CGIAR
SCIENCE FOR A FOOD-SECURE FUTURE

Technical REPORT

Desktop environmental flow assessment of the Omo-Gibe Basin, Ethiopia

Vincent G. L. Lishalan; John P. Simaika; Matthew McCartney

November 2026



Authors

Vincent G. L. Lishalan, Intern, International Water Management Institute (IWMI), Colombo, Sri Lanka

John P. Simaika, Senior Researcher – Aquatic Ecology, IWMI, Colombo, Sri Lanka

Matthew McCartney, Research Group Leader, Sustainable Water Infrastructure and Ecosystems, IWMI, Colombo, Sri Lanka

Author contributions

Conceptualization: John P. Simaika, Matthew McCartney; Data curation: Vincent G. L. Lishalan; Formal analysis: Vincent G. L. Lishalan, John P. Simaika; Methodology: Vincent G. L. Lishalan, John P. Simaika, Matthew McCartney; Project administration: John P. Simaika; Writing – original draft preparation: Vincent G. L. Lishalan, John P. Simaika; Writing – review and editing: John P. Simaika, Matthew McCartney; Visualization: Vincent G. L. Lishalan, John P. Simaika; Investigation: Vincent G. L. Lishalan, John P. Simaika; Supervision: John P. Simaika, Matthew McCartney.

Acknowledgments

This work was conducted as part of the CGIAR Multifunctional Landscapes Program. We would like to thank all funders who supported this research through their contributions to the CGIAR Trust Fund (www.cgiar.org/funders/). We also thank colleagues and partners who contributed insights, data, review comments, and technical feedback during the development of this publication. AI-use disclosure: This publication made limited use of AI tools for language refinement and formatting support. All AI-assisted content was reviewed and approved by the authors. The data, analysis, and substantive content remain entirely the work and responsibility of the authors.

CGIAR Multifunctional Landscapes Program

The CGIAR Multifunctional Landscapes (MFL) Program supports integrated management of land, water, and biodiversity to strengthen resilience, improve livelihoods, and promote inclusive governance through holistic landscape approaches that protect ecosystems and enable diverse stakeholders to co-create solutions.

Learn more about the program: https://www.cgiar.org/cgiar-research-portfolio-2025-2030/multifunctional-landscapes/

Suggested citation

Lishalan, Vincent G. L., John P. Simaika, and Matthew McCartney. 2026. *Desktop Environmental Flow Assessment of the Omo-Gibe Basin, Ethiopia.* International Water Management Institute (IWMI).

© 2026 International Water Management Institute. Some rights reserved. This work is licensed under a Creative Commons Attribution 4.0 International License (CC BY 4.0).

Front cover photo: Upper catchment of the Omo-Gibe River, Ethiopia. (Photo by Wolde Mekuria)
Back cover photo: Upper catchment of the Omo-Gibe River, Ethiopia. (Photo by Wolde Mekuria)

Disclaimer

This publication has been prepared as an output of the CGIAR Multifunctional Landscapes Program and has not been independently peer reviewed. Responsibility for editing, proofreading, and layout, opinions expressed, and any possible errors lies with the authors and not the institutions involved. Boundaries used in the maps do not imply the expression of any opinion whatsoever on the part of CGIAR concerning the legal status of any country, territory, city, or area or of its authorities, or concerning the delimitation of its frontiers or boundaries. Borders are approximate and cover some areas for which there may not yet be full agreement.

Contents

Acronyms and abbreviations 5

Summary 6

1. Introduction 6

1.1 Study objectives 10

2. Methodology 11

2.1 Study area and overall approach 11
2.2 Naturalized streamflow modeling 13
2.2.1 Observed streamflow data 13
2.2.2 Candidate predictors and data sources 15
2.2.3 Feature selection 16
2.2.4 Model structure and hyperparameters 17
2.2.5 Training, testing, and evaluation 18
2.2.6 Selection of the assessment sites and application of the model 18
2.3 Desktop e-flow assessment 21
2.3.1 Method selection 21
2.3.2 Model inputs 21
2.3.3 Transfer of South African parameters 21
2.3.4 Model runs and outputs 22
2.4 Stakeholder questionnaire 23
2.4.1 Questionnaire design and distribution 23
2.4.2 Data analysis 23

3. Results 24

3.1 Naturalized streamflow modeling 24
3.1.1 Predictor selection 24
3.1.2 Model performance 24
3.1.3 Simulated and observed hydrographs 26
3.1.4 Naturalized flow at the assessment sites 27
3.2 Environmental water requirements 30
3.2.1 Total-flow requirements 30

4. Discussion 41

4.1 Naturalized flow model 41
4.2 Environmental water requirements 42
4.3 Scope and missing components 42
4.4 Stakeholder perspectives 43
4.5 Limitations and next steps 43
4.6 Towards a digital twin of the Omo-Gibe Basin 43

5. Conclusions 44

References 44

Appendix A. Naturalized streamflow model 49

A.1 Candidate predictors 49
A.2 Performance at the gauged stations 51
A.3 Comparison with GLDAS-2.1 52

Appendix B. Environmental flow assessment 58

B.1 RDRM inputs 58
B.2 Natural flow 60
B.3 Requirements for all ecological categories 60
B.4 Monthly requirements at the sites 63

Appendix C. Stakeholder questionnaire 68

Section A. Respondent information 68
Section B. Data availability and accessibility 69

Appendix D. Pathway to a digital twin of the Omo-Gibe Basin 75

Acronyms and abbreviations

BFI	Baseflow index
CHIRPS	Climate Hazards Group InfraRed Precipitation with Station data
CV	Coefficient of variation
EIS	Ecological Importance and Sensitivity
ELOHA	Ecological limits of hydrologic alteration
ERA5	Land European Centre for Medium-Range Weather Forecasts reanalysis, version 5 (land component)
EWR	Environmental water requirement
FD / FI / FS	Fast deep / fast intermediate / fast shallow (velocity-depth classes)
FDC	Flow duration curve
GLDAS	Global Land Data Assimilation System
HydroBASINS	Hydrological sub-basin boundaries derived from HydroSHEDS
HydroRIVERS	River network derived from HydroSHEDS
HydroSHEDS	Hydrological data and maps based on Shuttle Elevation Derivatives at multiple Scales
IoT	Internet of Things
IWMI	International Water Management Institute
Log-NSE	Nash-Sutcliffe efficiency of log-transformed discharge
LIMCOM	Limpopo Watercourse Commission
MAR	Mean annual runoff
MoWE	Ministry of Water and Energy, Ethiopia
MR	Minimum requirement (digital twin)
NOAA	National Oceanic and Atmospheric Administration
NSE	Nash-Sutcliffe efficiency
PBIAS	Percent bias
PES	Present Ecological State
RDRM	Revised Desktop Reserve Model
REST	Representational state transfer (web service interface)
SD	Standard deviation
SPATSIM	Spatial and Time Series Information Modelling
SRTM	Shuttle Radar Topography Mission
SWAT+	Soil and Water Assessment Tool (revised version)
UN	United Nations
UNEP	United Nations Environment Programme
USGS	United States Geological Survey
XGBoost	Extreme gradient boosting

Summary

Freshwater ecosystems are in steep decline, and altered river flow is one of the main causes. The Omo-Gibe Basin in south-western Ethiopia supplies more than 80% of the inflow to Lake Turkana and is regulated by the Gibe hydropower cascade, yet no basin-wide environmental flow (e-flow) assessment exists. This study made a desktop e-flow assessment at 15 ungauged sites in the upper, middle, and lower sub-basins. Monthly naturalized flow was simulated with an XGBoost model driven by open-access satellite and reanalysis data (CHIRPS and ERA5-Land) and trained on four gauging stations. Environmental water requirements for ecological categories A to D were then estimated with the Revised Desktop Reserve Model (RDRM) in SPATSIM, using parameters transferred from South African reference rivers. A questionnaire of water professionals documented data gaps and barriers to implementation. The flow model reproduced seasonal flow well at the gauged stations (test-period NSE 0.78 to 0.89) but transferred poorly to ungauged catchments in a leave-one-station-out test. At the sub-basin outlets, the total environmental flow requirement was about 36% to 38% of natural mean annual runoff for category A and 18% to 19% for category D, and most of it came from the high-flow component. Low flows and the lower basin were the least certain, owing to doubtful channel inputs and few downstream gauges. The results are first-order estimates; more gauging, field data, ecological assessment, and agreed target categories are needed before they are used to set dam operating rules. These data, models, and stakeholder insights could also form the first components of a digital twin of the basin, and a phased pathway towards one is outlined.

1. Introduction

Freshwater ecosystems cover a small part of the Earth's surface, yet they host about one-third of all vertebrate species (Tickner et al. 2020b). They are also among the most threatened ecosystems on the planet. Wetlands are disappearing about three times faster than forests, and the populations of freshwater vertebrates have fallen more than twice as steeply as those on land or in the sea (Tickner et al. 2020b). Much of this loss is our own making. Dams, diversions, abstraction, pollution, and the conversion of floodplains change how much water a river carries, when it arrives, and where it can go. Altered flow is one of the five main causes of freshwater biodiversity loss, together with pollution, habitat degradation, overexploitation, and invasive species (Tickner et al. 2020b). These changes matter to people as well as wildlife. Healthy flows support benefits such as flood risk management (floodplains that store floodwater), floodplain fisheries (seasonal flooding that triggers fish spawning and feeding), and resilience to climate change (Arthington et al. 2024), and these benefits are put at risk when flows are altered.

The trend can still be reversed, and the need to act has become urgent. Tickner et al. (2020b) proposed an Emergency Recovery Plan to "bend the curve" of freshwater biodiversity loss, that is, to turn a trajectory of decline into one of recovery. The plan has six priority actions, and the first is to accelerate the implementation of environmental flows. The others are to improve water quality, protect and restore critical habitats, manage the exploitation of freshwater species and riverine aggregates, control invasive species, and restore river connectivity (Tickner et al. 2020b).

Governments have begun to respond. The Kunming-Montreal Global Biodiversity Framework, agreed in December 2022, includes the restoration of 30% of the world's degraded inland waters (UNEP 2023). In March 2023, a coalition of governments launched the Freshwater Challenge at the UN Water Conference. It is the largest initiative so far to restore degraded rivers, lakes, and wetlands, and it aims to restore 300,000 km of rivers and 350 million hectares of wetlands by 2030 (UNEP 2023). Action at the local or catchment level will not be enough on

its own. The decisions of the sectors that drive freshwater degradation, such as water management, agriculture, energy, and fisheries, must also change (Tickner et al. 2025).

Restoring a river is not possible without water of the right quantity, timing, and quality, and this is what environmental flows provide. E-flows therefore sit at the center of the response. They address flow alteration directly, and they are an important part of integrated strategies against the multiple threats to freshwater biodiversity (Arthington et al. 2024). Progress has been made, and e-flows are now being implemented in rivers and wetlands around the world (Tickner et al. 2020a), but implementation has been slower than needed and geographically uneven (Arthington et al. 2024). This study contributes to closing that gap in the Omo-Gibe Basin of Ethiopia.

Environmental flows, or e-flows, are the quantity, timing, and quality of water flows needed to sustain freshwater and estuarine ecosystems and the livelihoods and well-being of the people who depend on them (International River Foundation 2007). The definition was updated in 2018 to refer to the flows and levels that sustain aquatic ecosystems, which in turn support human cultures, economies, and livelihoods (Arthington et al. 2018). The idea rests on the natural flow regime, which is the pattern of low flows, floods, and seasonal change to which river life has adapted (Poff et al. 1997). An e-flow does not mean that all water must stay in the river. It identifies how much water, and what flow pattern, should be kept for the ecosystem when the remaining water is shared among other users.

An e-flow assessment is the scientific process used to estimate these flows. Its main purposes are to maintain ecological integrity, reduce the effects of infrastructure such as dams, protect downstream environments, and conserve water quality. The results give water managers a defensible basis for sharing water between the river and other users (Arthington et al. 2018).

More than 200 e-flow methods have been used in dozens of countries, and they fall into four main groups (Tharme 2003). Hydrological methods set requirements from the flow record alone. Hydraulic rating methods use the way channel width and depth change with discharge. Habitat simulation methods link flow to the habitat available for selected species, and holistic methods combine hydrology, hydraulics, ecology, and expert judgment (Tharme 2003). The more detailed approaches need field surveys, ecological data, and specialist teams, which many basins lack. Regional frameworks such as the ecological limits of hydrologic alteration (ELOHA) take another route and relate the degree of flow alteration to the ecological response across many rivers in a region (Poff et al. 2010).

Desktop methods were developed for basins where detailed data are not available. They use existing hydrological records and parameters transferred from similar rivers to give a rapid first estimate of flow requirements (Hughes and Hannart 2003), and have been applied in Ethiopia in the Blue Nile Basin (McCartney et al. 2009; Reitberger and McCartney 2011). In South Africa, this approach led to the Revised Desktop Reserve Model (RDRM; Hughes et al. 2012, 2014), which runs within the SPATSIM (Spatial and Time Series Information Modelling) software (Hughes and Forsyth 2006). The RDRM links hydrology, hydraulics, and ecology to estimate how much water a river needs to remain in a chosen ecological category (Hughes and Louw 2010). This study used the RDRM because the Omo-Gibe Basin has no field surveys or earlier e-flow studies. The model needs a monthly naturalized flow series for each site, that is, the flow that would occur without dams and abstractions. Most of the main stem of the Omo-Gibe is ungauged and the observed flows are altered by dams, so this series had to be simulated.

The Omo-Gibe Basin in south-western Ethiopia illustrates these pressures. It supplies more than 80% of the inflow to Lake Turkana, and its flow regime has already changed because of dams and land-use change (Worku et al. 2014). The Gibe I, II, and III hydropower plants have a combined capacity of more than 2,500 MW, and the Koysha dam is under construction (MoWE 2026). E-flows are needed in the basin for four interlinked reasons. First, they can guide the release rules of the dam cascade, so that hydropower operation does not remove

the seasonal pattern of flow downstream. Second, they help preserve the ecosystems and wildlife of the river and its floodplain. Third, they help maintain the water level of Lake Turkana, which depends on the Omo. Fourth, and perhaps most importantly, they can help sustain the communities that depend on the river for water, fishing, grazing, and farming.

