03 - EVALUATING THE WFLOW HYDROLOGICAL MODEL PERFORMANCE WITH TWO DIFFERENT RAINFALL DATASETS IN THE BOYNE CATCHMENT, IRELAND

Safa Mohammed[1] and Ahmed Nasr[1]

[1]*School of Transport and Civil Engineering, TU Dublin, Dublin, Ireland*

Abstract

Physically based distributed hydrological models provide a powerful tool for short-term flow predictions and long-term water management planning. Their strength lies in their ability to capture the spatially distributed nature of rainfall-runoff processes by integrating inputs and parameters that consider the inherent spatial heterogeneity of meteorological forcing variables and catchment features such as topography, soil composition, land cover, and river networks. The present study aims to develop a physically based distributed hydrological model for the Boyne catchment in Ireland, using the wflow model. The Boyne wflow model employs a 30 arcsec (~1 km) resolution grid with parameters obtained from global remote sensing and GIS data. These include: (i) MERIT Hydro and Lake Hydro for static geospatial data; (ii) SoilGrids distributed maps for soil-related parameters; and (iii) CORINE maps for land cover-related parameters. Additionally, the performance of the Boyne wflow model has been assessed using two distinct grided rainfall datasets at spatial resolution of 0.25 km. The first dataset is derived from Met Éireann synoptic stations network, while the second dataset is ERA5 dataset provided by the European Centre for Medium-Range Weather Forecasts (ECMWF). Evaluation during calibration and validation periods was conducted using two metrics, Nash-Sutcliffe Efficiency (NSE) and Kling-Gupta Efficiency (KGE), along with visual examination of flow hydrographs. Modelling results revealed minor underestimation in peak flows when the model was forced by the ERA5 rainfall dataset compared to the model driven by the ground truth gridded rainfall data. This is evident in the average KGE values of the two models, with ERA5 data driven model at 0.75 and the ground truth gridded rainfall data driven model 0.8. The findings of this study provide an initial yet encouraging understanding of creating an accurate distributed hydrological model for Irish catchments utilising global data and satellite data. This model can be employed effectively in simulating flow patterns and predicting river floods.

1. INTRODUCTION

Over the course of recent decades, there has been significant progress in the development of hydrological models aimed at improving our comprehension of rainfall-runoff processes. These models have become indispensable tools for conducting scenario simulations, making predictions, and facilitating decision-making in the management of river basins (Cloke and Pappenberger, 2009; Wannasin *et al.*, 2021).

Ireland, characterised by its temperate climate and abundant precipitation, presents a unique hydrological landscape. The country's susceptibility to rainfall-induced flooding events underscores the critical importance of understanding the region's hydrology. The heightened focus on flood risk management was prompted by a significant flood event in the Shannon River catchment during the 2015/2016 winter. The meteorological winter of 2015/2016 in Ireland stands out as an extraordinary period, marked by the breaking of numerous climate records and the occurrence of high-impact weather events that led to significant disruptions, primarily due to flooding and high winds (Sherlock and Duffy, 2019). As a result, there is an urgent need for the development of comprehensive hydrological models tailored to Ireland's diverse basins.

Hydrological models can be categorised based on their approach to spatial representation (Singh and Woolhiser, 2003). There are two primary categories: lumped models and distributed models. Lumped models simplify the mathematical representation of rainfall-runoff processes by using averaged parameters that apply to the entire basin. However, these parameters in lumped models need to be calibrated, as they cannot be directly measured. Consequently, the applicability of lumped models is limited to basins where gauge data is available (Crawford and Linsley, 1966). With the advancements in computer technology, the development of distributed models has emerged as a viable alternative to describe rainfall-runoff processes in a spatially distributed manner. These models employ parameters directly linked to the physical characteristics of the basin, such as topography, soil composition, land cover, and river network. Furthermore, they consider the spatial variability present in both physical characteristics and meteorological inputs. Consequently, distributed models offer a more comprehensive and precise representation of hydrological processes within a basin (Wicks and Bathurst, 1996). Given their capacity to meet diverse spatial modelling requirements, including predicting the extent of floods and droughts, distributed models have seen extensive utilisation in the planning and management of water resources (Alaminie *et al.*, 2023).

Despite the advantages they offer, it is essential to acknowledge that distributed hydrological models entail a significantly higher level of complexity and application challenges compared to lumped models. Originally introduced in the 1970s, distributed models were initially envisioned to operate without the need for prior calibration, relying on the direct determination of model parameters from field data. This implies that the accuracy of geospatial and spatial meteorological input data plays a crucial role in influencing the outcomes of the model (Abbott *et al.*, 1986).

In practical terms, however, obtaining such in situ data is not always feasible, particularly in regions characterised by ungauged or sparsely gauged basins. Consequently, similar to lumped models, certain parameters within distributed models still require calibration (Refsgaard, 1997). The process of calibrating a distributed model typically involves assigning distinct parameter values to various grid cells. While this approach is necessary, it introduces a considerable risk of overparameterisation, demands substantial computational resources, particularly at fine spatial resolutions, and can impede the application of the model in extensive basins (Beven, 2006).

Recent decades have witnessed remarkable advancements in remote sensing and geographic information system (GIS) data, encompassing valuable information on factors such as soil properties, land cover, and climate. These breakthroughs have the potential to bridge gaps in field data availability and provide innovative opportunities for spatial calibration and validation. Numerous remote sensing and GIS datasets are now accessible in the global public domain, characterised by commendable spatial and temporal resolutions (Fortin *et al.*, 2001). Combined with the computational capabilities of robust computer resources, this strategic approach minimises the need for extensive parameter calibration in distributed models. It offers a promising possibility for developing more practical and dependable distributed models, especially in basins where data is scarce or limited (Wannasin *et al.*, 2021).

Accurate representation of precipitation inputs is crucial for the reliability and effectiveness of hydrological models, as it directly influences the prediction of runoff, flooding, and water resource availability. In this context, the choice of rainfall data sources and their quality is of paramount importance. One notable advancement in the field of meteorological data is the introduction of the ERA5 (fifth generation of the ECMWF Reanalysis) rainfall datasets by the European Centre for Medium-Range Weather Forecasts (ECMWF) (Hersbach *et al.*, 2020). ERA5 datasets have gained considerable attention in recent years for their comprehensive and high-resolution records of meteorological variables, including precipitation, spanning several decades. This introduction of ERA5 rainfall data into the meteorological and climatological landscape has brought unprecedented opportunities for improving the precision and reliability of hydrological models.

