regulated Dartmouth reservoir in Lake Hume aiming to assess the impact of including reservoir management on model performance. The inclusion of reservoir management significantly improved the model performance at the Tallandoon station demonstrating the value of incorporating reservoir management practices into hydrological models, particularly in systems with significant anthropogenic controls (Fig. 8). The WWH model consistently underestimated streamflow along the flow duration curve, while the LA-MO overestimated streamflow. The incorporation of reservoir management routines in LA-MO (LA-MO-RES model) resolved the overestimation of high streamflow, aligning the simulations closely with observations, whilst at the low extremes, LA-MO-RES only showed a slight underestimation. The slope and magnitude of the flow duration curve further underscore the importance of reservoir management. The WWH model produced a slope aligned with the observations, but its magnitude was inaccurate. LA-MO-RES significantly improved both the slope and magnitude highlighting the ability of reservoir management to balance high and low streamflow dynamics.

Seasonal patterns in monthly mean streamflow were also much better captured with the inclusion of reservoir management routines. WWH underestimated monthly streamflow for all months except July to September, while LA-MO significantly overestimated monthly streamflow from May to November. LA-MO-RES corrected the LA-MO biases, bringing the simulated monthly streamflow closer to observations. Moreover, as shown in the scatterplot, both WWH and LA-MO have wide deviations from daily observations, with the former model underestimating and the latter model overestimating streamflow. LA-MO-RES significantly reduced the spread, aligning simulated streamflow more closely with the 1:1 line. This improvement indicates enhanced consistency and accuracy in

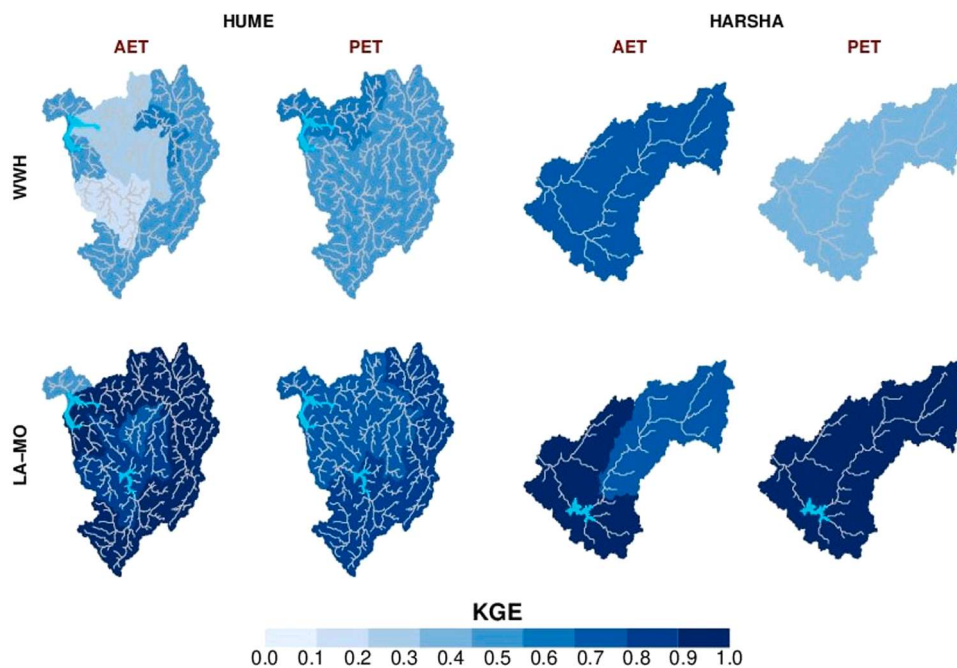


Fig. 7. AET and PET performance in terms of KGE at daily timestep using the default WWH model (top; not parameterized to account for AET and PET performance) and the locally adapted multi-objective LA-MO models (bottom) for the Lake Hume (first two columns) and Harsha Lake (third and fourth column) case studies. NASA MODIS AET and PET products are used as reference.

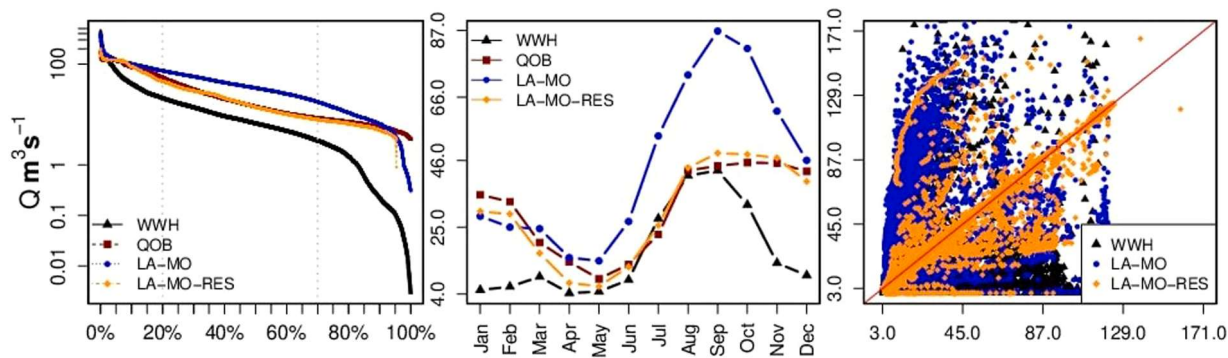


Fig. 8. The flow duration curve, monthly mean streamflow and scatter plot of daily streamflow (observations on the horizontal axis) for the Tallandoon station downstream of the regulated Dartmouth reservoir in Lake Hume. The maroon colour represents the observed streamflow (QOB), black the default WWH model, blue the locally adapted multi-objective model (LA-MO), and orange the locally adapted multi-objective model with reservoir calibration (LA-MO-RES).

streamflow, particularly for extreme conditions, due to the explicit representation of reservoir operations.