Flow requirements have been considered in the basin mainly in connection with dam planning. The additional study of downstream impacts prepared for Gibe III (currently the most downstream hydropower scheme) proposed two measures. The first was a provisional minimum ecological flow of about 25 m³/s, to be released if the power plant is not operating. This value was taken from the lowest monthly average dry-season flow recorded between 1964 and 2001, and it is higher than the 15 m³/s required by the construction contract, which predated the study. The report describes it as an interim value, to be replaced once the geomorphological and ecological response of the river is better understood (Agriconsulting S.p.A. and Mid-Day International Consulting Engineers 2009). The second measure was a controlled flood of about 10 days in September, with flows of 1,000 to 1,200 m³/s at the dam, intended to reproduce an average natural flood in the lower Omo (Agriconsulting S.p.A. and Mid-Day International Consulting Engineers 2009). Both measures were based mainly on existing hydrological records. A doctoral study later examined how climate change and human activities affect streamflow, hydropower, and environmental flow in the basin (Woldemariam 2019). To our knowledge, however, no published e-flow assessment covers the basin from its headwaters to the lower Omo.

E-flow assessments depend on reliable flow records, and these are limited in the basin. Ethiopia's Ministry of Water and Energy (MoWE 2026) lists 52 hydrological stations in the Omo-Gibe Basin. Four of them have telemetry (Gibe I, Gibe III, Gibe near Asendabo, and Gibe near Abelti), and 12 are non-functional or closed (MoWE 2026). The stations are concentrated in the upper and central parts of the basin, around towns such as Jimma, Wolkite, and Areka, with few along the lower river toward Lake Turkana (MoWE 2026; Figure 1.1). Station condition is also a problem. Unstable riverbeds, damaged staff gauges, and variable observer performance reduce data quality. Of the 42 stations assessed by MoWE in 2026, 22 had been rehabilitated, 16 were awaiting rehabilitation, and 4 new stations had yet to be established (MoWE 2026). The Gilgel Gibe I reservoir shows the size of the gap: seven gauging stations are needed to measure its inflows and outflow, but only two are operating (MoWE 2026). The MoWE Basin Monitoring Strategy for 2026 to 2035 proposes 32 new stations, which would bring the basin total to 74 (MoWE 2026). The four stations with long records used in this study (Megecha, Ajancho, Asendabo, and Abelti) all lie in the northern part of the basin, so flows in most of the middle and lower reaches must be simulated.

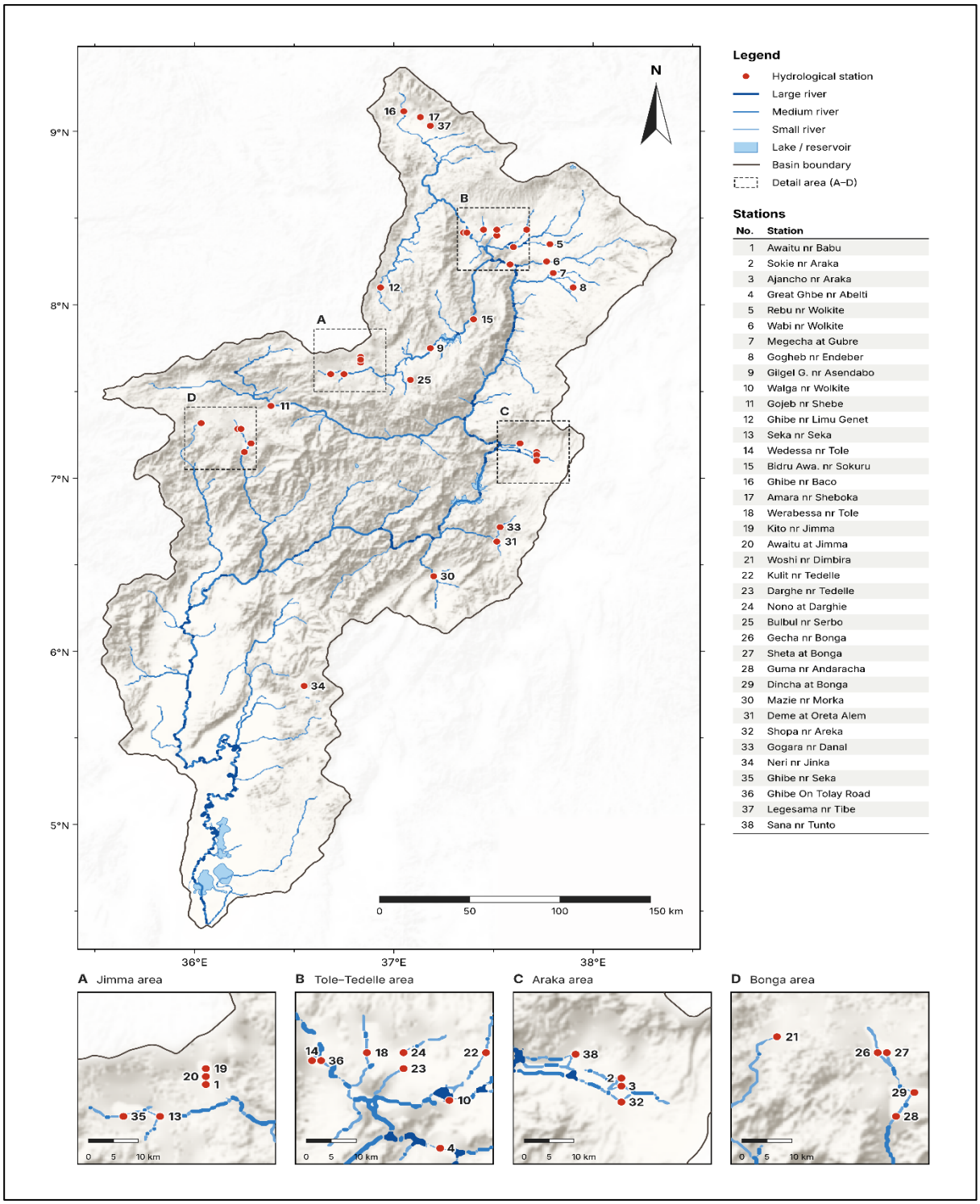


Figure 1.1. Hydrological station network of the Omo-Gibe Basin.

Source: Gauge stations from MoWE (2026). Basin and sub-basin boundaries from HydroBASINS and river network from HydroRIVERS (Lehner and Grill 2013). Reservoirs digitized from Google satellite imagery. Base layers: Imagery Google and Esri, USGS, NOAA.

Producing first-order e-flow estimates is only the first step. To guide dam releases, abstraction permits, and protection of the lower Omo, the estimates need to be updated as new observations become available and tested against scenarios of dam operation, abstraction, and climate. One way to organize this is a digital twin of the basin: a virtual representation of the river system that links data, models, and decision tools. The International Water Management Institute (IWMI) has built such a system for the Limpopo River Basin (Jolaiya et al. 2025). Building a twin was not an objective of this study. However, the data, models, and stakeholder information assembled here could form its first components. This report outlines what further work this would require.

1.1 Study objectives

The general objective of this study was to conduct a desktop e-flow assessment in the Omo-Gibe Basin. Three specific objectives were set. The first was to develop a machine learning model that estimates daily and monthly naturalized streamflow in ungauged catchments, using open-access satellite remote sensing data, and to apply it at 15 ungauged sites, five in each of the upper, middle, and lower sub-basins. The second was to use the RDRM in SPATSIM to estimate the environmental water requirements at these sites for ecological categories A to D, with hydraulic and ecological parameters transferred from South African rivers. The third was to document data availability, data gaps, and the barriers to applying e-flows in the basin through a questionnaire survey of water professionals.

2. Methodology

2.1 Study area and overall approach

The Omo-Gibe Basin lies in south-western Ethiopia and covers about 79,000 km², which gives it the third largest water potential of any basin in the country, after the Abay (Blue Nile) and Baro-Akobo (Worku et al. 2014). It drains the southern Ethiopian highlands southward and ends in Lake Turkana, in Kenya. The lake has no surface outlet, so the basin is endorheic, and it supplies more than 80% of the lake's inflow (Worku et al. 2014). Topography changes sharply along the basin. The northern upper part is mountainous, with elevations of up to 4,000 m, while the lower part is flat and lies at roughly 365 m where the river meets the lake (Worku et al. 2014).

The climate follows the same gradient. The highlands are wet, humid, and cool, whereas the lowlands are hot and semi-arid to arid (Worku et al. 2014). Most rain falls in the wet season from June to September, and the months from October to May are comparatively dry (Worku et al. 2014). On the aridity index, parts of the middle basin are humid, other regions are dry sub-humid, and the ungauged lower part is semi-arid (Worku et al. 2014).

Land use in the basin has changed considerably. Comparing land cover maps for the 1990s and 2009, Worku et al. (2014) found that forest, woodland, woody grassland, and shrubland declined by 53% while grassland and cropland expanded by 83%, with the largest changes in the humid parts of the basin. The basin was long regarded as close to pristine, but hydropower development began with the Gibe I, II, and III dams (Worku et al. 2014). The water of the basin supports hydropower, irrigated agriculture, downstream livelihoods, and the aquatic ecosystems of the lower Omo. Most of the main stem is ungauged, and the flow of the middle and lower reaches is altered by the dams.

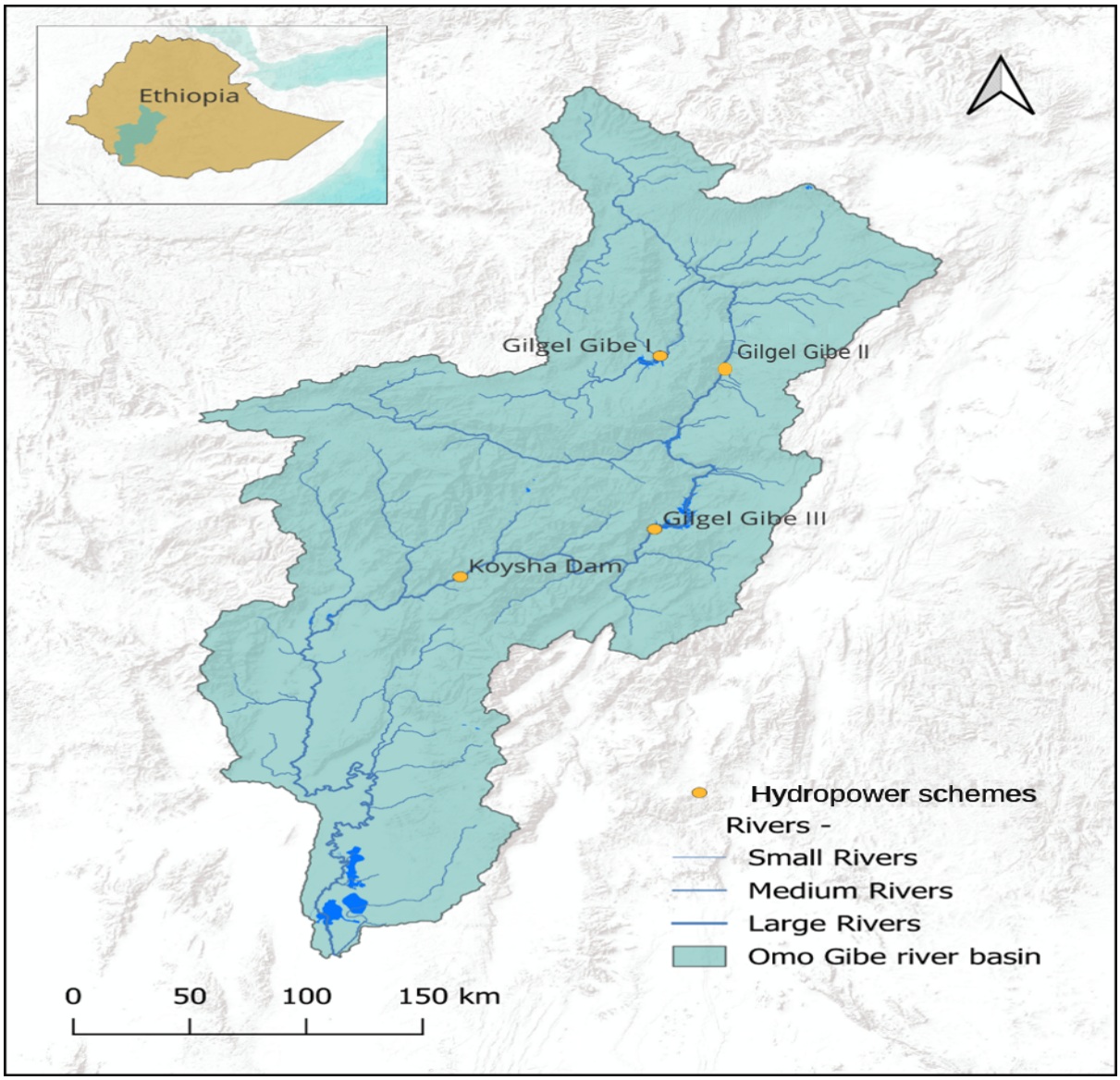


Figure 2.1. Map of the Omo-Gibe Basin showing the positions of the hydropower schemes, including the Koysha dam (under construction).

Source: Basin and sub-basin boundaries from HydroBASINS and river network from HydroRIVERS (Lehner and Grill 2013). Locations of the hydropower schemes digitized by the author with reference to Carr (2017). Reservoirs digitized from Google satellite imagery. Base layers: Imagery Google and Esri, USGS, NOAA.

The study had three linked components. First, naturalized flow was simulated at the assessment sites. Second, these flows were used as the hydrological input to a desktop environmental flow (e-flow) assessment. Third, a stakeholder questionnaire was administered to document data availability, data gaps, and the barriers to implementing e-flows in the basin.

2.2 Naturalized streamflow modeling

2.2.1 Observed streamflow data

Daily discharge records for 1981 to 2005 were obtained from four gauging stations, located in the upper watershed: Megecha near Gubere, Ajancho near Areka, Gilgel Gibe near Asendabo, and Gibe at Abelti (Table 2.1). The records were originally collected by the Ministry of Water and Energy (MoWE) of Ethiopia, the national authority that holds the hydrological station records for the basin (MoWE n.d.). The same stations have been used in earlier analyses of the flow regime of the basin (Worku et al. 2014).