This paper addresses these critical considerations by pursuing two key objectives. Firstly, we aim to assess the ability of the Wflow.jl model (van Verseveld *et al.*, 2022) to replicate observed runoff within the Boyne Catchment which is characterised by diverse land uses, topographies, and an extensive river network. Such diversity makes the selected catchment an ideal study area for evaluating hydrological model performance. Secondly, we seek to examine the implications of utilising different rainfall datasets on model results. In particular, the model's performance is assessed using two distinct rainfall datasets as inputs: one derived from conventional rain gauge observations and the other from advanced remote sensing and weather radar technology ERA5 dataset. The results obtained from this comparative analysis will provide valuable insights into the model's sensitivity to rainfall data sources and its effectiveness in representing the hydrological processes in this diverse and dynamic catchment.

2. STUDY AREA

The study area encompasses the Boyne Catchment, located in the eastern part of Ireland (Figure 1). This catchment extends between latitudes 53.35°N and 53.84°N and longitudes 6.29°W and 7.01°W, covering a diverse range of topographical features, land uses, and hydrological characteristics. With an area of approximately 2690 km^2, the catchment is characterised by a mix of agricultural areas, urban-residential zones, natural landscapes, and water bodies. The Boyne Catchment's topography exhibits variations in elevation, ranging from the lowest point at sea level to the highest peak at 330 meters above sea level.

The Boyne catchment experiences wet winters and drier summers, making it susceptible to rainfall-induced flooding events. This distinct weather pattern is a consequence of Ireland’s temperate climate which is marked by its proximity to the Atlantic Ocean, exerting a significant influence on local weather patterns.

Flowing through the catchment is the Boyne River, which ultimately discharges into the Irish Sea. The river network within the catchment is extensive, comprising numerous tributaries and smaller streams that contribute to the overall hydrological behaviour.

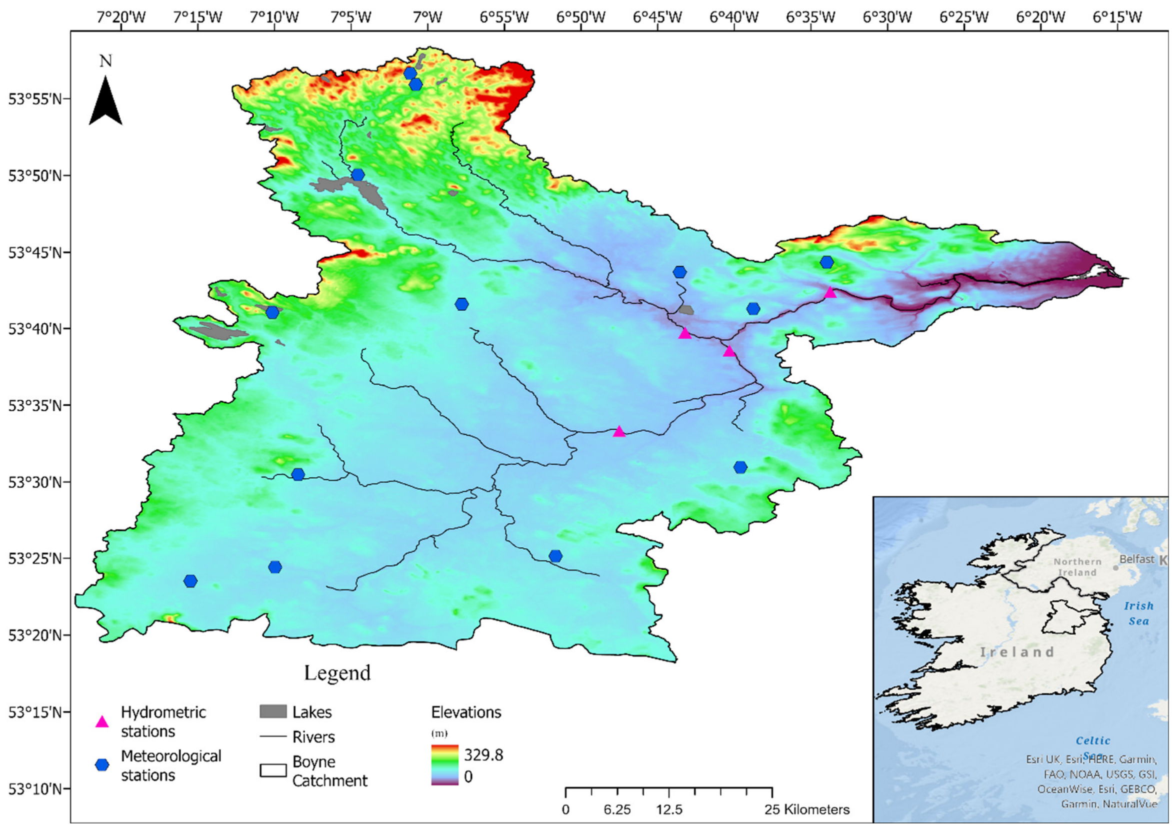


***Figure 1**: Overview of the study area: Boyne Catchment*

3. METHODOLOGY

3.1 Data acquisition and pre-processing

This study requires integration of geo-spatial physical, meteorological, and hydrological data for configuring the model. Geo-spatial physical data takes on the role of depicting the basin's morphological, physical, soil, and land use properties for each grid of the computational domain within the Wflow_sbm model. Meteorological data serves as the dynamic forces driving the model, while and hydrological data assumes a pivotal function in parameter optimisation and the subsequent analysis of results.

3.1.1 Geo-spatial physical (static) data

This dataset encompasses hydrography data such as basin maps, elevations from Digital Elevation Model (DEM), and land surface slope. All the hydrography data were obtained from merit_hydro_1k at 30 arcsec resolution (Eilander et al., 2021). Additionally, the static data consists of river system details, comprising a river network map and the bank full discharge data for each reach within the rivers. These river system data were initially obtained from Lin et al. (2020) and were subsequently verified by cross-referencing with the Ireland river network downloaded from the EPA website (https://gis.epa.ie).

Additionally, this study considered static maps related to lakes, reservoirs, and glaciers, although, in the context of the Boyne catchment, there was no reservoirs or glaciers, lakes data was downloaded from hydroLAKES (Messager et al., 2016). Land cover types, soil types, leaf area index (LAI) climatology maps per month are also included in the model. The land cover was obtained from Corine Land Cover (CLC) 2018 (European Environment Agency, 2018), the soil from SoilGrids (250 m resolution) (Hengl *et al.*, 2017) and the LAI dataset has been extracted from MCD15A3H Version 6 (Myneni, R., Knyazikhin, Y., Park, T., 2015).