6. Discussion

6.1. Customized LSHM operationalization

The final pillar of the framework targets the operationalization of the customized LSHM further impacting local decision-making. Here we argue that customizing local LSHMs for operational services offers a pragmatic approach to enhancing hydrological predictions and services while addressing practical constraints in model development and implementation. Developing new hydrological models and operational services is generally time-intensive and resource-demanding, making it often impractical to build entirely new systems for local applications. By leveraging an existing LSHM setup, such as the one of the WWH model, the time and effort required for service development can be reduced, allowing the focus to shift to fine-tuning and optimizing the model for specific local contexts. Using an established LSHM as a reference also enables a more efficient workflow as model developers can refine parameters, integrate local data, and adjust outputs to meet operational requirements with minimal changes to the core model structure.

An additional advantage of using LSHM-based systems lies in their reliance on open datasets. This feature is particularly beneficial in regions where local data are classified, confidential or unavailable. The WWH model for example employs open datasets for its baseline configuration, allowing customization to proceed even in data-scarce environments. In these cases, operationalization can be maintained through reliance on openly accessible data sources, such as EOs and GRDC stations which, while of potentially lower quality than local in-situ data, still provide the necessary input for viable hydrological predictions. This approach not only expands the applicability of LSHM-based services in data-scarce regions, but also allows for the operational integration of EO data that, although inherently variable in quality, are available at spatial and temporal scales suitable for local customization.

The proposed customization of local LSHMs can be of high value to the United Nations Early Warnings for All (EW4All) initiative (<https://www.un.org/en/climatechange/early-warnings-for-all>). EW4All highlights the urgent need to expand early warning systems globally, with the goal of ensuring comprehensive coverage that protects all populations from hydro-meteorological hazards. Many nations and regions currently lack dedicated early warning systems, creating a demand for adaptable and scalable solutions. The LSHM-based customization approach presents a viable strategy for supporting the EW4All initiative by enabling a rapid setup of local services derived from LSHM systems. By using an existing LSHM framework, regions without established early warning systems can deploy a foundational model that is subsequently customized with locally relevant parameters, open EO datasets, and other available data streams. This streamlined approach enables faster implementation and provides a critical tool for delivering timely and actionable warnings, thereby contributing to the global effort to build resilience against natural hazards.

Finally, we focused here on the local customization of LSHMs, yet this raises important considerations about how such refinements can be integrated into large-scale modelling efforts. LSHMs are typically designed to perform consistently across multiple basins and diverse hydrological regimes, which often necessitates generalizations that overlook site-specific processes. Consequently, scaling up local improvements is not always straightforward. To bridge this gap, there is a need to develop frameworks that enable LSHMs to evolve through the integration of local data and knowledge without compromising their large-scale performance. For instance, local customizations such as model re-delineation or the integration of high-quality in-situ data could support the refinement of datasets used in LSHM identification. Overall, the scalability of local improvements depends on the structural flexibility of the model. In the case of WWH, parameterization is based on land use classes and elevation zones, which enables regionalization but does not easily accommodate site-specific parameter tuning linked to individual stations. Therefore, while local insights are invaluable, flexible and adaptive frameworks are needed to reconcile them with the design principles of LSHMs, enabling continuous model evolution that benefits both large-scale applicability and local relevance.

6.2. Customizing local LSHMs using additional innovative data sources

To advance the customization of local LSHMs, integrating “non-traditional” data sources offers promising avenues for improving model accuracy and enhancing local hydrological insights. While this study has demonstrated the value of in-situ data and EOs from MODIS, additional hydro-meteorological variables available over extended periods and at higher spatiotemporal resolutions are essential. Such key variables are soil moisture, snow water equivalent, fractional snow cover, total water storage anomalies, water extents and vegetation indices. These are increasingly becoming available from sources like the European Space Agency Climate Change Initiative (ESA CCI) and NASA at resolutions suitable for more targeted LSHM customization. In particular, snow and soil moisture have shown potential as critical predictive variables, especially in snow-dominated or high-latitude regions (Musuuza et al., 2023; Wanders et al., 2019). Integrating soil moisture and groundwater storage anomaly data could improve the model’s representation of antecedent conditions and seasonal hydrological shifts, especially in regions where snowpack dynamics heavily influence hydrological responses.

Effective parameter identification within this multi-variable framework relies on EO data quality, which this study did not specifically address. Continued efforts to enhance EO data in terms of resolution, accuracy and historical availability are essential, as higher-quality inputs directly strengthen parameter identification and boost model performance. Currently, the absence of uncertainty quantification in EO products limits the efficacy of the multi-variable parameter identification approach in river systems where data reliability may fluctuate, such as high-latitude regions or complex terrains with challenging snow and evapotranspiration dynamics. Incorporating uncertainty information into EO products could enable the framework to assign diagnostic weights, thereby improving model calibration and performance in areas where data quality is unreliable. Furthermore, using the same variable from multiple EO providers could address variability in data quality and improve the robustness of model simulations. For example, assessing different sources of evapotranspiration data, such as GLEAM (available online: <https://www.gleam.eu>) and LSA-SAF (available online: <https://landsaf.ipma.pt/en>) (both accessed on 09.01.2025), could uncover biases related to specific data sources and help evaluate potential model improvements across diverse environmental conditions. This approach, especially valuable where certain EO sources may lack accuracy or availability, could establish a benchmark for EO data quality assessment, aiding in the selection of optimal data sources for effective LSHM customization.