Table 2.1. Gauging stations used to develop the naturalized flow model, with the area of the upstream catchment.

Station name	Location description	Station number	Latitude (°N)	Longitude (°E)	Catchment area (km²)
Megecha	Megecha near Gubere (upper basin)	091005	8.1633	37.8383	1179.96
Ajancho	Ajancho River near Areka (middle basin)	092010	7.0833	37.8000	565.42
Asendabo	Gilgel Gibe near Asendabo (upper basin)	091008	7.7667	37.2167	3123.16
Abelti	Gibe at Abelti (upper basin)	091001	8.2333	37.5833	16000.42

Source: Station numbers and coordinates from MoWE (n.d.). Catchment areas are those of the catchment polygons used in the model.

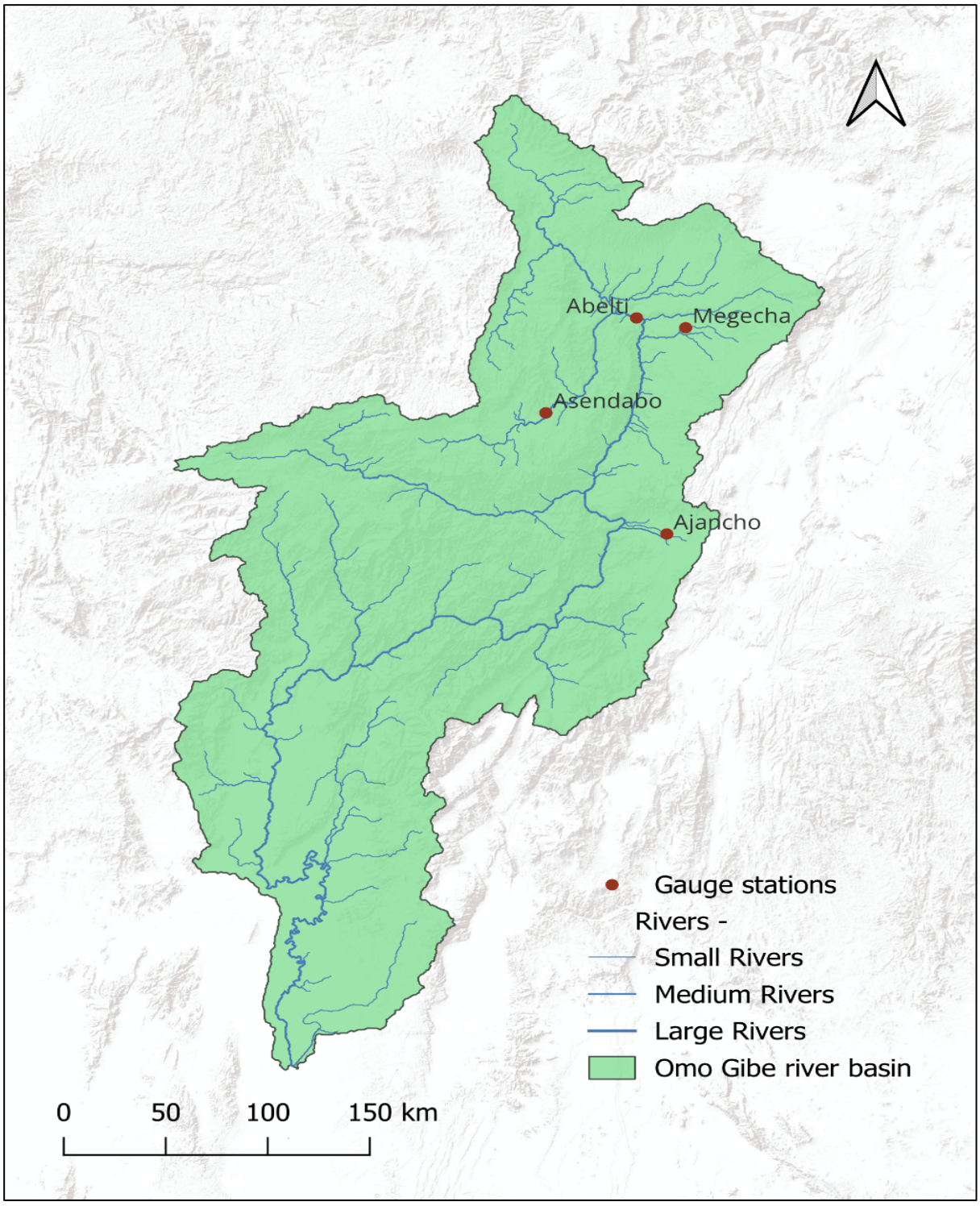


Figure 2.2. Locations of the river gauging stations used to develop the model. Three stations are in the upper basin and one (Ajancho) is in the middle basin.

Source: Gauge station locations from MoWE (n.d.). Basin boundary from HydroBASINS and river network from HydroRIVERS (Lehner and Grill 2013). Base layers: Imagery Google and Esri, USGS, NOAA.

The model was meant to simulate flow without major human influence, so the records had to be close to natural. Three points support this. First, the records end in 2005. The Gibe II dam began operating in 2010 (Avery and Tebbs 2018), and the model was trained only on the first two-thirds of each record, ending in 1997. Second, all four gauged catchments lie upstream of the Gibe II and Gibe III dams (Figures 2.1 and 2.2). Third, Worku et al. (2014) used the same stations, attributed the changes they found in the flow regime to land-cover change and not to regulation, and described the basin as long regarded as close to pristine. Two cautions apply. The Gibe I dam was commissioned in 2004 (Avery and Tebbs 2018), so the last two years of the record, which fall in the test period, will have affected flow at Abelti, located downstream of it. In addition, no records of abstraction were available, so upstream withdrawals could not be tested. We assumed that they were small relative to flow before 2005, and we treat this as a limitation (Section 4).

Naturalized flow here therefore means flow without dams and, as far as is known, without abstraction. It does not mean flow without land-cover change. The simulated flows represent the land cover and climate of 1981 to 2005, which already included the loss of forest and woodland and the expansion of cropland described in section 2.1 (Worku et al. 2014). The model has no land-cover predictor, so the relation between rainfall and runoff that it learned reflects that condition, and it was applied unchanged at the 15 assessment sites. The results are best read as natural flow under 1981 to 2005 conditions and not as flow under pristine or present-day land cover.

2.2.2 Candidate predictors and data sources

For the model, sixteen candidate variables were selected in six groups (Table 2.2). They were chosen to represent the main controls on how rainfall becomes streamflow, and so that every predictor could be obtained for the ungauged upstream catchments of the 15 assessment sites from global gridded products. This follows the approach of regional machine learning models, which combine meteorological time series with static catchment attributes so that a single model can represent many catchments (Kratzert et al. 2019b).

Catchment area was entered as its natural logarithm. It was used as the physical attribute because it sets the scale of the discharge a catchment can produce, and the logarithm reduces the large range in area among the stations and sites. Seasonality was represented by the sine and cosine of the day of the year, which describe the annual cycle without a break between 31 December and 1 January. This matters in a basin with a distinct wet season from June to September (Worku et al. 2014). Solar radiation and potential evapotranspiration describe the energy available to evaporate water, and so the water lost before it reaches the river. Soil moisture, together with its lagged value and its departure from the long-term mean, describes the antecedent wetness that controls how much rainfall becomes runoff. Precipitation was included as the main input of water, with lagged, cumulative, and anomaly terms. The short lag was meant to capture the immediate response of a catchment to rain, and the long cumulative sum to work as an alternative index of antecedent wetness to soil moisture, that is, of the water already stored in the catchment before a rain event.

Catchment area needs particular care because the model was applied to catchments much larger than the gauged ones. The four gauged catchments lie in the northern part of the basin (Table 2.1; Figure 2.2), whereas the assessment sites reach 70,567 km² (Table 3.1). A tree-based model cannot extrapolate. A catchment larger than the largest gauged one is treated by the model as if it had the area of the largest gauged catchment. Two features of the design limit the consequences. The model predicts runoff depth in millimeters and not discharge (section 2.2.3), so catchment area is not used to scale the output. It only helps the model to separate catchments, and discharge is obtained afterward by multiplying the predicted runoff depth by the true area. The simulated runoff depth at the large sites therefore depends mainly on rainfall, evaporation, and season. This is sound only if runoff depth does not change with

catchment size outside the training range, apart from the effect of climate. We could not test this, and it is probably least true in the semi-arid lower basin, where losses to the riverbed and floodplain increase with river length. It is flagged as a limitation (Section 4).

Table 2.2. Candidate predictor variables and their data sources.

Group	Variables	Data source
Physical	Natural logarithm of upstream catchment area	HydroSHEDS
Seasonality	Sine and cosine of day of year	Calculated from date
Solar forcing	Daily mean surface solar radiation and downward shortwave radiation flux	ERA5-Land
Evapotranspiration	Daily potential evapotranspiration, with rolling-mean and lagged versions	ERA5-Land
Soil moisture	Volumetric soil moisture (0-10 cm), with lagged and anomaly versions	ERA5-Land
Precipitation	Daily precipitation, with lagged, cumulative, and anomaly versions	CHIRPS

Daily precipitation was taken from the Climate Hazards Group InfraRed Precipitation with Station data (CHIRPS), which has a resolution of 0.05° (about 5 km; Funk et al. 2015). Solar radiation, potential evapotranspiration, and soil moisture were taken from the ERA5-Land reanalysis, which has a resolution of 0.1° (about 9 km; Muñoz-Sabater et al. 2021). Upstream catchment area was derived from the level 8 sub-basin polygons of HydroSHEDS, which are based on a digital elevation model with a resolution of 15 arc-seconds, or about 500 m (Lehner and Grill 2013). All hydro-climatic variables were spatially averaged over the entire catchment upstream of each gauging station and not extracted at a single point.

2.2.3 Feature selection

Highly collinear predictors were removed before the model was fitted. When two predictors carry the same information, their individual effects are hard to separate, and model estimation can become distorted once correlations exceed about 0.7 (Dormann et al. 2013). Redundant variables also add dimensions to the problem without adding information, and in a tree-based model they share importance between them, which makes the retained variables harder to interpret. Pairwise Spearman rank correlation coefficients were calculated among the 16 candidates. The correlation matrix was calculated from the training period only, that is, the first two thirds of each station record (1981 to late 1997), so the test years did not influence which predictors were kept. Spearman correlation was used because it measures monotonic relationships and is not affected by skewed distributions such as daily rainfall. When two variables had a coefficient above 0.75 or below -0.75, one of them was discarded. Five variables were retained for the final model: log catchment area, the sine transform of day of year, the 7-day rolling mean of potential evapotranspiration, 3-day lagged daily precipitation, and 60-day cumulative precipitation. Both a short lag and a long cumulative sum of

precipitation were kept because they represent different processes, and the 60-day sum was kept in preference to soil moisture, with which it was highly correlated (section 2.2.2).

The model predicted runoff depth and not discharge. Observed daily discharge Q (m^3/s) was converted to runoff depth R (mm/day) with $R = Q \times 86{,}400 / (A \times 1{,}000)$, where A is the catchment area in km^2, so that catchments of different sizes could be compared, and 86,400 is the number of seconds in a day. The target was the natural logarithm of one plus the runoff depth, $\ln(1 + R)$, which reduces the skew of the flow data (Log_Runoff in Figure A1). Predictions were back-transformed with exp(x) - 1. Back-transforming a model fitted in log space tends to underestimate the mean, so the result was multiplied by Duan's smearing factor, which is the mean of the exponentiated residuals of the training data (Duan 1983). The predicted runoff depth was then converted back to discharge with $Q = R \times A \times 1{,}000 / 86{,}400$. Monthly flows were calculated as the mean of the daily values.

2.2.4 Model structure and hyperparameters

Flow was modeled with extreme gradient boosting (XGBoost; Chen and Guestrin 2016), using the XGBRegressor class of the xgboost package in Python (xgboost version 3.4.1; Python version 3.13.9). XGBoost is not a random forest. A random forest grows many deep trees independently on bootstrap samples of the data and averages their predictions (Breiman 2001). Gradient boosting instead builds shallow trees one after another, and each new tree is fitted to the errors left by the trees before it, so that the ensemble gradually reduces its loss (Friedman 2001). XGBoost is a regularized and computationally efficient implementation of this idea (Chen and Guestrin 2016). Each tree used only a random fraction of the training samples and of the features, which gives the model a stochastic element as in stochastic gradient boosting (Friedman 2002), but the trees still depend on each other, and this is the main difference from a random forest. The algorithm was selected because it can represent nonlinear relationships and interactions among predictors and can be fitted to an observed record of this length (about 25 years of daily data per station, 1981 to 2005).

The hyperparameters were set as fixed values when the regressor was initialized in the model script and were not adjusted by an automated search. The model used 500 trees, a learning rate of 0.03, and a maximum tree depth of 6. Each tree was fitted to a random 80% of the training samples (subsample = 0.8) and a random 80% of the features (colsample_bytree = 0.8). The random seed was set to 42 so that the results could be reproduced (Table 2.3). All other settings were left at the package defaults.

Table 2.3. XGBoost hyperparameters used in the naturalized flow model.

Hyperparameter	Value	Description
n_estimators	500	Number of trees
learning_rate	0.03	Shrinkage applied to each tree
max_depth	6	Maximum depth of each tree
Subsample	0.8	Fraction of samples used per tree
colsample_bytree	0.8	Fraction of features used per tree

Hyperparameter	Value	Description
random_state	42	Seed for reproducibility

2.2.5 Training, testing, and evaluation

The first two-thirds of the observed record were used for training and the final third for testing, so that the model was evaluated in a later period than the one it was trained on. A 60-day buffer, equal to the longest predictor window, was left between the two sets. Daily flow and the rolling and cumulative predictors are strongly autocorrelated, and without a gap, information from the training period would leak into the test period (Roberts et al. 2017). Four configurations were compared: daily versus monthly time step, each with and without the Ajancho station. Model performance in the test period was evaluated at each station with the Nash-Sutcliffe efficiency (NSE) and with the coefficient of determination (R^2) and p-value of the regression of simulated on observed flow. NSE compares the squared errors of the model with the variance of the observed flow. A value of 1 means a perfect fit, and a value of 0 means that the model is no better than the mean of the observations. Because the errors are squared, NSE is most sensitive to errors in the high flows. To check bias and low-flow performance, NSE of log-transformed discharge and the percent bias (PBIAS) were also calculated. Simulated and observed hydrographs were compared visually. The configuration with the best performance was kept for the rest of the study.