3.1.2 Metrological (forcing) data

The meteorological data for forcing the Wflow_sbm model were precipitation, temperature, and solar radiation. the model was forced with two distinct datasets for gridded precipitation, the first dataset is ERA5 precipitation, with a spatial resolution of a regular lat-lon grid of 0.25 degrees (31 km) and daily temporal resolution (Hersbach *et al.*, 2020). The second dataset incorporated daily precipitation data from Met Éireann weather observation network (https://www.met.ie/climate/available-data/daily-data). Within the Boyne catchment area, a total of 13 weather stations were situated, among a network of 339 stations located across Ireland. It is crucial to note that the static and forcing maps in this study possess the same spatial domain and resolution. Consequently, any re-gridding of the forcing data was not supported. Instead, the data collected from the station-based observations underwent interpolation using the Inverse Distance Weighting (IDW) method. This process resulted in the creation of gridded precipitation data, aligning precisely with the 0.25-degree resolution of ERA5 precipitation dataset. Gridded temperature and solar radiation data were extracted from the daily ERA5 dataset. These datasets were subsequently utilised to compute potential evapotranspiration using the De Bruin method (De Bruin *et al.*, 2016).

3.1.3 Hydrological data

The required hydrological data encompassed daily streamflow observations from hydrometric gauges within the Boyne catchment. These gauges, namely Slane Castle, Navan Weir, Trim, and Liscarton. In this paper we demonstrated the results of Trim gauge. The data were acquired to align with the model's simulation period, which spanned from 2003 to 2022. It was ensured that the data provided had a daily temporal resolution and was expressed in units of cubic meters per second (m^3/sec).

3.2 Wflow_sbm model

Wflow.jl (v0.6.1) (van Verseveld *et al.*, 2022) is a state-of-the-art, open-source spatially distributed hydrological model that is equipped with multiple hydrologic model concepts, all designed for effective implementation written in Julia programming language. This framework represents a progression from the earlier Wflow framework (Schellekens *et al.*, 2020), which was Python-based PCRaster (Karssenberg *et al.*, 2010). Wflow.jl features a rich selection of vertical and lateral concepts, all of which are valuable tools in the use of hydrologic modelling.

In this study, we used Wflow_sbm, the primary hydrological modelling concept integrated within the Wflow.jl framework. Wflow_sbm stands as a representative of a group of hydrological models that all share the fundamental concept of the vertical Soil Bucket Model (SBM). This SBM concept is strongly influenced by Topog_SBM (Vertessy and Elsenbeer, 1999), which envisions the soil as a "bucket" encompassing both saturated and unsaturated stores, although it has undergone substantial modifications and improvements over time.

Wflow_sbm offers versatility by supporting various lateral concepts for water routing, covering river, overland, and subsurface flow. In this study, the Wflow_sbm model, as illustrated in Figure 2, adopts the kinematic-wave approach to manage the lateral components, aligning with approaches used in other hydrological odels such as TOPKAPI and 1K-DHM (Liu and Todini, 2002; Tanaka and Tachikawa, 2015), and Topog_SBM (Vertessy and Elsenbeer, 1999). Figure 2 provides a comprehensive visual representation of the essential processes and fluxes within the wflow_sbm model.

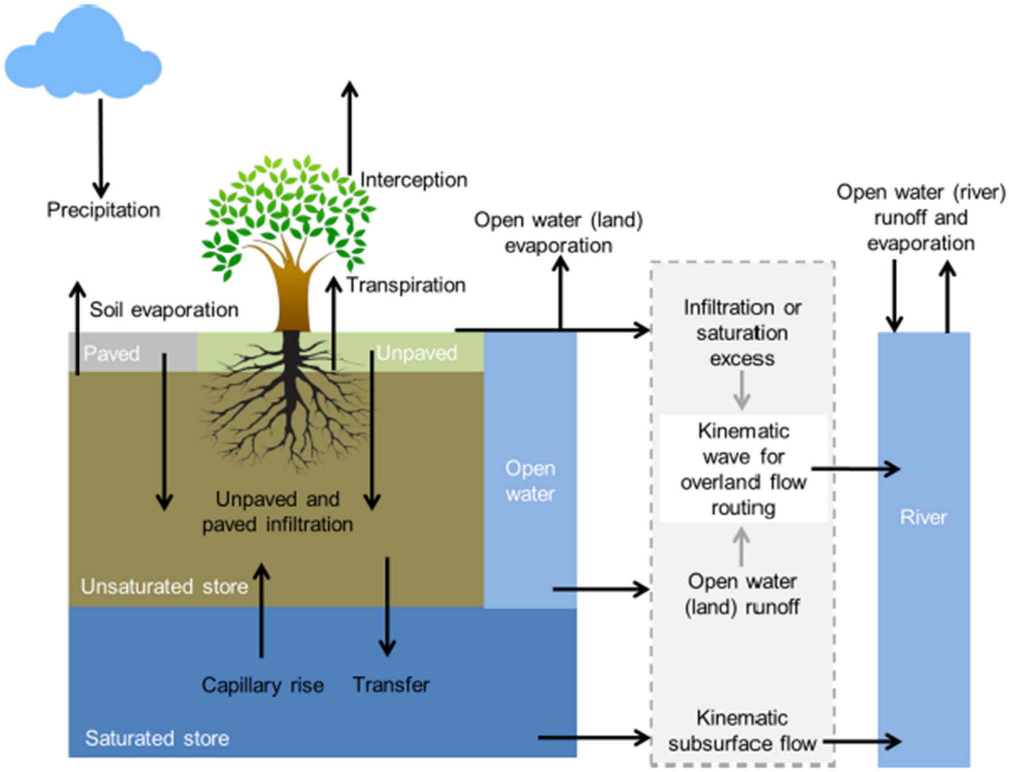


Figure 2: Schematic structure of water processes and fluxes in the Wflow_sbm model

3.3 Wflow_sbm parameterisation using HydroMT-Wflow

In this study, the parameter sets utilised were generated using the hydroMT software package (Eilander *et al.*, 2023). The specific data sources for deriving these parameter sets are outlined in the preceding sections. These parameters are initially estimated based on pedotransfer-functions (PTFs) at their original data resolution ('level-0') and are subsequently upscaled to the model resolution ('level-1') using upscaling operators (van Verseveld *et al.*, 2022).