Beyond EO data, further innovative data sources offer additional customization potential, especially for meeting specific local modelling needs. Low-cost monitoring systems, such as affordable sensor networks, present a viable solution for expanding data coverage, particularly in remote or under-monitored regions. These systems can capture real-time hydro-meteorological variables, enhancing the spatiotemporal detail needed for localized modelling in areas where EOs or conventional ground-based observations are sparse. Their accessibility allows for cost-effective data acquisition over extended periods, thus providing continuous data streams that can be readily assimilated into LSHMs. Unmanned Aerial Vehicles (UAVs, or drones) have also revolutionized localized hydrological data collection, enabling high-resolution aerial surveys of hydrological features and landscapes (Zhou et al., 2024). UAVs can capture detailed measurements of snow cover, vegetation, soil moisture, and surface water extents, providing insights that are challenging to obtain from satellite imagery alone. In floodplains or catchments with rapid topographical changes, UAVs data can aid in understanding the spatial dynamics of river systems, thereby refining model calibration in these dynamic landscapes.

Moreover, citizen science initiatives bring a novel and valuable dimension to hydrological data collection. By involving local communities, citizen science can produce vital ground-level observations (e.g., rainfall measurements, snowpack and streamflow monitoring), particularly in areas lacking conventional monitoring infrastructure (Nardi et al., 2022). Citizen-collected data allow for high-frequency observations, especially during extreme weather events, and foster a collaborative approach that enhances the model’s representativeness of local hydrological processes. While citizen science offers unique benefits, integrating these data into LSHMs requires robust quality assurance protocols to ensure reliability. Moving forward, a combined approach that leverages diverse data sources will better support parameter refinement and model setup, ensuring LSHMs are equipped to meet local hydrological needs providing substantial benefits for real-time applications such as flood risk assessment and optimized water resource management.

7. Conclusions

Herein we highlight the importance of customizing LSHMs to improve their applicability at the local scale, addressing limitations in representing local hydrological processes. Using the global WWH model as a benchmark, the study demonstrates significant improvements in modelling streamflow and evapotranspiration (both actual and potential) when LSHMs are refined to include local features such as lakes and reservoirs and region-specific hydrological processes. The findings emphasize the value of integrating local data and knowledge, re-delineating model setups and re-calibrating parameters, even in data-limited regions, such as the Lake Hume (Australia) and Harsha Lake (USA), to enhance model performance across multiple metrics and signatures. We highlight the potential of combining global hydrological models with localized refinements to improve decision-making in water resource management. By leveraging EOs and local data and knowledge, LSHMs can be effectively downscaled to meet user needs, offering practical tools for enhancing operational hydrological services and building resilience to hydrological challenges. Hence, the findings demonstrate the transformative potential of localized model customization in addressing global hydrological modelling challenges and paving the way for more accurate, adaptable and user-centred hydrological services.

The main conclusions from this study are as follows:

- The re-delineation of the Lake Hume and Harsha Lake catchments demonstrates the critical role of integrating local knowledge in refining LSHMs. Addressing limitations in sub-basin delineation, gauge placement, and reservoir integration significantly enhanced

the accuracy of hydrological system representation, providing a robust foundation for operational hydrological services and decision-making.

- EOs offer a unique opportunity to improve hydrological modelling by enabling spatially distributed assessments, which are particularly valuable in regions with limited in-situ observations. The study highlights the ability of EOs to support parameter identification, not only for streamflow but also for other critical fluxes, underlining their potential to fill data gaps in regions where in-situ data are unavailable.
- Combining multiple variables and metrics during model identification significantly improved streamflow performance (in terms of KGE) and robustness for both case studies compared to the benchmark WWH model. The sets of the "VMP-3 setups" (using different metrics for QOB, PET and AET) reduced performance variability and enabled better representation of diverse hydrological processes. These results highlight the importance of tailored metric and variable selection for effective local customization.
- Improvements in the spatial consistency of evapotranspiration fluxes underscore the importance of tailoring hydrological models to account for regional climatic conditions. Differences in dominant evapotranspiration processes between Lake Hume and Harsha Lake emphasize the necessity of localized customization to address region-specific limitations in AET and PET representation. These adjustments improved model accuracy and reduced biases in hydrological outputs, such as streamflow.
- Incorporating reservoir characteristics and management practices into the locally adapted multi-objective model significantly enhanced its ability to simulate regulated systems. While resolving such features in global-scale models may not always be feasible, insights from locally calibrated models can inform the parameterization of reservoirs and regulatory processes in LSHMs, bridging the gap between localized and large-scale modelling efforts.

CRedit authorship contribution statement

Pechlivanidis Ilias: Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Conceptualization. **Musuuza Jude Lubega:** Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Data curation.

Author contributions

IGP was responsible for the project management and funding acquisition and contributed to the basic idea, study design, result analysis and figures, interpretation of results, and writing of the manuscript. JLM contributed to the code development, model runs, result analysis and figures, interpretation of results, and writing of the manuscript.