The test-period scores reported here are slightly optimistic. The decisions to drop Ajancho, whose fit was poor (section 3.1.2), and to keep the monthly model were made after the test-period results had been seen, so the test data were not fully unseen when the final configuration was chosen.

A temporal test does not show how well the model transfers to catchments without a gauge, which is how it was used at the 15 assessment sites. A leave-one-station-out (spatial) cross-validation was therefore added (Roberts et al. 2017; Kratzert et al. 2019a). Each station of the selected model (Megecha, Asendabo, and Abelti) was left out in turn. The model was refitted on the training periods of the remaining stations, with the same predictors, hyperparameters, and smearing procedure, and its monthly simulation was compared with the observed flow at the omitted station over its full record and over its test period. Ajancho was treated in the same way, with the three stations of the selected model as training data. Because only two stations were left in each fit, the procedure was repeated with Ajancho included in the training data. The results are given in section 3.1.2.

2.2.6 Selection of the assessment sites and application of the model

The selected model was used to simulate naturalized flow at 15 sites, five in each of the upper, middle, and lower sub-basins (U_1 to U_5, M_1 to M_5, and D_1 to D_5; Table 2.4; Figure 2.3). The sites were chosen by random selection. Within each sub-basin, five locations on the river network were drawn at random, so that the choice did not depend on the judgment of the analyst and each sub-basin was represented by the same number of sites. Random selection has three consequences. First, the sites do not coincide with the decision points that matter most for management, such as dam outlets, the Kuraz irrigation area, the lower Omo floodplain, the delta, and protected areas, so the results describe the basin in general and not these places in particular. Second, some sites are on tributaries and not on the main stem (Table 2.4). Third, two pairs of sites (U_3 and U_5, and M_3 and M_4) have the same catchment area, so the 15 sites represent only 13 distinct catchments. The site with the largest catchment in each sub-basin (U_5, M_5, and D_5) was taken as the outlet of that sub-basin.

The upstream catchment of every site was taken from the HydroSHEDS level 8 sub-basin polygons, which are derived from a digital elevation model at 15 arc-second resolution (about 500 m). The CHIRPS (0.05°) and ERA5-Land (0.1°) data described in section 2.2.2 were then averaged over each upstream catchment, in the same way as for the gauged catchments, and the five predictors were calculated from these averages. The output was aggregated to a monthly series, which is the flow input required by the e-flow model. The annual coefficient of variation and the baseflow index of the simulated monthly series were also calculated for each site. The baseflow index was calculated with RDRM model in SPATSIM software.

Table 2.4. Location, catchment area, and position on the river network of the 15 assessment sites.

Site	Sub-basin	Latitude (°N)	Longitude (°E)	Catchment area (km²)	Main stem or tributary
U_1	Upper	8.56	37.31	6,763	Main stem
U_2	Upper	8.09	37.40	5,172	Tributary
U_3	Upper	7.89	37.60	21,429	Main stem
U_4	Upper	7.80	37.27	3,123	Tributary
U_5	Upper	7.73	37.56	21,429	Main stem
M_1	Middle	7.02	37.45	31,193	Main stem
M_2	Middle	7.05	37.42	33,302	Main stem
M_3	Middle	6.86	37.38	34,473	Main stem
M_4	Middle	6.74	37.25	34,473	Main stem
M_5	Middle	6.71	36.79	45,529	Main stem
D_1	Lower	6.53	36.49	49,440	Main stem
D_2	Lower	6.04	35.99	56,788	Main stem
D_3	Lower	5.38	35.94	60,852	Main stem
D_4	Lower	5.00	36.05	69,096	Main stem
D_5	Lower	4.72	35.99	70,567	Main stem

Source: Site selection by the author; catchment areas from the HydroSHEDS level 8 sub-basin polygons (Lehner and Grill 2013).

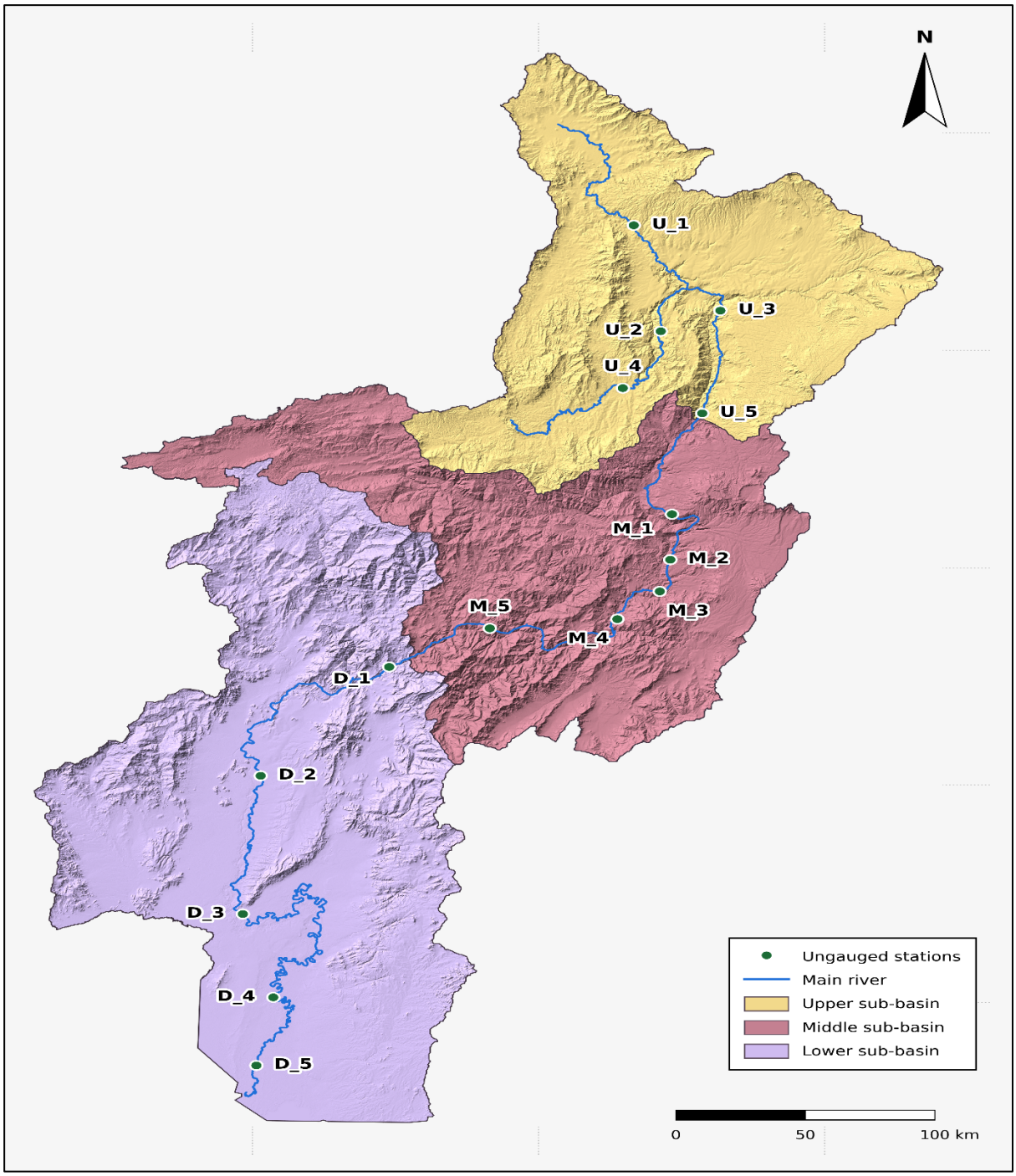


Figure 2.3. Selected sites for the e-flow assessment, with the Upper, Middle, and Lower Omo boundaries.

Source: Sub-basin and river layers from HydroBASINS and HydroRIVERS (Lehner and Grill 2013). The upper and middle sub-basins are separated at the Gibe II scheme and the middle and lower sub-basins at the Koysha dam. Site locations as listed in Table 2.4. Hill shade derived from the SRTM 30 m digital elevation model (Farr et al. 2007), obtained through Google Earth Engine.

2.3 Desktop e-flow assessment

2.3.1 Method selection

E-flow methods range from simple hydrological rules to detailed holistic assessments (Tharme 2003). A desktop hydrological method was adopted because the Omo-Gibe Basin has no field surveys, habitat data, or earlier e-flow studies, all of which the more detailed approaches require. The Revised Desktop Reserve Model (RDRM) within the SPATSIM software was used (SPATSIM version 3.00.0009). SPATSIM (Spatial and Time Series Information Modelling) was developed at the Institute for Water Research, Rhodes University, and links spatial and time-series data with hydrological models (Hughes and Forsyth 2006). The RDRM is built on the earlier desktop Reserve model (Hughes and Hannart 2003; Hughes and Louw 2010) and was revised to support rapid assessment (Hughes et al. 2012, 2014). It estimates the flow required for a river to remain in a chosen ecological category, from A (natural) to D (largely modified) (Kleynhans and Louw 2007).

2.3.2 Model inputs

Three sets of inputs were required. The first was the monthly naturalized flow series for each site (Section 2.2). The second was a set of hydraulic parameters describing the channel: geomorphic zone, valley slope, channel width and depth, and bankfull discharge. The third was a set of ecological parameters describing how aquatic biota respond to depth and velocity, namely the velocity-depth weights and the maximum baseflow percentile. Hydraulic parameters such as maximum channel width and valley slope were extracted from Google Earth imagery (captured on 1 January 2021, showing normal dry-season low-flow stage conditions). The geomorphological zone (ranging from 1 to 6) was determined based on the valley slope in accordance with the RDRM guidance, while bankfull discharge was automatically calculated by the RDRM using empirical hydraulic geometry relationships linked to the channel dimensions and hydrological variability, and the other parameters were set using the South African reference rivers described in section 2.3.3. Several of the values in Table B1 are doubtful, and their effect on the results is discussed in Section 4.

2.3.3 Transfer of South African parameters

The RDRM was developed and calibrated on South African rivers, as no calibrated parameter set exists for Ethiopian rivers. One South African reference river was therefore adopted for each sub-basin (Table 2.5). The reference rivers were chosen with the geomorphological zonation of Rowntree et al. (2000), which classifies channels from steep mountain streams (zone 1) to wide lowland rivers (zone 6). Channel type controls how depth and velocity change with discharge, and the transfer of parameters between rivers of similar type is the standard way of extending the desktop Reserve approach to unsurveyed rivers (Hughes and Hannart 2003; McCartney et al. 2009; Reitberger and McCartney 2011; Hughes et al. 2014). Thina-2 was matched to the steeper, narrower headwater channels of the upper sub-basin, Asseg-2 to the moderate-gradient middle reaches, and Augr-3 to the wide, low-gradient river of the lower basin. A different reference river was used in each sub-basin because channel form changes strongly along the river, from steep and narrow in the headwaters to wide and flat in the lowlands, and a single parameter set could not describe all three.

The reference rivers were chosen by channel type, not by climate or biota. The climate and flow descriptors below give context only and were not used as matching criteria. Thina-2 lies in the Thina River catchment of the Eastern Cape, where steep, deeply incised valleys store little sediment (Van der Waal 2014, 38–39). Rainfall in the upper catchment is about 707–928 mm per year and peaks in summer as high-intensity thunderstorms, with perennial flow and short, steep hydrographs (Van der Waal 2014, 39). The catchment is undammed, although

grazing-related degradation has made flow flashier (Van der Waal 2014, 39). Asseg-2 lies on the Assegaai River in Mpumalanga, in a summer-rainfall catchment with mean annual rainfall of about 960 mm (about 820–1,000 mm across the catchment) (Du Plessis 2023, 3–8). The Assegaai is impounded at Heyshope Dam, so the natural character of Asseg-2 depends on its position relative to the dam. Augr-3 represents a semi-arid lowland reach of the Orange River. Mean annual rainfall across the Orange basin is about 325 mm (range 35–1,040 mm), and most of the basin is semi-arid (Chisanga et al. 2022, 1279, 1284). The Orange is heavily regulated, with dams, evaporation and abstraction reducing natural flow by about half (Chisanga et al. 2022, 1277).

The velocity-depth weights and the baseflow limit were not set by the author. They were taken unchanged from the calibrated parameter sets for Thina-2, Asseg-2, and Augr-3 in the SPATSIM RDRM database (Table 2.6). The maximum baseflow parameter on the flow duration curve is 15 for Thina-2, 20 for Asseg-2, and 40 for Augr-3. The fast shallow (FS), fast intermediate (FI), and fast deep (FD) weights also differ among the sites. In the wet season they are 4, 5, and 6 for Thina-2, 24, 48, and 60 for Asseg-2, and 8, 14, and 19 for Augr-3. In the dry season, the FD weight is zero at all three sites. Hydrological variability also differs (8.31, 3.38, and 7.08). The geomorphic zones also differ among the sites within each Omo-Gibe sub-basin (Table B1), so using one reference river per sub-basin is a simplification. A sensitivity analysis, running each sub-basin with other reference rivers and varying the parameters, was beyond the scope of this study and is listed as a limitation (Section 4).

Table 2.5. South African reference rivers used for the hydraulic and ecological parameters.