For instance in this study, soil-related parameters were estimated using the Pedo-Transfer Function (PTF) used for estimating soil properties is based on Brakensiek (Brakensiek, Rawls and Stephenson, 1984) and land cover-related parameters were directly taken from the Corine 2018 dataset. Some of the parameters have unavailable PTFs, and for these, constant values were assigned, in accordance with the default values developed by Schellekens et al. (2019), This left only one soil-related parameter, KsatHorFrac [–], to undergo calibration. KsatHorFrac serves as a multiplication factor applied to KsatVer (vertical saturated conductivity) to determine horizontal saturated conductivity for the computation of lateral subsurface flow.

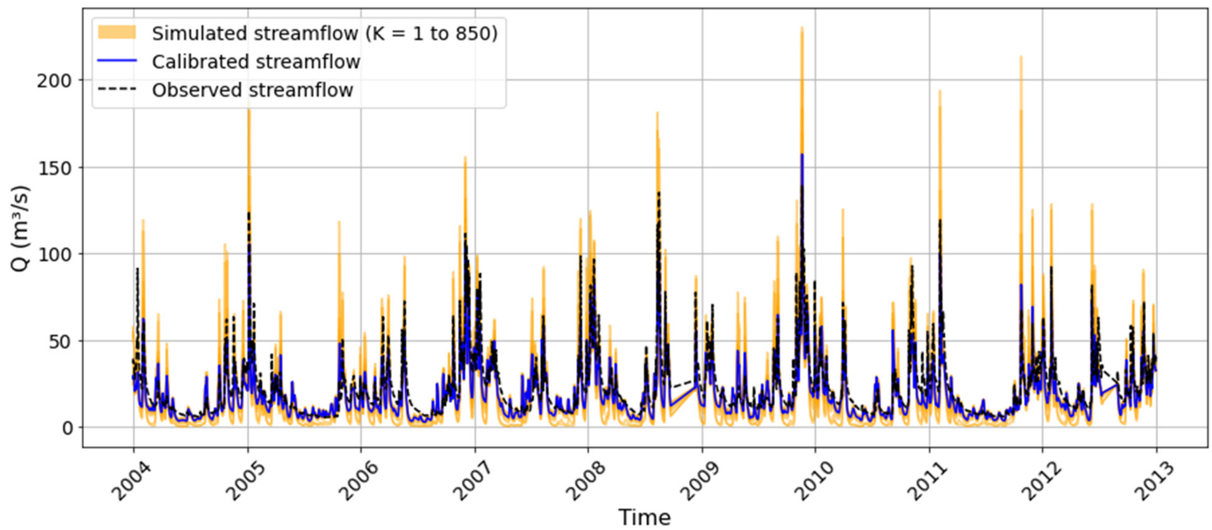
3.4 Model application, calibration and validation

The Wflow_sbm model was applied to the Boyne catchment at a spatial resolution of 30 arcsec (approximately 1 km) and a daily temporal resolution. The calibration process involved manual adjustments of the KsatHorFrac parameter, spanning the years 2003 to 2012, encompassing both wet and dry periods, with parameter values ranging from 100 to 800. Subsequently, the optimized KsatHorFrac values were validated for the period spanning 2013 to 2022. The Kling Gupta efficiency (KGE) (Gupta *et al.*, 2009) and NSE (Nash and Sutcliffe, 1970) were used as objective functions during the calibration of the KsatHorFrac parameter and as performance metrics to evaluate the accuracy of the Wflow_sbm model.

4. RESULTS AND DISCUSSION

Overall, the Wflow_sbm model, when forced by both ERA5 precipitation dataset and Ground station dataset, effectively replicates the daily streamflow observed in the Boyne catchments. The daily hydrographs of the Boyne catchment throughout the calibration period (2004 to 2012), shown in Figures 3 and 4, demonstrate remarkable correspondence in peak timing and seasonal fluctuations between the observed and simulated streamflows. The simulated streamflows are depicted as orange-coloured bands, representing a range. Within this range, the simulated streamflow obtained using the calibrated KsatHorFrac value is illustrated as a continuous solid line, while the observed streamflow is represented by a dashed black line for comparison. This finding aligns with the results of a study conducted by Wannasin *et al.* (2021) , in which they investigated the impact of employing various pedotransfer functions for estimating soil map parameters in the Greater Chao Phraya River (GCPR) basin in Thailand, focusing on its influence on daily streamflow.

In terms of comparing the two precipitation datasets, initially, there appears to be no distinguishable difference between the simulations produced by the model when driven by either ERA5 or ground station datasets. To prevent any presentation bias during the calibration period, we deliberately chose the years 2005 and 2008 as representative years, which respectively reflect dry and wet periods.



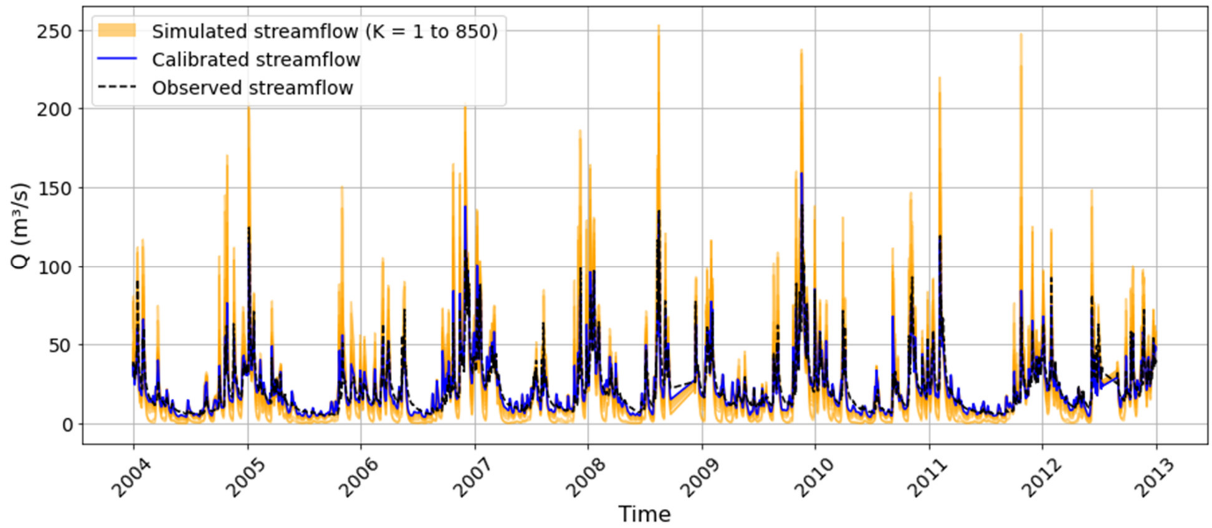


***Figure 3**: Daily streamflow observations and simulations (with Varied KsatHorFrac Values Ranging from 1 to 850), Including the Calibrated Simulation) for the ERA5 dataset (Upper) and Ground-stations dataset (lower)*