Declaration of Competing Interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Ilias Pechlivanidis reports financial support and article publishing charges were provided by Horizon 2020 European Innovation Council Fast Track to Innovation. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

Financial support: This study was partially funded by the EU Horizon 2020 project Prime Water (Delivering advanced predictive tools from medium to seasonal range for water dependent industries exploiting the cross-cutting potential of EO and hydroecological modeling) under the Grant Agreement 870497. In addition, this study was partially funded by the EU Horizon 2020 project I-CISK (Innovating climate services through integrating scientific and local knowledge) under Grant Agreement 101037293.

Appendix

Here we present a summary of the dataset used in the two case studies, including their characteristics ([Table A1](#)).

Table A1
Information on the datasets used, including their time coverages and spatiotemporal resolutions

Hydrologic variable (abbreviation; unit)	Provider	Availability window	Spatial resolution [m]	Temporal resolution
Actual evapotranspiration (AET; mm)	National Aeronautics and Space Administration (NASA)	2001–2019	500	8-day aggregate
Potential evapotranspiration (PET; mm)	National Aeronautics and Space Administration (NASA)	2001–2019	500	8-day aggregate

(continued on next page)

Table A1 (continued)

Hydrologic variable (abbreviation; unit)	Provider	Availability window	Spatial resolution [m]	Temporal resolution
Streamflow (QOB; m ³ ·s ⁻¹)	For Lake Hume: Commonwealth Scientific and Industrial Research Organisation (CSIRO)	1890–2020	-	Daily
	For Harsha Lake: United States Geological Survey (USGS)	2000–2018	-	Daily
	Global Runoff Data Centre (GRDC)	1890–2018	-	Daily

Here we present the sensitive WWH model parameters and their ranges (Table A2) that have been used in the parameter identification process.

Table A2

Sensitive WWH model parameters controlling different processes and their ranges

Process	Parameter	Description	Units	Range (Min - Max)
Evapotranspiration	lp	Soil water limit for PET	-	0.1–0.9
	epotdist	Exponent for dependency of evapotranspiration on depth	m ⁻¹	0.1–5
	kc	Crop coefficient	-	0.1–1.6
	cevpcorr	Evapotranspiration correction	mm·°C·day ⁻¹	0.00016–0.9
Recession	rrcs3	Recession coefficient for slope dependence in the upper soil layer	day ⁻¹	0.0001–0.1
	srrcs	Recession coefficient for surface runoff	day ⁻¹	0.01–1
	rrcs1	Recession coefficient for uppermost soil layer	day ⁻¹	0.01–0.75
Water storage	wcep	Effective porosity	-	0.505–0.9
	wcwp	Wilting point	-	0.1–0.394
Macropore flow	mperc1	Maximum percolation capacity from the uppermost soil layer to the next	mm·day ⁻¹	1–50