Sub-basin	Sites	Reference Site & River
Upper	U_1 to U_5	Thina-2 (Thina River)
Middle	M_1 to M_5	Asseg-2 (Assegaai River)
Lower	D_1 to D_5	Augr-3 (Orange River reach)

Table 2.6. RDRM parameters used for the reference sites (SPATSIM parameter database)

Parameter	Thina-2	Asseg-2	Augr-3
Max baseflow FDC	15	20	40
Wet-season weights (FS / FI / FD)	4 / 5 / 6	24 / 48 / 60	8 / 14 / 19
Dry-season weights (FS / FI / FD)	3 / 1 / 0	11 / 9 / 0	3 / 2 / 0
Wet / dry stress at 0 FS	9 / 5	9 / 9	9 / 9
Valley slope (fraction)	0.001	0.005	0.002
Hydrological variability	8.31	3.38	7.08

2.3.4 Model runs and outputs

The model was run at each of the 15 sites for ecological categories A to D. The RDRM splits the requirement into a low-flow part, calculated for every month from the baseflow component of the natural flow, which maintains the baseflows of the river through the year, and a high-flow part, which provides flood events. Two results were extracted for each category: the low-flow

environmental water requirement (EWR) and the total EWR, which is the sum of the low-flow and high-flow parts (Hughes and Louw 2010; Hughes et al. 2014). Both were expressed in million cubic meters per year (Mm^3/yr) and as a percentage of natural mean annual runoff (MAR). Natural flow accumulates downstream, so the requirement of each sub-basin was represented by its outlet site (U_5, M_5, and D_5), and the other sites were used to show how the requirement varies within a sub-basin. The requirements were calculated from the simulated series as is and are given as single values. The bias of the flow model (section 3.1.3) and the year-to-year variability of the flow were not carried into them.

2.4 Stakeholder questionnaire

2.4.1 Questionnaire design and distribution

An online questionnaire was used to collect the views of people involved in water management in the basin. It was sent after the May 2026 e-flow workshop (Simaika and Haile 2026a) to its 25 participants. Responses were presented at the September 2026 workshop (Simaika and Haile 2026b). It contained 15 questions (Annex C) on three themes: the type and availability of data held, key data gaps, and the barriers to managing the basin and applying e-flows. Most questions were closed (single or multiple choice), and some asked respondents to select up to three options. Several open-ended questions asked for data sources, priority river reaches, training needs, and recommendations.

The introduction to the questionnaire explained its purpose, stated that participation was voluntary and could be stopped at any time, and said that individual responses would not be shared and that results would be presented only in aggregate form.

2.4.2 Data analysis

Responses were exported from the survey platform and checked for completeness. Fourteen of the 25 recipients responded (56%). Closed questions were analyzed with descriptive statistics, counting the respondents who selected each option, and responses were grouped by organization type (university or research institution, government ministry or public authority, consultancy or technical advisory firm, and independent expert). Ratings of monitoring coverage were compared between the reaches upstream and downstream of the Gibe cascade. Some questions asked for up to three choices, but the survey platform did not enforce this limit. All selections were counted, and the number of respondents who chose more than the allowed number is reported with each result. Free-text answers given instead of a listed option were kept as separate categories and are labeled as such. The open-ended answers (questions 10, 11, 14, and 15) were read several times and grouped into recurring themes such as priority reaches, named data sources, training needs, and recommendations. Because of the small sample, the results were treated as a snapshot of expert opinion and not as statistically representative. The numbers given for themes are numbers of respondents, and an answer could fall under more than one theme. Names and contact details given in answers to question 10 are not reported.

3. Results

The results are presented in three parts: first, the performance of the naturalized streamflow model and the flows it simulated at the 15 assessment sites; second, the environmental water requirements estimated with the Revised Desktop Reserve Model (RDRM); and third, the results of the stakeholder questionnaire.

3.1 Naturalized streamflow modeling

3.1.1 Predictor selection

Spearman rank correlation among the 16 candidate predictors showed several groups of highly collinear variables (Figure A1). After one variable was removed from each pair with a correlation above 0.75 or below -0.75, five predictors remained: log catchment area, the sine transform of day of year, the 7-day rolling mean of potential evapotranspiration, 3-day lagged precipitation, and 60-day cumulative precipitation. The retained predictors were only weakly to moderately correlated with one another ($|\rho| \leq 0.67$), and their correlations with log transformed runoff were 0.15, -0.61, -0.62, 0.35, and 0.80, respectively.

Although certain variables such as log catchment area ($\rho = 0.15$) and 3-day lagged precipitation ($\rho = 0.35$) displayed lower Spearman rank correlations with runoff, they were intentionally retained because they capture independent, physically distinct dimensions of the hydrological cycle essential for a nonlinear tree-based model. Specifically, the 3-day lag captures the immediate catchment response and storm routing time, whereas the 60-day sum serves as an index of long-term antecedent wetness. Alternative indicators such as soil moisture were excluded due to excessive rank collinearity with the 60-day sum ($\rho = 0.78$ to 0.85), and log catchment area was preserved to provide the fundamental spatial scaling required to convert unit runoff into volumetric discharge across sub-basins of different sizes.

3.1.2 Model performance

Four model configurations were compared: daily and monthly time steps, each with and without the Ajancho station (Figure 3.1). Ajancho was the weakest station in both time steps. When it was included, the coefficient of determination (R^2) at the four stations ranged from 0.01 to 0.79 at the daily time step and from 0.13 to 0.88 at the monthly time step ($p < 0.001$ for all stations). The reason for the weak fit at Ajancho was not examined here. It may reflect uncertainty in the record combined with the station's very low flow and by far the smallest catchment (565 km²), a regime in which regression trees are sensitive to minor noise in the data and to coarse spatial scaling. Dropping the station also removes the only small, low-flow catchment from the training data, which matters because small tributaries are simulated at some assessment sites, for example U_4 (3,123 km²). This is treated as a limitation in section 4. Leaving Ajancho out raised the mean R^2 of the models from 0.50 to 0.64 at the daily time step and from 0.66 to 0.84 at the monthly time step, but this is mainly because the poorly fitting station was no longer in the average. At the three remaining stations the mean R^2 was 0.66 with and 0.64 without Ajancho at the daily time step, and 0.84 with and 0.84 without it at the monthly time step, and R^2 at Megecha fell from 0.84 to 0.79 (Table A2). Monthly R^2 was higher than daily R^2 at every station in both configurations. For the finalized monthly model without Ajancho, all station regressions yielded high statistical significance (Megecha: $p = 4.36 \times 10^{-27}$, Asendabo: $p = 5.69 \times 10^{-37}$, and Abelti: $p = 8.84 \times 10^{-48}$). The monthly model without Ajancho had the highest mean R^2 of the four configurations. It was selected and used for all later steps, and only its results are shown below. Because this choice was made after the test-period results had been compared, the scores below are slightly optimistic (section 2.2.5).

A leave-one-station-out cross-validation was run to test transfer to catchments without a gauge, and it performed poorly. At the omitted stations NSE ranged from -1,216 to -0.4, PBIAS

from -90% to +3,586%, and R^2 from 0.36 to 0.86. The high R^2 at Megecha, Asendabo, and Abelti shows that the seasonal timing was reproduced when each was left out, but the flow volume was not, so none of the three stations was reproduced adequately. Adding Ajancho to the training data reduced the bias at Megecha (from +3,586% to +76%) and changed the results at Asendabo and Abelti very little. This is expected with so few stations. Only three or four gauged catchments were available, all in the northern basin, so each fit saw only two or three distinct catchment areas. A tree-based model cannot extrapolate beyond the areas it has seen, so the omitted station was effectively given the runoff level of the training station with the nearest area.

The temporal split, and not leave-one-station-out cross-validation, was used to develop and select the model for three reasons. First, the model is trained to reproduce the timing and seasonal shape of flow from rainfall and evaporation, and a temporal split tests this directly, using years the model has not seen at stations whose records are available. Second, cross-validation in space needs enough gauged catchments to cover the range of conditions to which the model is applied. Only three or four were available, so each left-out fit was trained on two or three catchment areas, and the result mainly reflects how close the omitted catchment's area was to a training catchment. A score obtained in this way is a weak basis for choosing between model configurations. Third, the temporal split allowed the four configurations to be compared on the same footing. Leave-one-station-out was therefore treated as a separate check of spatial transfer and not as a selection criterion. Its poor result is a limitation of the study, and it means that the flows at the 15 assessment sites cannot be called validated. The temporal test shows only how well the model reproduces flow at stations whose earlier records it has seen. The site flows are therefore treated as first estimates and were checked against GLDAS-2.1 runoff, the Gibe III inflow estimate, and published estimates of the inflow to Lake Turkana (section 4). More gauged stations in the middle and lower basin would be needed to test spatial transfer properly.

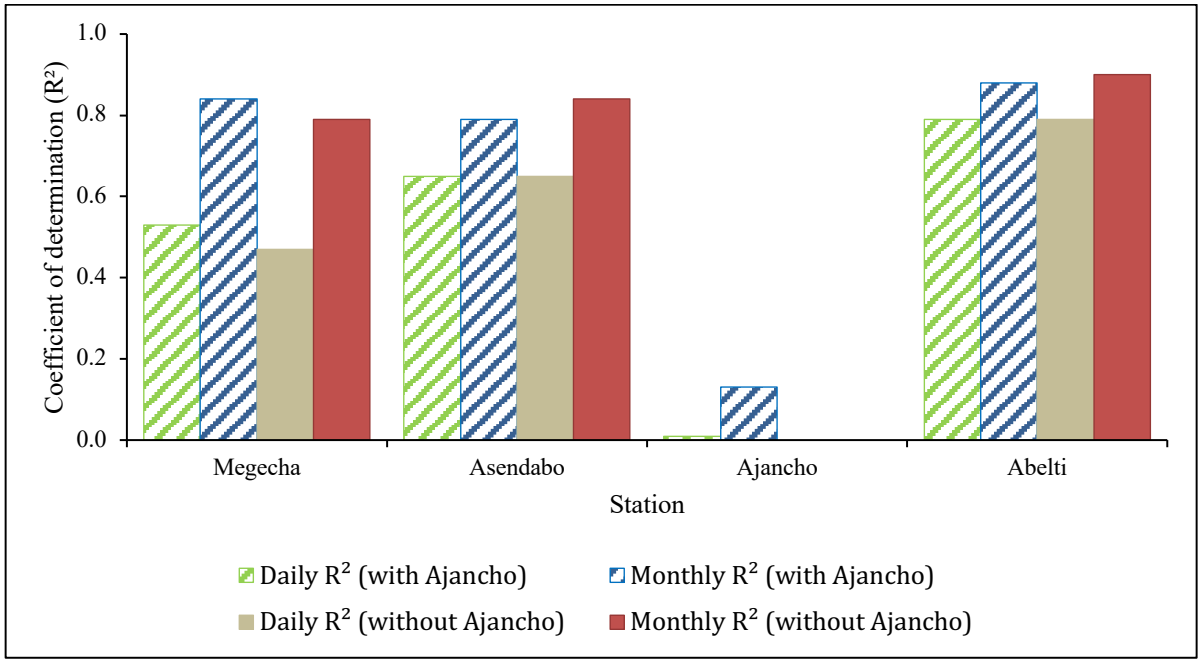


Figure 3.1. Daily and monthly coefficient of determination (R^2) at each gauging station for the models with (hatched bars) and without (solid bars) the Ajancho station.

3.1.3 Simulated and observed hydrographs

Figures 3.2 to 3.4 compare simulated and observed monthly flows for the selected model over the whole record, including the training period. Gaps in the observed record are shown as breaks in the observed line. In the test period, which began in late 1997, the Nash-Sutcliffe efficiency (NSE) was 0.78 at Megecha, 0.82 at Asendabo, and 0.89 at Abelti, and R^2 was 0.79, 0.84, and 0.90, respectively. The simulated flows followed the timing and the seasonal shape of the observed flows at all three stations. Individual peaks were not always matched: the 1998 peak at Asendabo was underestimated, for example, and one of the peaks in 2003 at Megecha was overestimated. Because NSE is based on squared errors, it is most sensitive to the high flows. To complement this, the logarithmic Nash-Sutcliffe efficiency (Log-NSE) and percent bias (PBIAS) were evaluated: Log-NSE values were 0.61 at Megecha, 0.79 at Asendabo, and 0.68 at Abelti, while PBIAS was +6.0% at Megecha, -13.2% at Asendabo, and -13.3% at Abelti. These metrics indicate that seasonal timing and high flows were reproduced well, but the negative PBIAS at the two larger stations indicates a systematic tendency to underestimate total runoff volume. The environmental water requirements in section 3.2 were calculated from the simulated flow without correcting for this bias.

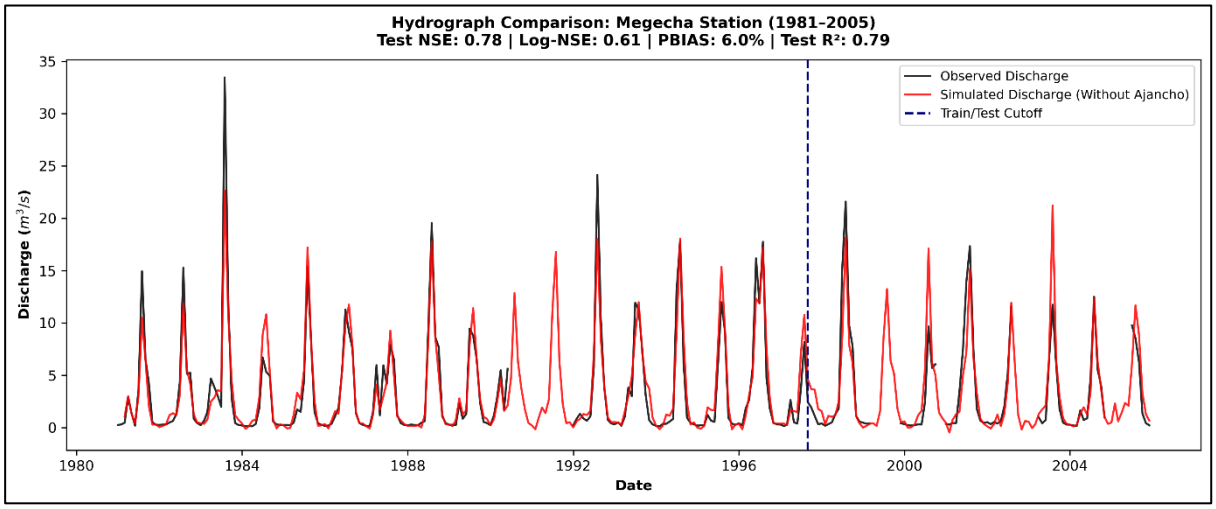


Figure 3.2. Observed and simulated monthly discharge at Megecha for the selected model (without Ajancho). The dashed vertical line marks the end of the training period, and breaks in the observed line are gaps in the record.