During the calibration of the KsatHorFrac parameter, it became evident that lower KsatHorFrac values resulted in inadequate simulations, as indicated by low NSE (e.g., -0.20) and KGE (e.g., 0.147) values, with a notable instance observed at KsatHorFrac 50. In contrast, with the ERA5 dataset the model achieved higher NSE and KGE values at the same KsatHorFrac of 50, with values of 0.27 and 0.46, respectively. Conversely, for higher KsatHorFrac values, the model, driven by both datasets, demonstrated excellent performance, as indicated by the high NSE and KGE values (Table 1).

***Table 1**: Wflow_sbm performance in simulating daily streamflow with ERA5 and Ground-stations datasets.*

ERA5					Ground station			
	Calibration		Validation		Calibration		Validation	
K	NSE	KGE	NSE	KGE	NSE	KGE	NSE	KGE
1	0.11	0.38			-0.39	0.07		
50	0.27	0.46			-0.20	0.147		
150	0.54	0.64			0.19	0.34		
250	0.67	0.75			0.45	0.50		
350	0.72	0.80			0.61	0.62		
450	0.75	0.80			0.71	0.71		
550	0.76	0.78	0.77	0.84	0.77	0.78		
650	0.76	0.75			0.81	0.84		
750	0.76	0.72			0.83	0.88		
850	0.75	0.70			0.85	0.91	0.84	0.84

The calibrated KsatHorFrac value for the Boyne catchment is 550 when the model is driven by the ERA5 dataset, whereas it is 850 when forced by the ground station dataset. The range of simulated daily streamflows, spanning KsatHorFrac values from 1 to 850, is notably broader for the model driven by the ground station dataset. This is visually evident (Figure 4 and 5) and supported by the data in Table 1, where the range extends from -0.39 to 0.85 for NSE and from 0.07 to 0.91 for KGE. In contrast, for the model driven by ERA5, the range is narrower, with values of 0.11 to 0.75 for NSE and 0.38 to 0.7 for KGE. This implies that the ERA5 dataset introduces a smaller degree of uncertainty in the streamflow results in comparison to the ground station dataset. On the other hand, with their calibrated KsatHorFrac values, the models derived from the ERA5 and ground station datasets provide daily streamflow estimates for the Boyne catchment that are comparable, with slightly better performance observed for the model forced by the ground station dataset, as demonstrated in Figure 4 and the goodness-of-fit indicators in Table 1.

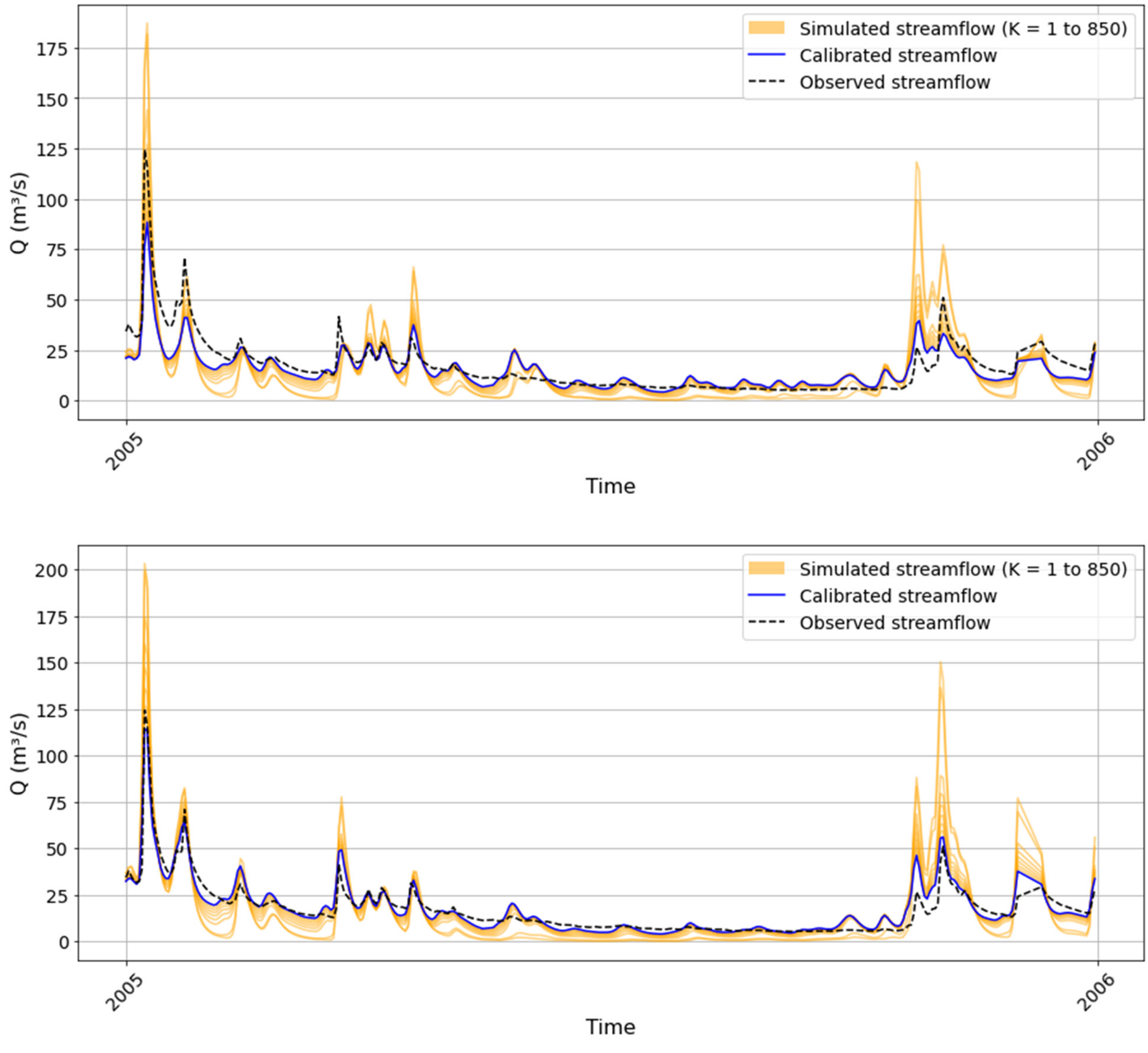


Figure 4*: Observed and simulated daily streamflows for the Boyne catchment in 2005 (dry period) for the ERA5 dataset (Upper) and Ground-stations dataset (lower)*