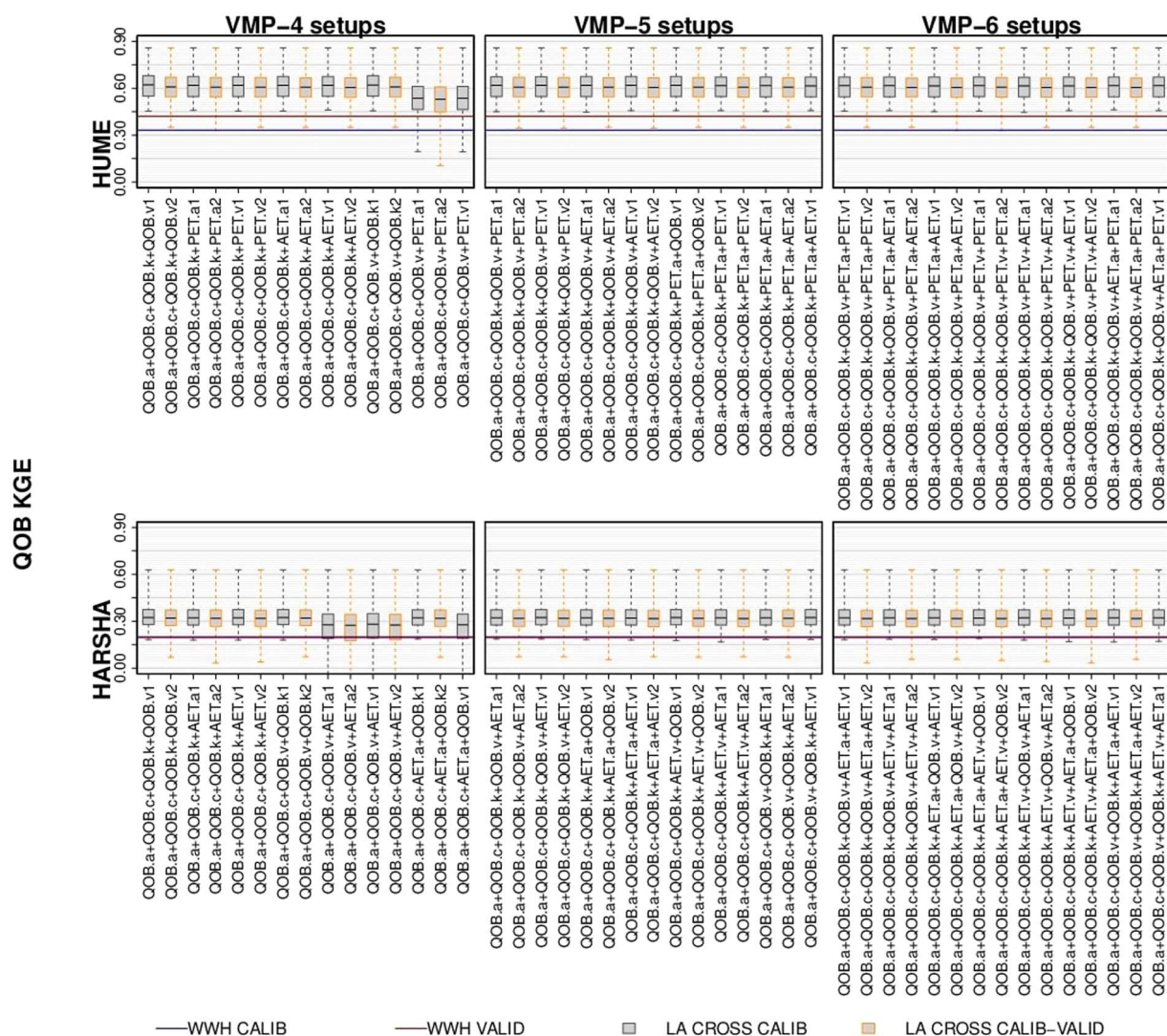


Figure A1. Model performance for 4–6 variable-metric-period (VMP) combinations based on the identified locally adaptive (LA) model for Lake Hume (top; unregulated Jingellic station) and Harsha Lake (bottom; unregulated Perintown station). The horizontal lines represent the performance of the default WWH model in the calibration (blue) and validation (red) periods. The data sets QOB, AET and PET are the streamflow and the actual and potential evapotranspiration, respectively, while the metrics a, c, k and v are the absolute relative error, correlation coefficient, KGE and ratio of standard deviations in the simulations and observations, respectively. The values 1 and 2 at the end of the combination name indicate the calibration and validation periods, respectively

Data availability

Data will be made available on request.

References

- Alfieri, L., Avanzi, F., Delogu, F., Gabellani, S., Bruno, G., Campo, L., Libertino, A., Massari, C., Tarpanelli, A., Rains, D., Miralles, D.G., Quast, R., Vreugdenhil, M., Wu, H., Brocca, L., 2022. High-resolution satellite products improve hydrological modeling in northern Italy. *Hydrol. Earth Syst. Sci.* 26, 3921–3939. <https://doi.org/10.5194/hess-26-3921-2022>.
- Andersson, J.C.M., Olsson, J., Van De Beek, R. (C.Z.), Hansryd, J., 2022. OpenMRG: open data from microwave links, radar, and gauges for rainfall quantification in Gothenburg, Sweden. *Earth Syst. Sci. Data* 14, 5411–5426. <https://doi.org/10.5194/essd-14-5411-2022>.
- Arheimer, B., Pimentel, R., Isberg, K., Crochemore, L., Andersson, J.C.M., Hasan, A., Pineda, L., 2020. Global catchment modelling using World-Wide HYPE (WWH), open data and stepwise parameter estimation. *Hydrol. Earth Syst. Sci.* 24, 535–559. <https://doi.org/10.5194/hess-24-535-2020>.
- Beck, H.E., Pan, M., Lin, P., Seibert, J., Van Dijk, A.I.J.M., Wood, E.F., 2020. Global fully distributed parameter regionalization based on observed streamflow from 4,229 headwater catchments. *JGR Atmos.* 125, e2019JD031485. <https://doi.org/10.1029/2019JD031485>.