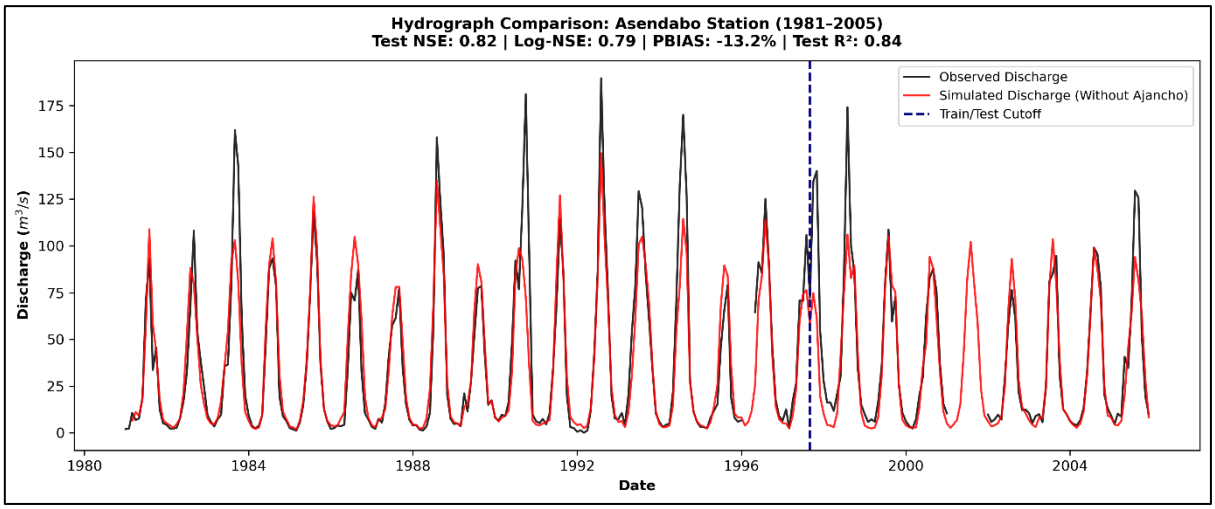


Figure 3.3. Observed and simulated monthly discharge at Asendabo for the selected model (without Ajancho). The dashed vertical line marks the end of the training period, and breaks in the observed line are gaps in the record.

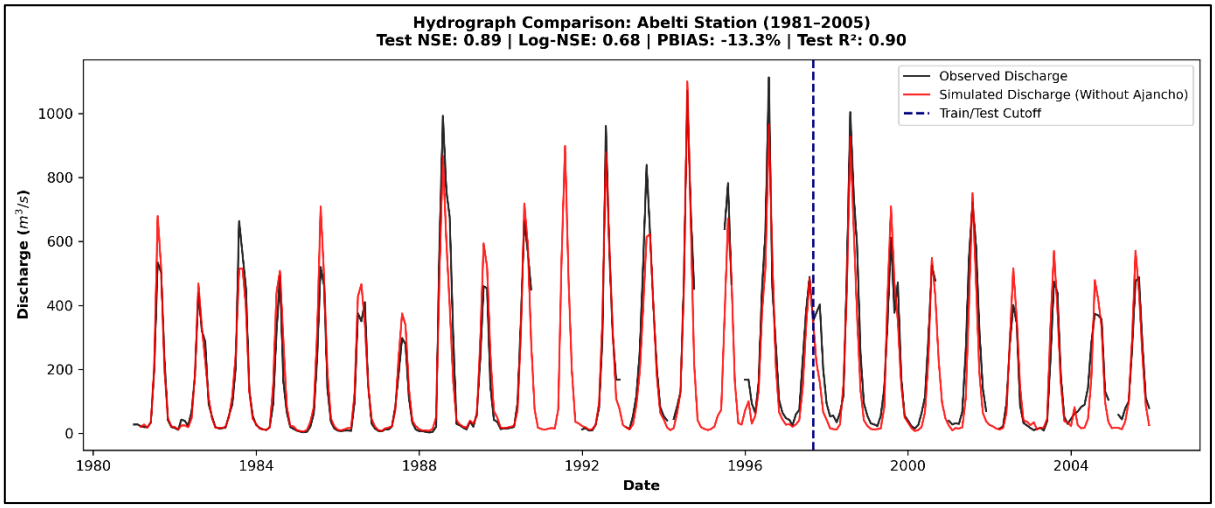


Figure 3.4. Observed and simulated monthly discharge at Abelti for the selected model (without Ajancho). The dashed vertical line marks the end of the training period, and breaks in the observed line are gaps in the record.

3.1.4 Naturalized flow at the assessment sites

The selected model was used to simulate naturalized flow at the fifteen assessment sites (Figure 3.5; Table 3.1). Natural mean annual runoff (MAR) increased downstream, from the upper to the lower sub-basin. At the site with the largest catchment in each sub-basin (U_5, M_5, and D_5), which represents the outlet of the sub-basin, the simulated MAR was 7,517 Mm³ (238 m³/s) in the upper sub-basin, 13,683 Mm³ (434 m³/s) in the middle sub-basin, and 18,430 Mm³ (584 m³/s) in the lower sub-basin. These sites drain 21,429, 45,529, and 70,567 km², respectively. Within a sub-basin, MAR does not rise steadily with site number. The sites were chosen at random on the river network, so some of them are on tributaries and not on the main stem. The smallest MAR, 1,222 Mm³, belongs to U_4, which has the smallest catchment (3,123 km²). U_3 and U_5 have the same catchment area, as do M_3 and M_4, and their MARs differ by less than 0.1%. D_4 and D_5 lie close together, and their MARs differ

by less than 1% (18,598 and 18,430 Mm³) even though D_5 has the slightly larger catchment, because the added semi-arid area lowers the mean runoff depth. This difference is much smaller than the uncertainty of the model. The error bars in Figure 3.5 show the interannual variability of the simulated annual flow (±1 SD).

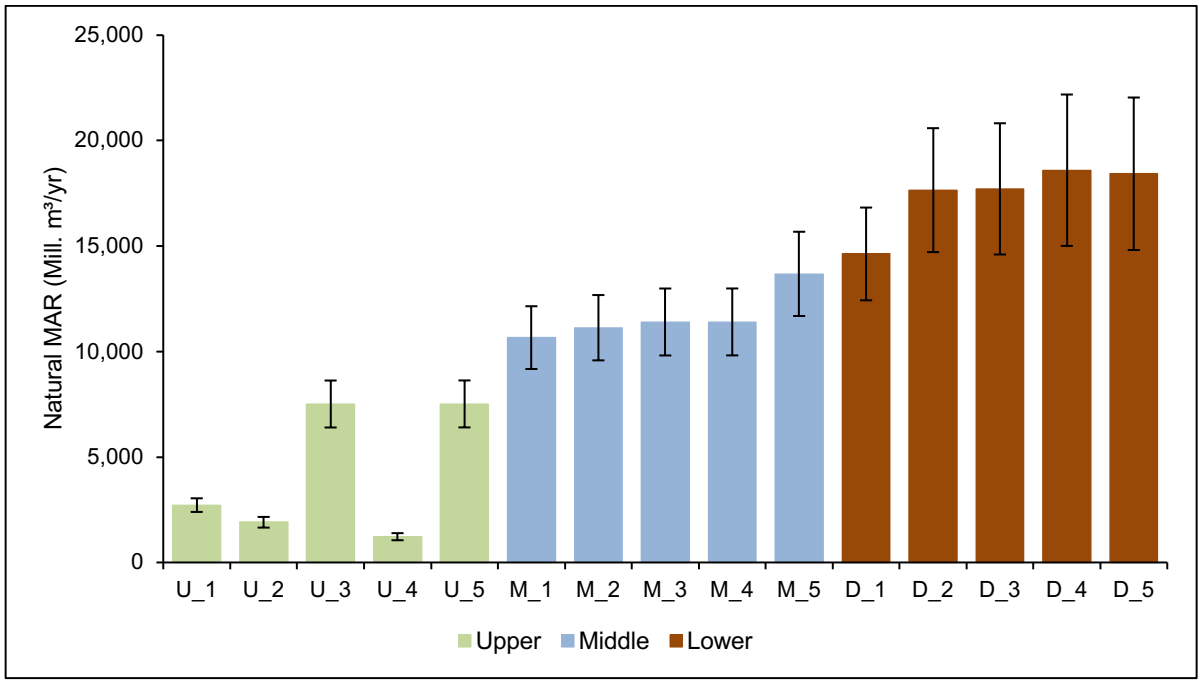


Figure 3.5. Natural mean annual runoff (MAR) at the 15 assessment sites, simulated with the naturalized flow model. Error bars show ±1 standard deviation of the simulated annual flow.

Flow was strongly seasonal (Figure 3.6). At the outlet of each sub-basin (U_5, M_5, and D_5), the simulated mean monthly flow peaked in August and was lowest in February. The four months from July to October carried 79% of the annual flow at U_5, 70% at M_5, and 61% at D_5. Across the 15 sites, the annual coefficient of variation of the simulated flow ranged from 0.12 to 0.20, and the baseflow index ranged from 0.32 to 0.39. These statistics come from simulated and not observed flow. The annual coefficients of variation are low and are probably smaller than the true values, because tree-based models damp extremes (section 4).

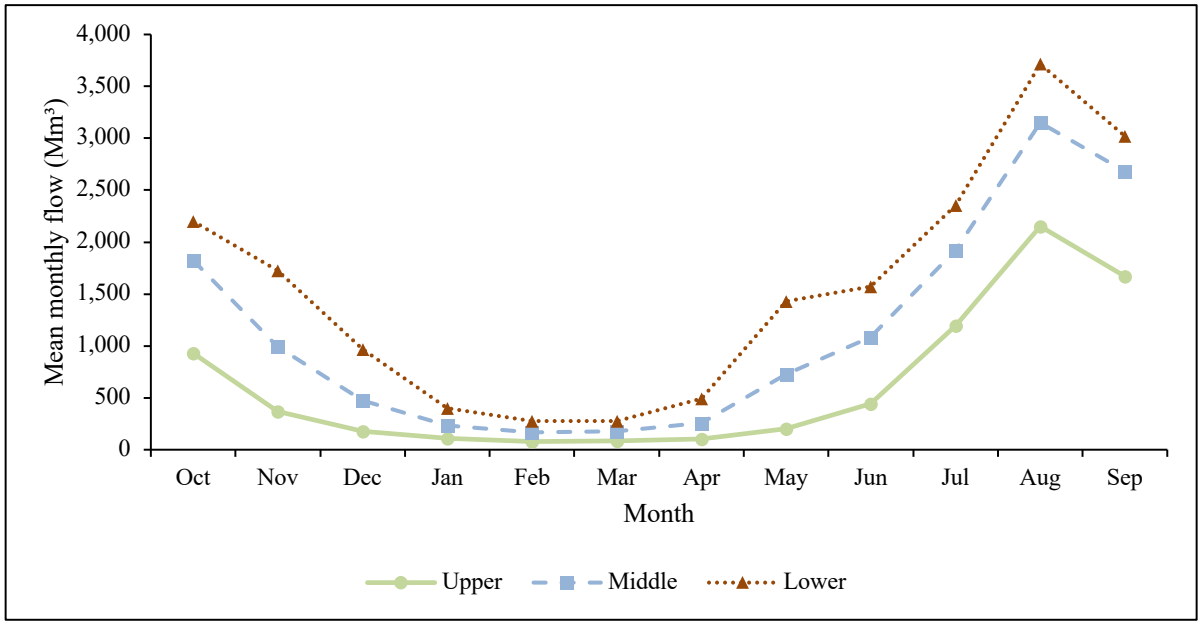


Figure 3.6. Mean monthly natural flow at the outlet of the upper (U_5), middle (M_5), and lower (D_5) sub-basins, by month of the hydrological year (October to September).

Table 3.1. Catchment area and natural mean annual runoff (MAR) at the 15 assessment sites, with the annual coefficient of variation (CV) and baseflow index (BFI) of the simulated flow.

Site	Catchment area (km²)	MAR (Mm³/yr)	MAR (m³/s)	Annual CV	BFI
U_1	6,763	2,719	86.2	0.12	0.32
U_2	5,172	1,906	60.4	0.13	0.35
U_3	21,429	7,513	238.1	0.15	0.35
U_4	3,123	1,222	38.7	0.14	0.35
U_5	21,429	7,517	238.2	0.15	0.35
M_1	31,193	10,661	337.8	0.14	0.35
M_2	33,302	11,131	352.7	0.14	0.35
M_3	34,473	11,403	361.3	0.14	0.36
M_4	34,473	11,405	361.4	0.14	0.36
M_5	45,529	13,683	433.6	0.15	0.37
D_1	49,440	14,629	463.6	0.15	0.36
D_2	56,788	17,654	559.4	0.17	0.37
D_3	60,852	17,714	561.3	0.18	0.37
D_4	69,096	18,598	589.4	0.19	0.38
D_5	70,567	18,430	584.0	0.20	0.39

3.2 Environmental water requirements

The RDRM was run at the 15 sites for ecological categories A to D. For each category the model gives a low-flow requirement, which maintains the baseflows of the river, and a total requirement, which also includes the high-flow (flood) component (section 2.3.4). Both are reported in million cubic meters per year (Mm³/yr) and as a percentage of natural MAR. Because natural flow accumulates downstream, the requirement of a sub-basin is described by its most downstream site (U_5, M_5, and D_5). The results for all sites are given in Tables 3.2 and 3.3. No Present Ecological State or Ecological Importance and Sensitivity assessment was made, so no ecological category is recommended for any reach, and the results were not compared with present-day or regulated flows (section 4).