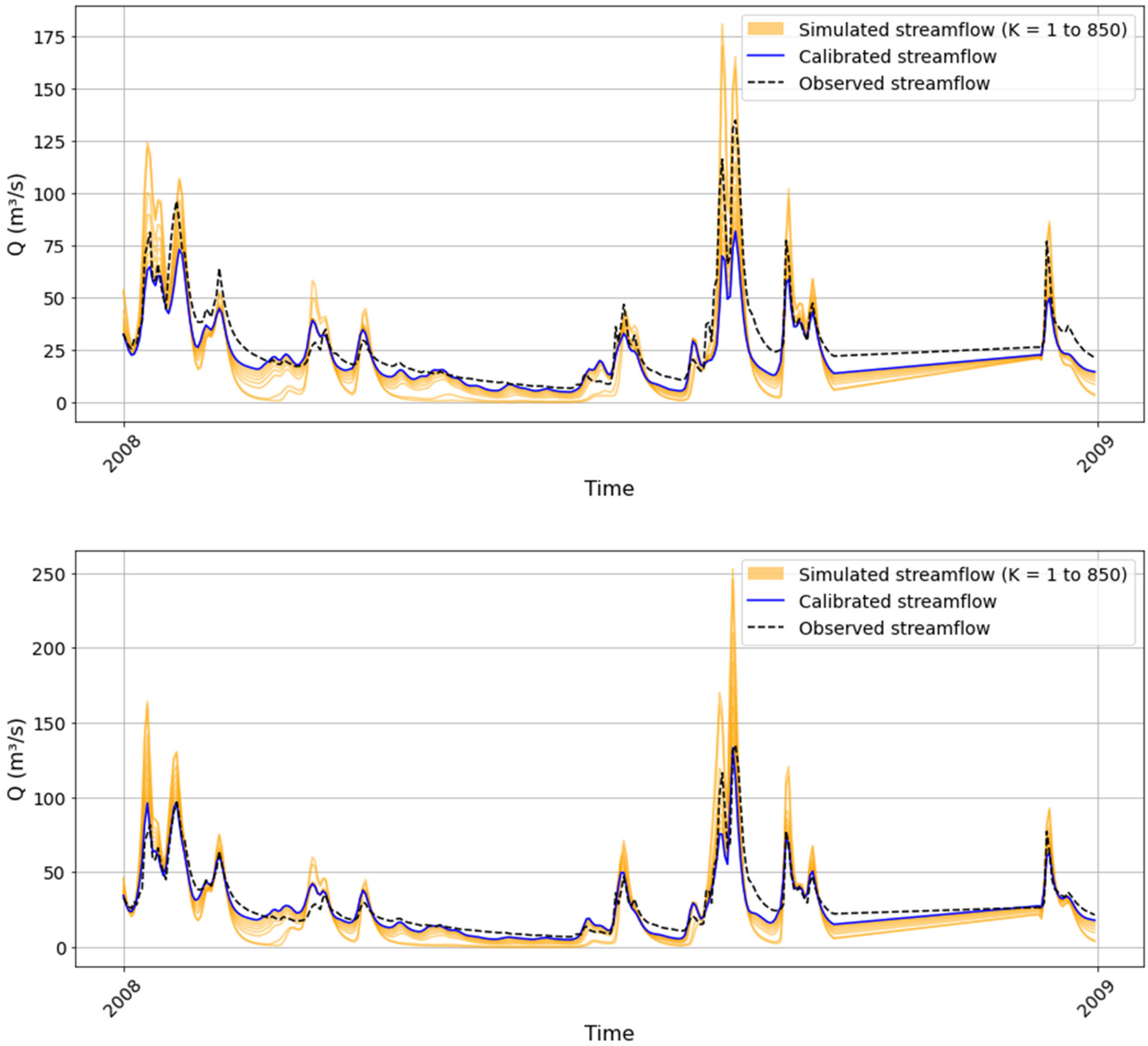


***Figure 5**: Observed and simulated daily streamflows for the Boyne catchment in 2008 (wet period) for the ERA5 dataset (Upper) and Ground-stations dataset (lower)*

Overall, throughout the validation period, the model demonstrated highly promising performance when driven by both precipitation datasets (ERA5 and ground stations); however, modelling results revealed minor underestimation in peak flows when the model was forced by the ERA5 rainfall dataset compared to the model driven by the ground truth gridded rainfall data. We specifically selected the timeframe from mid-2015 to mid-2018 to represent the model simulations in the validation period, using both ERA5 and ground station datasets (Figure 6). This period encompasses the extreme weather event of the winter 2015/2016 and several instances of limited streamflows, making it a valuable indicator for visual interpretation of the results. It can be clearly shown, in contrast to the calibration period, the model driven by the ground station dataset tends to overestimate the peaks but accurately replicates the baseflow. Conversely, the model forced by the ERA5 dataset exhibits the ability to capture both peaks and baseflows.