- Beck, H.E., Van Dijk, A.I.J.M., Levizzani, V., Schellekens, J., Miralles, D.G., Martens, B., De Roo, A., 2017. MSWEP: 3-hourly 0.25° global gridded precipitation (1979–2015) by merging gauge, satellite, and reanalysis data. *Hydrol. Earth Syst. Sci.* 21, 589–615. <https://doi.org/10.5194/hess-21-589-2017>.
- Berg, P., Almén, F., Bozhinova, D., 2021. HydroGFD3.0 (Hydrological Global Forcing Data): A 25 km global precipitation and temperature data set updated in near-real time. *Earth Syst. Sci. Data* 13, 1531–1545. <https://doi.org/10.5194/essd-13-1531-2021>.
- Beven, K., Binley, A., 1992. The future of distributed models: model calibration and uncertainty prediction. *Hydrol. Process.* 6, 279–298. <https://doi.org/10.1002/hyp.3360060305>.
- Beven, K., Freer, J., 2001. Equifinality, data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems using the GLUE methodology. *J. Hydrol.* 249, 11–29. [https://doi.org/10.1016/S0022-1694\(01\)00421-8](https://doi.org/10.1016/S0022-1694(01)00421-8).
- Bierkens, M.F.P., Bell, V.A., Burek, P., Chaney, N., Condon, L.E., David, C.H., de Roo, A., Döll, P., Drost, N., Famiglietti, J.S., Flörke, M., Gochis, D.J., Houser, P., Hut, R., Keune, J., Kollet, S., Maxwell, R.M., Reager, J.T., Samaniego, L., Sudicky, E., Sutanudjaja, E.H., van de Giesen, N., Winsemius, H., Wood, E.F., 2015. Hyper-resolution global hydrological modelling: what is next? “Everywhere and locally relevant.”. *Hydrol. Process.* 29, 310–320. <https://doi.org/10.1002/hyp.10391>.
- Brendel, C., Capell, R., Bartosova, A., 2023. To tame a land: Limiting factors in model performance for the multi-objective calibration of a pan-European, semi-distributed hydrological model for discharge and sediments. *J. Hydrol.: Reg. Stud.* 50, 101544. <https://doi.org/10.1016/j.ejrh.2023.101544>.
- Clark, M.P., Bierkens, M.F.P., Samaniego, L., Woods, R.A., Uijlenhoet, R., Bennett, K.E., Pauwels, V.R.N., Cai, X., Wood, A.W., Peters-Lidard, C.D., 2017. The evolution of process-based hydrologic models: historical challenges and the collective quest for physical realism. *Hydrol. Earth Syst. Sci.* 21, 3427–3440. <https://doi.org/10.5194/hess-21-3427-2017>.
- Clemenzi, I., Gustafsson, D., Marchand, W.-D., Norell, B., Zhang, J., Pettersson, R., Allan Pohjola, V., 2023. Impact of snow distribution modelling for runoff predictions. *Hydrol. Res.* 54, 633–647. <https://doi.org/10.2166/nh.2023.043>.
- Copernicus Climate Change Service, 2020. Near surface meteorological variables from 1979 to 2018 derived from bias-corrected reanalysis. <https://doi.org/10.24381/CDS.20D54E34>.
- Dasgupta, A., Arnal, L., Emerton, R., Harrigan, S., Matthews, G., Muhammad, A., O'Regan, K., Pérez-Ciria, T., Valdez, E., Van Osnabrugge, B., Werner, M., Buontempo, C., Cloke, H., Pappenberger, F., Pechlivanidis, I.G., Prudhomme, C., Ramos, M., Salamon, P., 2025. Connecting hydrological modelling and forecasting from global to local scales: perspectives from an international joint virtual workshop. *J. Flood Risk Manag.* 18, e12880. <https://doi.org/10.1111/jfr3.12880>.
- Davy, M.C., Koen, T.B., 2013. Variations in soil organic carbon for two soil types and six land uses in the Murray Catchment, New South Wales, Australia. *Soil Res* 51, 631. <https://doi.org/10.1071/SR12353>.
- De Lannoy, G.J.M., Bechtold, M., Albergel, C., Brocca, L., Calvet, J.-C., Carrassi, A., Crow, W.T., De Rosnay, P., Durand, M., Forman, B., Geppert, G., Girotto, M., Hendricks Franssen, H.-J., Jonas, T., Kumar, S., Lievens, H., Lu, Y., Massari, C., Pauwels, V.R.N., Reichle, R.H., Steele-Dunne, S., 2022. Perspective on satellite-based land data assimilation to estimate water cycle components in an era of advanced data availability and model sophistication. *Front. Water* 4, 981745. <https://doi.org/10.3389/frwa.2022.981745>.
- Du, T.L.T., Lee, H., Bui, D.D., Graham, L.P., Darby, S.D., Pechlivanidis, I.G., Leyland, J., Biswas, N.K., Choi, G., Batelaan, O., Bui, T.T.P., Do, S.K., Tran, T.V., Nguyen, H.T., Hwang, E., 2022. Streamflow prediction in highly regulated, transboundary watersheds using multi-basin modeling and remote sensing imagery. *e2021WR031191 Water Resour. Res.* 58. <https://doi.org/10.1029/2021WR031191>.
- Girons Lopez, M., Bosshard, T., Crochemore, L., Pechlivanidis, I.G., 2025. Leveraging GCM-based forecasts for enhanced seasonal streamflow prediction in diverse hydrological regimes. *J. Hydrol.* 650, 132504. <https://doi.org/10.1016/j.jhydrol.2024.132504>.
- Girons Lopez, M., Crochemore, L., Pechlivanidis, I.G., 2021. Benchmarking an operational hydrological model for providing seasonal forecasts in Sweden. *Hydrol. Earth Syst. Sci.* 25, 1189–1209. <https://doi.org/10.5194/hess-25-1189-2021>.