3.2.1 Total-flow requirements

The total-flow EWR decreased from category A to category D at every site (Table 3.2; Figures 3.7 and 3.8). At the outlets of the three sub-basins, the category A requirement was 2,707 Mm³/yr (36.0% of MAR) at U_5, 4,903 Mm³/yr (35.8%) at M_5, and 6,944 Mm³/yr (37.7%) at D_5. Under category D, it fell to 1,450 Mm³/yr (19.3%), 2,638 Mm³/yr (19.3%), and 3,324 Mm³/yr (18.0%), respectively. In relative terms, the requirement was therefore similar in the three sub-basins. Across all 15 sites, the category A requirement ranged from 28.9% of MAR (M_1) to 50.5% (U_4). D_5 needed more water than D_4 (6,944 and 5,436 Mm³/yr in category A) although the two sites have almost the same MAR. The difference comes from the channel inputs to the model: the maximum channel width was about 912 m at D_5 and about 100 m at D_4, and the valley slope was lower at D_5 (0.004 against 0.008).

Table 3.2. Total-flow environmental water requirements in Mm³/yr, with the percentage of natural MAR in brackets.

Site	Category A	Category B	Category C	Category D
U_1	1,227 (45.1)	1,006 (37.0)	826 (30.4)	660 (24.3)
U_2	714 (37.5)	611 (32.0)	503 (26.4)	372 (19.5)
U_3	2,387 (31.8)	1,939 (25.8)	1,627 (21.7)	1,310 (17.4)
U_4	617 (50.5)	509 (41.6)	400 (32.7)	303 (24.8)
U_5	2,707 (36.0)	2,125 (28.3)	1,778 (23.6)	1,450 (19.3)
M_1	3,086 (28.9)	2,488 (23.3)	2,085 (19.6)	1,673 (15.7)
M_2	3,893 (35.0)	3,041 (27.3)	2,519 (22.6)	2,090 (18.8)
M_3	3,907 (34.3)	3,038 (26.6)	2,490 (21.8)	2,073 (18.2)
M_4	4,294 (37.6)	3,449 (30.2)	2,869 (25.2)	2,346 (20.6)
M_5	4,903 (35.8)	3,907 (28.6)	3,293 (24.1)	2,638 (19.3)
D_1	5,067 (34.6)	3,838 (26.2)	3,113 (21.3)	2,429 (16.6)
D_2	5,623 (31.9)	4,378 (24.8)	3,536 (20.0)	2,829 (16.0)
D_3	5,883 (33.2)	4,541 (25.6)	3,652 (20.6)	2,863 (16.2)
D_4	5,436 (29.2)	4,265 (22.9)	3,446 (18.5)	2,812 (15.1)
D_5	6,944 (37.7)	5,383 (29.2)	4,289 (23.3)	3,324 (18.0)

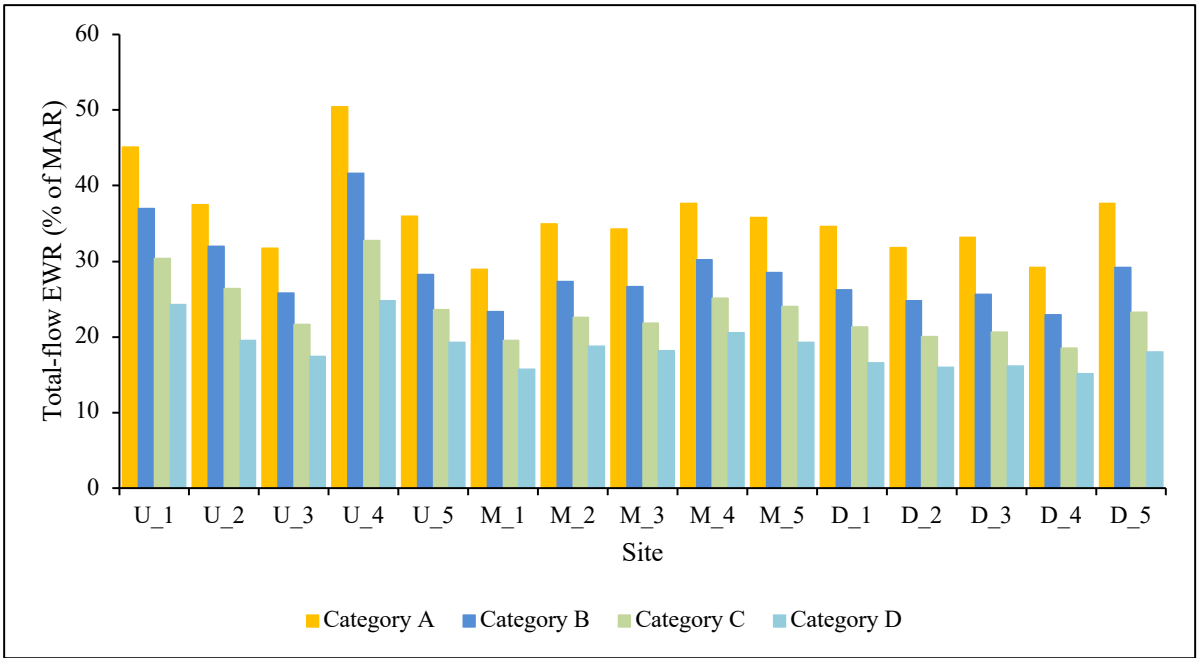


Figure 3.7. Total-flow environmental water requirement as a percentage of natural MAR at the 15 sites, for ecological categories A to D.

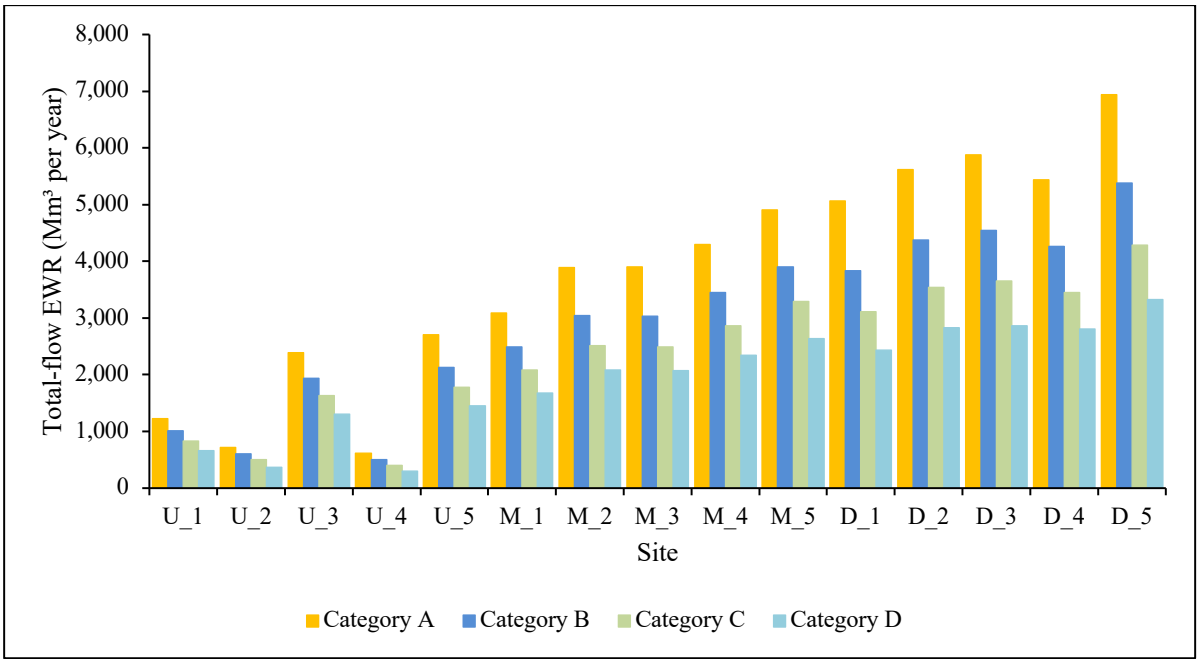


Figure 3.8. Total-flow environmental water requirement (Mm³ per year) at the 15 sites, for ecological categories A to D.

The total-flow EWR increased almost in proportion to natural MAR (Figure 3.9), and MAR explained 96% of the variation in every category. This is expected because the naturalized flow is the main input of the model. The relationship is not exact because channel inputs differ among sites and the transferred parameters differ among sub-basins, each of which used a different South African reference river (section 2.3.3).

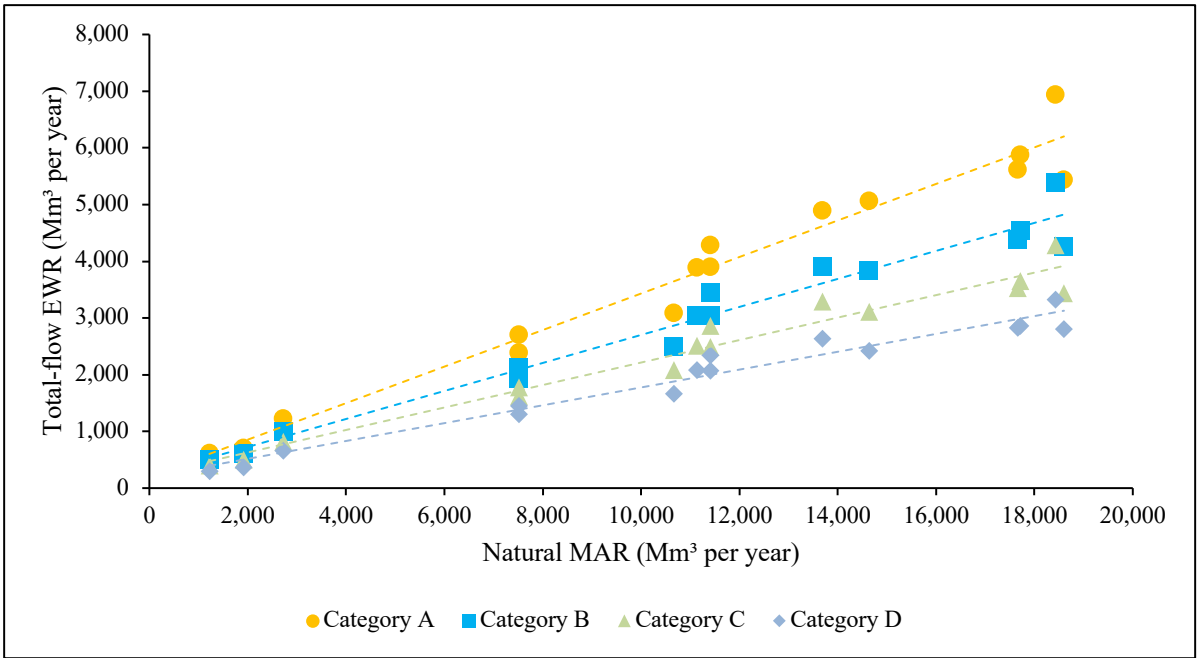


Figure 3.9. Relationship between total-flow environmental water requirement and natural MAR at the 15 sites, for ecological categories A to D. Dashed lines are linear regressions (R^2 = 0.96 for all four categories).

3.2.2 Low-flow requirements

The low-flow EWR was much smaller than the total-flow EWR (Table 3.3; Figures 3.10 and 3.11). At the outlets of the sub-basins, the category A low-flow requirement was 317 Mm³/yr (4.2% of MAR) at U_5, 906 Mm³/yr (6.6%) at M_5, and 2,576 Mm³/yr (14.0%) at D_5. Under category D, it was 8 Mm³/yr (0.1%), 292 Mm³/yr (2.1%), and 626 Mm³/yr (3.4%). The low-flow requirement fell more steeply between categories than the total-flow requirement. Averaged over the 15 sites, low flows made up 21% of the total requirement in category A and 8% in category D. The low-flow requirement was less closely related to MAR than the total requirement (R^2 = 0.73 in category A and 0.50 in category D). Expressing it against the natural baseflow, calculated as MAR multiplied by the baseflow index, gave a similar relationship (R^2 = 0.76 and 0.54). The low-flow component depends more on the channel inputs of the model. At D_5, for example, the category A low-flow requirement was about twice that at D_4 (2,576 and 1,262 Mm³/yr) for almost the same MAR.

Table 3.3. Low-flow environmental water requirements in Mm³/yr, with the percentage of natural MAR in brackets.

Site	Category A	Category B	Category C	Category D
U_1	185 (6.8)	119 (4.4)	72 (2.6)	37 (1.4)
U_2	164 (8.6)	130 (6.8)	100 (5.2)	33 (1.8)
U_3	240 (3.2)	119 (1.6)	72 (1.0)	19 (0.3)
U_4	155 (12.6)	108 (8.8)	59 (4.8)	20 (1.6)
U_5	317 (4.2)	140 (1.9)	47 (0.6)	8 (0.1)
M_1	696 (6.5)	457 (4.3)	331 (3.1)	219 (2.1)
M_2	415 (3.7)	149 (1.3)	61 (0.5)	23 (0.2)
M_3	440 (3.9)	156 (1.4)	62 (0.5)	24 (0.2)
M_4	804 (7.1)	551 (4.8)	405 (3.5)	272 (2.4)
M_5	906 (6.6)	606 (4.4)	442 (3.2)	292 (2.1)
D_1	1,374 (9.4)	710 (4.9)	396 (2.7)	219 (1.5)
D_2	1,411 (8.0)	758 (4.3)	424 (2.4)	223 (1.3)

Site	Category A	Category B	Category C	Category D
D_3	1,758 (9.9)	1,088 (6.1)	637 (3.6)	343 (1.9)
D_4	1,262 (6.8)	595 (3.2)	320 (1.7)	168 (0.9)
D_5	2,576 (14.0)	1,750 (9.5)	1,100 (6.0)	626 (3.4)

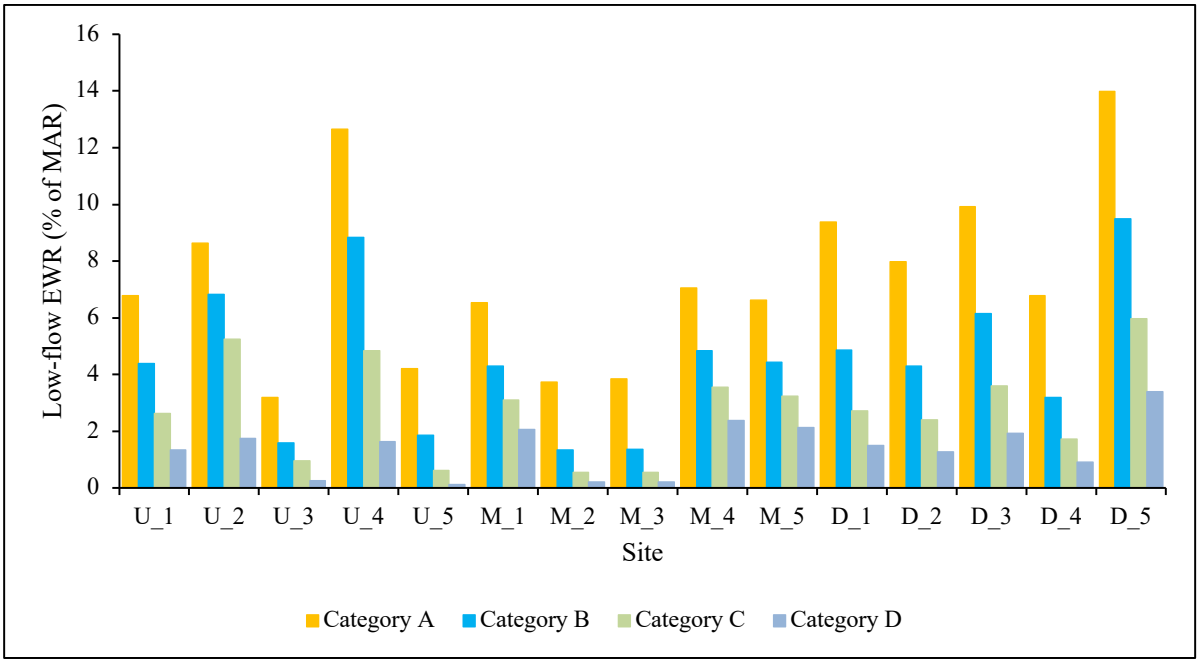


Figure 3.10. Low-flow environmental water requirement as a percentage of natural MAR at the 15 sites, for ecological categories A to D.