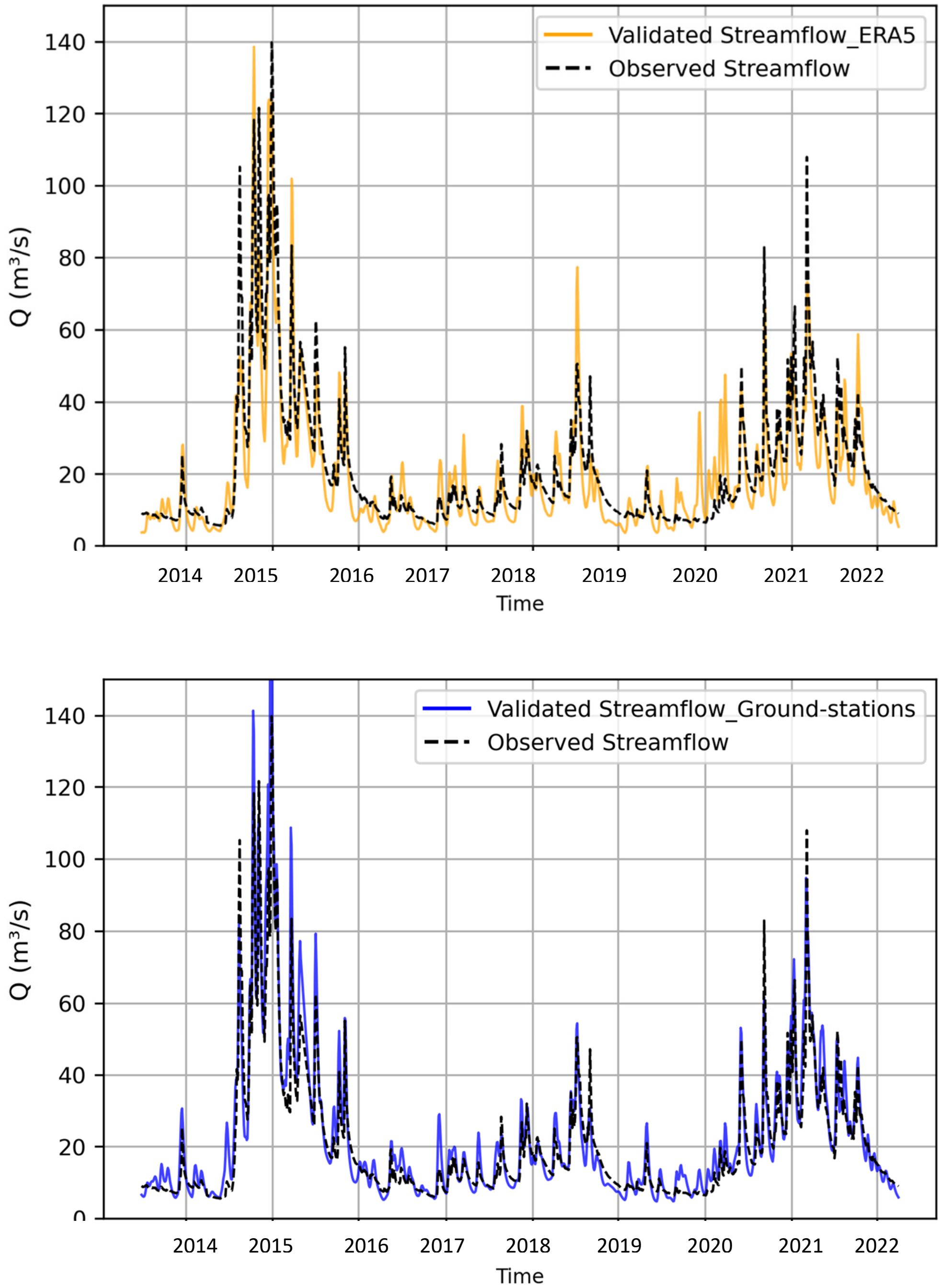


***Figure 6**: Observed and validated daily streamflows for the Boyne catchment for the ERA5 dataset (orange_upper) and Ground-stations dataset (blue_lower). From (July 2015 to May 2018)*

5. CONCLUSION

This study marks the groundbreaking use of the Wflow_sbm model to generate daily streamflow simulation within an Irish catchment. Our approach entailed configuring the model using global geospatial datasets and, where possible, validating it using locally accessible data sources. Additionally, we forced the model with meteorological data from both global and in situ sources, specifically utilising ERA5 and ground-based stations precipitation datasets. This approach effectively addresses the common challenge of limited in situ data, often associated with fully physical distributed models that necessitate extensive data resources, which may not always be readily available.

The daily streamflows of the Boyne catchment were successfully replicated during the validation period from 2013 to 2022, achieving high goodness-of-fit measures for NSE and KGE. This was accomplished using two distinct datasets: ERA5 from the European Centre for Medium-Range Weather Forecasts (ECMWF) and the ground-station precipitation dataset from the Met Éireann weather observation network.

The results highlight the model's sensitivity to horizontal saturated conductivity in the computation of lateral subsurface flow, as indicated by the KsatHorFrac parameter, which varies across a range of values from 1 to 850. This sensitivity becomes evident through the varying model responses. It aligns with the recommendations in existing literature, emphasizing the impact of this parameter on streamflow generated by different datasets.

Drawing a definitive conclusion regarding the performance of the two precipitation datasets is challenging due to several factors. First and foremost, we conducted manual calibration for only one parameter of the Wflow_sbm model, as recommended in the literature. There remains a need for further investigation and calibration of the remaining parameters for which pedotransfer functions are not available. Additionally, it is crucial to acknowledge that the choice of objective functions for calibration is subjective. The selection of the best parameter values for calibration is strongly influenced by the chosen objective function. For instance, the NSE score may favour flashiness over base flow. Furthermore, the KGE metric warrants a more comprehensive examination of its components, as each component serves as a valuable indicator of uncertainty and the model's tendency to overestimate, or underestimate mean streamflows.

In light of these findings, there is a clear need for further research into the utilization of seamless parameter maps derived from global data for distributed hydrological modelling. Additionally, it is advisable to investigate the impact of employing various global meteorological datasets on model performance.

6. ACKNOWLEDGEMENT

The research conducted in this publication was funded by the Irish Research Council under 2021 Government of Ireland Postgraduate (GOIPG) Scholarship under the grant number [GOIPG/2021/1388].

Also, the authors would like to extend their gratitude to Willem J. van Verseveld for his invaluable contribution of leaf area index (LAI) climatology maps extracted and prepared for Ireland, which were used in this study. His expertise and generosity greatly enriched the quality and scope of our research.

7. REFERENCES

Abbott, M. B. *et al.* (1986) ‘An introduction to the European Hydrological System - Systeme Hydrologique Europeen, “SHE”, 1: History and philosophy of a physically-based, distributed modelling system’, *Journal of Hydrology*, 87(1–2), pp. 45–59. doi: 10.1016/0022-1694(86)90114-9.

Alaminie, A. A. *et al.* (2023) 'Nested hydrological modeling for flood prediction using CMIP6 inputs around Lake Tana, Ethiopia', *Journal of Hydrology: Regional Studies*, 46. doi: 10.1016/j.ejrh.2023.101343.