- Guo, Y., Zhang, Y., Zhang, L., Wang, Z., 2021. Regionalization of hydrological modeling for predicting streamflow in ungauged catchments: a comprehensive review. *WIREs Water* 8, e1487. <https://doi.org/10.1002/wat2.1487>.
- Gupta, H.V., Kling, H., Yilmaz, K.K., Martinez, G.F., 2009. Decomposition of the mean squared error and NSE performance criteria: implications for improving hydrological modelling. *J. Hydrol.* 377, 80–91. <https://doi.org/10.1016/j.jhydrol.2009.08.003>.
- Gupta, H.V., Perrin, C., Blöschl, G., Montanari, A., Kumar, R., Clark, M., Andréassian, V., 2014. Large-sample hydrology: a need to balance depth with breadth. *Hydrol. Earth Syst. Sci.* 18, 463–477. <https://doi.org/10.5194/hess-18-463-2014>.
- Haylock, M.R., Hofstra, N., Klein Tank, A.M.G., Klok, E.J., Jones, P.D., New, M., 2008. A European daily high-resolution gridded data set of surface temperature and precipitation for 1950–2006. *J. Geophys. Res.* 113, 2008JD010201. <https://doi.org/10.1029/2008JD010201>.
- Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., De Chiara, G., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R.J., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., De Rosnay, P., Rozum, I., Vamborg, F., Villaume, S., Thépaut, J., 2020. The ERA5 global reanalysis. *Quart. J. R. Meteor. Soc.* 146, 1999–2049. <https://doi.org/10.1002/qj.3803>.
- Hrachowitz, M., Savenije, H.H.G., Blöschl, G., McDonnell, J.J., Sivapalan, M., Pomeroy, J.W., Arheimer, B., Blume, T., Clark, M.P., Ehret, U., Fenicia, F., Freer, J.E., Gelfan, A., Gupta, H.V., Hughes, D.A., Hut, R.W., Montanari, A., Pande, S., Tetzlaff, D., Troch, P.A., Uhlenbrook, S., Wagener, T., Winsemius, H.C., Woods, R.A., Zehe, E., Cudennec, C., 2013. A decade of Predictions in Ungauged Basins (PUB)—a review. *Hydrol. Sci. J.* 58, 1198–1255. <https://doi.org/10.1080/02626667.2013.803183>.
- Hundecha, Y., Arheimer, B., Berg, P., Capell, R., Musuza, J.L., Pechlivanidis, I., Photiadou, C., 2020. Effect of model calibration strategy on climate projections for hydrological indicators at a continental scale. *Clim. Change.* <https://doi.org/10.1007/s10584-020-02874-4>.
- Hundecha, Y., Arheimer, B., Donnelly, C., Pechlivanidis, I., 2016. A regional parameter estimation scheme for a pan-European multi-basin model. *J. Hydrol.: Reg. Stud.* 6. <https://doi.org/10.1016/j.ejrh.2016.04.002>.
- Kauffeldt, A., Halldin, S., Rodhe, A., Xu, C.-Y., Westerberg, I.K., 2013. Disinformative data in large-scale hydrological modelling. *Hydrol. Earth Syst. Sci.* 17, 2845–2857. <https://doi.org/10.5194/hess-17-2845-2013>.
- Khaki, M., Hoteit, I., Kuhn, M., Forootan, E., Awange, J., 2019. Assessing data assimilation frameworks for using multi-mission satellite products in a hydrological context. *Sci. Total Environ.* 647, 1031–1043. <https://doi.org/10.1016/j.scitotenv.2018.08.032>.
- Klemes, V., 1986. Operational testing of hydrological simulation models. *Hydrol. Sci. J.* 31, 13–24. <https://doi.org/10.1080/02626668609491024>.
- Knoben, W.J.M., Woods, R.A., Freer, J.E., 2018. A quantitative hydrological climate classification evaluated with independent streamflow data. *Water Resour. Res.* 54, 5088–5109. <https://doi.org/10.1029/2018WR022913>.
- Kouki, K., Räisänen, P., Luojus, K., Luomaranta, A., Riihelä, A., 2022. Evaluation of Northern Hemisphere snow water equivalent in CMIP6 models during 1982–2014. *Cryosphere* 16, 1007–1030. <https://doi.org/10.5194/tc-16-1007-2022>.
- Kuentz, A., Arheimer, B., Hundecha, Y., Wagener, T., 2017. Understanding hydrologic variability across Europe through catchment classification. *Hydrol. Earth Syst. Sci.* 21, 2863–2879. <https://doi.org/10.5194/hess-21-2863-2017>.
- Kumar, R., Livneh, B., Samaniego, L., 2013. Toward computationally efficient large-scale hydrologic predictions with a multiscale regionalization scheme. *Water Resour. Res.* 49, 5700–5714. <https://doi.org/10.1002/wrcr.20431>.
- Mao, F., Khamis, K., Clark, J., Krause, S., Buytaert, W., Ochoa-Tocachi, B.F., Hannah, D.M., 2020. Moving beyond the technology: a socio-technical roadmap for low-cost water sensor network applications. *Environ. Sci. Technol.* 54, 9145–9158. <https://doi.org/10.1021/acs.est.9b07125>.
- Mu, Q., Zhao, M., Running, S.W., 2011. Improvements to a MODIS global terrestrial evapotranspiration algorithm. *Remote Sens. Environ.* 115, 1781–1800. <https://doi.org/10.1016/j.rse.2011.02.019>.
- Musuza, J.L., Crochemore, L., Pechlivanidis, I.G., 2023. Evaluation of earth observations and in situ data assimilation for seasonal hydrological forecasting. *Water Resour. Res.* 59, e2022WR033655. <https://doi.org/10.1029/2022WR033655>.

- Musuza, J.L., Gustafsson, D., Pimentel, R., Crochemore, L., Pechlivanidis, I., 2020. Impact of satellite and in situ data assimilation on hydrological predictions. *Remote Sens.* 12, 811. <https://doi.org/10.3390/RS12050811>.
- Nardi, F., Cudennec, C., Abrate, T., Allouch, C., Annis, A., Assumpção, T., Aubert, A.H., Béro, D., Braccini, A.M., Buytaert, W., Dasgupta, A., Hannah, D.M., Mazzoleni, M., Polo, M.J., Sæbø, Ø., Seibert, J., Tauro, F., Teichert, F., Teutonico, R., Uhlenbrook, S., Wahrman Vargus, C., Grimaldi, S., 2022. Citizens AND HYdrology (CANDHY): conceptualizing a transdisciplinary framework for citizen science addressing hydrological challenges. *Hydrol. Sci. J.* 67, 2534–2551. <https://doi.org/10.1080/02626667.2020.1849707>.
- Nijzink, R.C., Almeida, S., Pechlivanidis, I.G., Capell, R., Gustafssons, D., Arheimer, B., Parajka, J., Freer, J., Han, D., Wagener, T., van Nooijen, R.R.P., Savenije, H.H.G., Hrachowitz, M., 2018. Constraining conceptual hydrological models with multiple information sources. *Water Resour. Res.* 54, 8332–8362. <https://doi.org/10.1029/2017WR021895>.
- Pagliaro, L., Bouraoui, F., Diels, J., Willems, P., McIntyre, N., 2019. Investigating regionalization techniques for large-scale hydrological modelling. *J. Hydrol.* 570, 220–235. <https://doi.org/10.1016/j.jhydrol.2018.12.071>.
- Pandeya, B., Uprety, M., Paul, J.D., Sharma, R.R., Dugar, S., Buytaert, W., 2021. Mitigating flood risk using low-cost sensors and citizen science: a proof-of-concept study from western Nepal. *J. Flood Risk Manag.* 14, e12675. <https://doi.org/10.1111/jfr3.12675>.
- Pechlivanidis, I.G., Arheimer, B., 2015. Large-scale hydrological modelling by using modified PUB recommendations: the India-HYPE case. *Hydrol. Earth Syst. Sci.* 19, 4559–4579. <https://doi.org/10.5194/hess-19-4559-2015>.
- Pechlivanidis, I.G., Crochemore, L., Rosberg, J., Bosshard, T., 2020. What are the key drivers controlling the quality of seasonal streamflow forecasts? *Water Resour. Res.* 56, e2019WR026987. <https://doi.org/10.1029/2019wr026987>.
- Pechlivanidis, I.G., Jackson, B., McIntyre, N., Wheeler, H.S., 2011. Catchment scale hydrological modelling: a review of model types, calibration approaches and uncertainty analysis methods in the context of recent developments in technology and applications. *Glob. NEST J.* 13, 193–214.
- Pianosi, F., Sarrazin, F., Wagener, T., 2015. A matlab toolbox for global sensitivity analysis. *Environ. Model. Softw.* 70, 80–85. <https://doi.org/10.1016/j.envsoft.2015.04.009>.
- Pimentel, R., Arheimer, B., Crochemore, L., Andersson, J.C.M., Pechlivanidis, I.G., Gustafsson, D., 2023. Which potential evapotranspiration formula to use in hydrological modeling world-wide? *Water Resour. Res.* 59, e2022WR033447. <https://doi.org/10.1029/2022WR033447>.
- Rakovec, O., Mizukami, N., Kumar, R., Newman, A.J., Thober, S., Wood, A.W., Clark, M.P., Samaniego, L., 2019. Diagnostic evaluation of large-domain hydrologic models calibrated across the contiguous United States. *JGR Atmos.* 124, 13991–14007. <https://doi.org/10.1029/2019JD030767>.
- Samaniego, L., Kumar, R., Zink, M., 2013. Implications of parameter uncertainty on soil moisture drought analysis in Germany. *J. Hydrometeorol.* 14, 47–68. <https://doi.org/10.1175/JHM-D-12-075.1>.
- Schneider, U., Becker, A., Finger, P., Meyer-Christoffer, A., Ziese, M., 2018. GPCC full data monthly version 2018.0 at 0.25°: monthly land-surface precipitation from rain-gauges built on GTS-based and Historic. Data: Gridded Mon. Totals. https://doi.org/10.5676/DWD_GPCC/FD_M_V2018_025.
- Seibert, J., Van Meerveld, H.J. (Ilja), 2016. Hydrological change modeling: challenges and opportunities. *Hydrol. Process.* 30, 4966–4971. <https://doi.org/10.1002/hyp.10999>.
- Sun, Q., Miao, C., Duan, Q., Ashouri, H., Sorooshian, S., Hsu, K., 2018. A review of global precipitation data sets: data sources, estimation, and intercomparisons. *Rev. Geophys.* 56, 79–107. <https://doi.org/10.1002/2017RG000574>.
- Tang, G., Clark, M.P., Papalexiou, S.M., 2022. EM-Earth: the ensemble meteorological dataset for planet earth. *Bull. Am. Meteorol. Soc.* 103, E996–E1018. <https://doi.org/10.1175/BAMS-D-21-0106.1>.
- Thiemig, V., Gomes, G.N., Sköien, J.O., Ziese, M., Rauthe-Schöch, A., Rustemeier, E., Rehfeldt, K., Walawender, J.P., Kolbe, C., Pichon, D., Schweim, C., Salamon, P., 2022. EMO-5: a high-resolution multi-variable gridded meteorological dataset for Europe. *Earth Syst. Sci. Data* 14, 3249–3272. <https://doi.org/10.5194/essd-14-3249-2022>.
- Wagener, T., Gupta, H.V., 2005. Model identification for hydrological forecasting under uncertainty. *Stoch. Environ. Res. Risk Assess.* 19, 378–387. <https://doi.org/10.1007/s00477-005-0006-5>.
- Wanders, N., Thober, S., Kumar, R., Pan, M., Sheffield, J., Samaniego, L., Wood, E.F., 2019. Development and evaluation of a Pan-European multi-model seasonal hydrological forecasting system. *J. Hydrometeorol.* 20, 99–115. <https://doi.org/10.1175/JHM-D-18-0040.1>.
- Zhou, Z., Riis-Klinkvort, L., Jørgensen, E.A., Lindenhoff, C., Frías, M.C., Vesterhauge, A.R., Olesen, D.H., Lavish, M., Dobrovolskiy, A., Kadek, A., Orlic, N., Grubesa, T., Drmić, L., Grosen, H., Nielsen, S., Wennerberg, D., Fagerström, V., Axén, J., Gustafsson, D., Bauer-Gottwein, P., 2024. Measuring river surface velocity using UAS-borne doppler radar. *Water Resour. Res.* 60, e2024WR037375. <https://doi.org/10.1029/2024WR037375>.