Beven, K. (2006) 'A Manifesto for the Equifinality Thesis Keith Beven Lancaster University, UK', (1992).

Brakensiek, D. L., Rawls, W. J. and Stephenson, G. R. (1984) 'Modifying SCS hydrologic soil groups and curve numbers for rangeland soils', *1984 Annual Meeting ASAE, Pacific Northwest Region*, paper no., pp. 1–13.

De Bruin, H. A. R. *et al.* (2016) 'Thermodynamically based model for actual evapotranspiration of an extensive grass field close to FAO reference, suitable for remote sensing application', *Journal of Hydrometeorology*, 17(5), pp. 1373–1382. doi: 10.1175/JHM-D-15-0006.1.

Cloke, H. L. and Pappenberger, F. (2009) 'Ensemble flood forecasting: A review', *Journal of Hydrology*, 375(3–4), pp. 613–626. doi: 10.1016/j.jhydrol.2009.06.005.

Crawford, N. . and Linsley, R. K. (1966) 'Digital Simulation in Hydrology', *Contemporary Hydrology*, pp. 157–158.

Eilander, D. *et al.* (2021) 'A hydrography upscaling method for scale-invariant parametrization of distributed hydrological models', *Hydrology and Earth System Sciences*, 25(9), pp. 5287–5313. doi: 10.5194/hess-25-5287-2021.

Eilander, D. *et al.* (2023) 'HydroMT: Automated and reproducible model building and analysis', *Journal of Open Source Software*, 8(83), p. 4897. doi: 10.21105/joss.04897.

Fortin, J.-P. *et al.* (2001) 'Distributed Watershed Model Compatible with Remote Sensing and GIS Data. I: Description of Model', *Journal of Hydrologic Engineering*, 6(2), pp. 91–99. doi: 10.1061/(asce)1084-0699(2001)6:2(91).

Gupta, H. V. *et al.* (2009) 'Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling', *Journal of Hydrology*, 377(1–2), pp. 80–91. doi: 10.1016/j.jhydrol.2009.08.003.

Hengl, T. *et al.* (2017) 'SoilGrids250m: Global gridded soil information based on machine learning', *PLoS ONE*, 12(2), p. e0169748. doi: 10.1371/journal.pone.0169748.

Hersbach, H. *et al.* (2020) 'The ERA5 global reanalysis', *Quarterly Journal of the Royal Meteorological Society*, 146(730), pp. 1999–2049.

Karssenberg, D. *et al.* (2010) 'A software framework for construction of process-based stochastic spatio-temporal models and data assimilation', *Environmental Modelling and Software*, 25(4), pp. 489–502. doi: 10.1016/j.envsoft.2009.10.004.

Lin, P. *et al.* (2020) 'Global Estimates of Reach-Level Bankfull River Width Leveraging Big Data Geospatial Analysis', *Geophysical Research Letters*, 47(7). doi: 10.1029/2019GL086405.

Liu, Z. and Todini, E. (2002) 'Towards a comprehensive physically-based rainfall-runoff model', *Hydrology and Earth System Sciences*, 6(5), pp. 859–881. doi: 10.5194/hess-6-859-2002.

Messager, M. L. *et al.* (2016) 'Estimating the volume and age of water stored in global lakes using a geo-statistical approach', *Nature Communications*, 7. doi: 10.1038/ncomms13603.

Myneni, R., Knyazikhin, Y., Park, T. (2015) 'MCD15A3H MODIS/Terra+Aqua Leaf Area Index/FPAR 4-day L4

Global 500m SIN Grid V006 [Data set]', *NASA EOSDIS Land Processes DAAC*. doi: https://doi.org/10.5067/MODIS/MCD15A3H.006.

Nash, J. E. and Sutcliffe, I. V. (1970) 'River flow forecasting through conceptual models Part I - A discussion of principles', *Journal of Hydrology*, 10(3), pp. 282–290.

Refsgaard, J. C. (1997) 'Parameterisation, calibration and validation of distributed hydrological models', *Journal of Hydrology*, 198(1–4), pp. 69–97. doi: 10.1016/S0022-1694(96)03329-X.

Schellekens, J. *et al.* (2020) 'openstreams/wflow: Bug fixes and updates for release 2020.1.2'. Available at: https://zenodo.org/record/4291730.

Sherlock, E. and Duffy, S. (2019) '05-Establishing the flood forecast centre and expanding met Eireann's rainfall radar network'. Available at: http://hydrologyireland.ie/wp-content/uploads/2019/11/05-Eoin-Sherlock-Establishing-the-Flood-Forecast-Centre-and-Expanding-Met-Eireanns-Radar.pdf.

Singh, V. P. and Woolhiser, D. A. (2003) 'Mathematical Modeling of Watershed Hydrology', *Perspectives in Civil Engineering: Commemorating the 150th Anniversary of the American Society of Civil Engineers*, pp. 345–367. doi: 10.1061/(asce)1084-0699(2002)7:4(270).

Tanaka, T. and Tachikawa, Y. (2015) 'Testing the applicability of a kinematic wave-based distributed hydrological model in two climatically contrasting catchments', *Hydrological Sciences Journal*, 60(7–8), pp. 1361–1373. doi: 10.1080/02626667.2014.967693.

van Verseveld, W. J. *et al.* (2022) 'Wflow_sbm v0.6.1, a spatially distributed hydrologic model: from global data to local applications', *Geoscientific Model Development Discussions*, pp. 1–52. Available at: https://doi.org/10.5194/gmd-2022-182.

Vertessy, R. A. and Elsenbeer, H. (1999) 'Distributed modeling of storm flow generation in an Amazonian rain forest catchment: Effects of model parameterization', *Water Resources Research*, 35(7), pp. 2173–2187. doi: 10.1029/1999WR900051.

Wannasin, C. *et al.* (2021) 'Daily flow simulation in Thailand Part I: Testing a distributed hydrological model with seamless parameter maps based on global data', *Journal of Hydrology: Regional Studies*, 34. doi: 10.1016/j.ejrh.2021.100794.

Wicks, J. M. and Bathurst, J. C. (1996) 'SHESED: A physically based, distributed erosion and sediment yield component for the SHE hydrological modelling system', *Journal of Hydrology*, 175(1–4), pp. 213–238. doi: 10.1016/S0022-1694(96)80012-6.