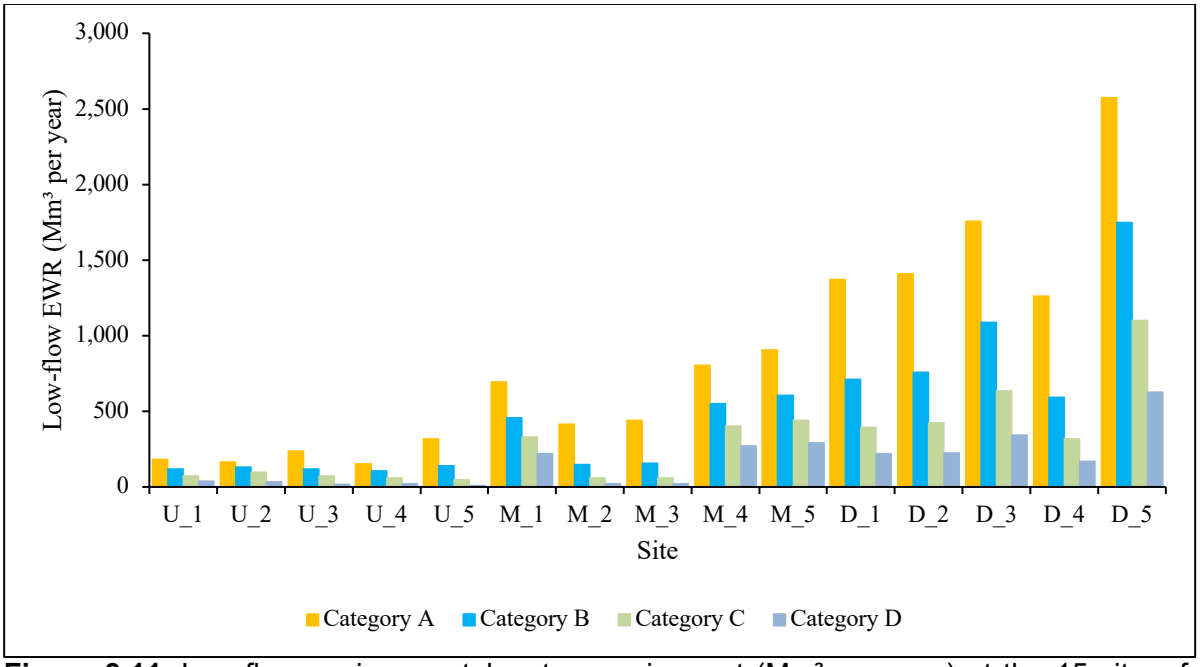


Figure 3.11. Low-flow environmental water requirement (Mm³ per year) at the 15 sites, for ecological categories A to D.

3.2.3 Seasonal distribution of the requirements

The monthly total-flow requirement followed the seasonal pattern of the natural flow (Figures 3.12 to 3.14), as expected because the RDRM takes the naturalized flow series as its main input. The figures show the median (50 percent exceedance) monthly values at the outlet of each sub-basin (U_5, M_5, and D_5). At all three sites, natural flow peaked in August and was lowest in February, and the requirement of every category also peaked in August. In August, the category A requirement was 912 Mm³ at U_5 against 2,025 Mm³ of natural flow, 1,492 against 3,071 Mm³ at M_5, and 2,058 against 3,846 Mm³ at D_5. In February, the category A requirement was 1, 33, and 36 Mm³ at the three sites, against natural flows of 74, 165, and 278 Mm³. The difference between the categories was largest in the wet season and small in the dry season.

The dry-season requirements were low relative to natural flow. In February the category A requirement was 2% of natural flow at U_5, 20% at M_5, and 13% at D_5. At U_5 it was 1.4 Mm³, which is about 0.6 m³/s, against about 30 m³/s of natural flow and against the interim minimum release of 25 m³/s proposed for Gibe III, which amounts to about 65 Mm³ in a 30-day month (Agriconsulting S.p.A. and Mid-Day International Consulting Engineers 2009). The requirement also fell sharply from February to March at all five lower sites, for example from 36.1 to 4.3 Mm³ at D_5 in category A, although natural flow rose slightly (from 278 to 280 Mm³ at D_5). The fall does not appear at the upper and middle sites or in the natural flow that was used as input, so it arises inside the RDRM and not in the flow model. We could not trace its cause from the model output. It is probably an artifact of the parameters transferred from the reference river Augr-3 (a semi-arid, highly variable reach with low dry-season weights; Table 2.6) and of the channel inputs in Table B1, and we treat the March values at the lower sites, and the very low dry-season values at some upper and middle sites, as unreliable (section 4).

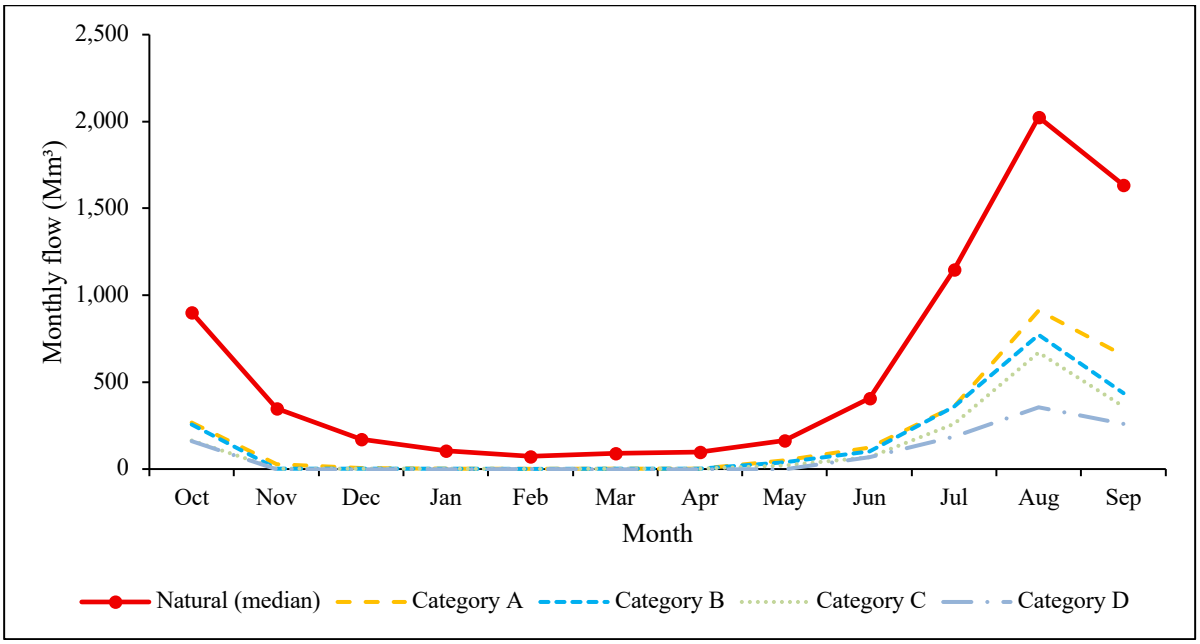


Figure 3.12. Median monthly natural flow and total-flow requirement for ecological categories A to D at the outlet of the upper sub-basin (U_5), by month of the hydrological year.

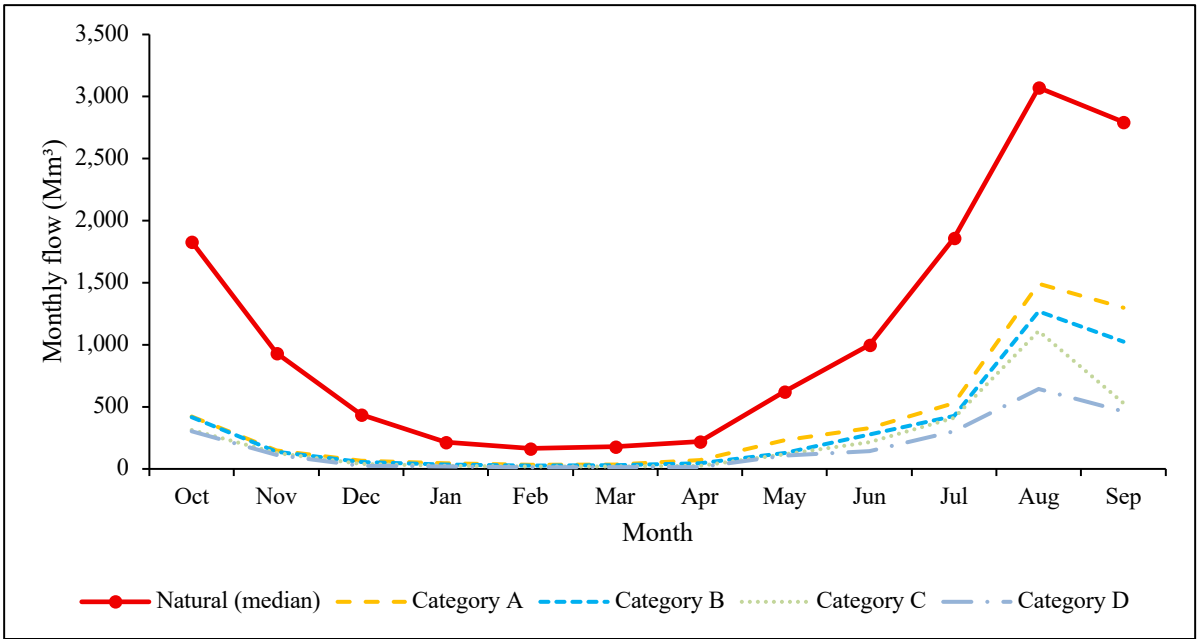


Figure 3.13. Median monthly natural flow and total-flow requirement for ecological categories A to D at the outlet of the middle sub-basin (M_5), by month of the hydrological year.

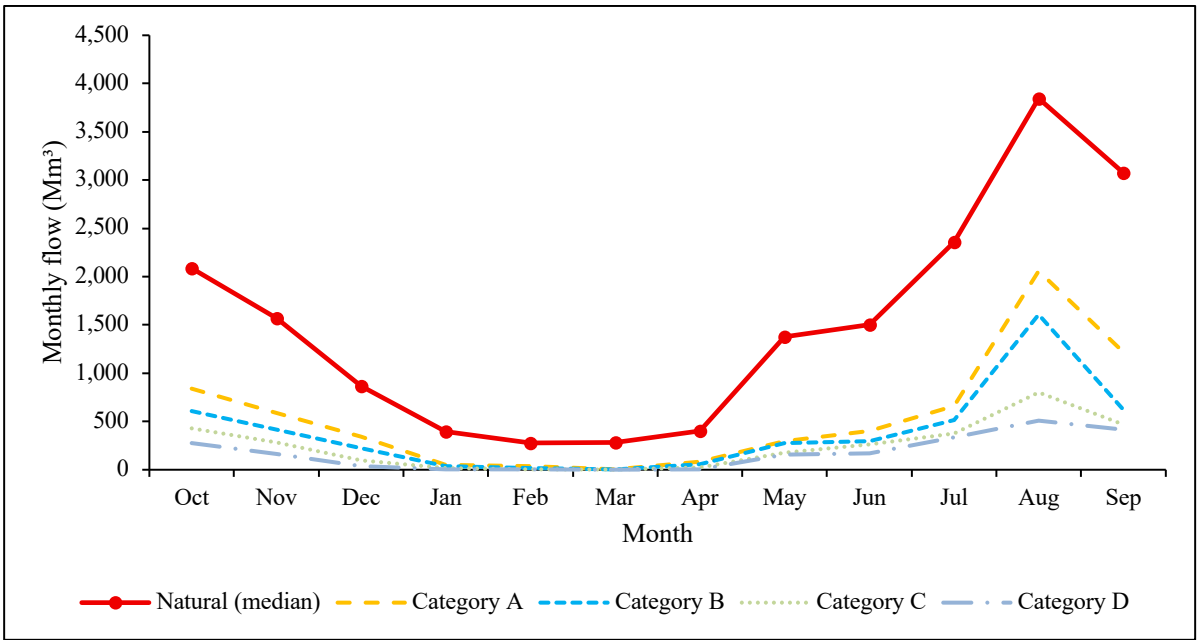


Figure 3.14. Median monthly natural flow and total-flow requirement for ecological categories A to D at the outlet of the lower sub-basin (D_5), by month of the hydrological year.

3.3 Stakeholder questionnaire

Fourteen of the 25 people who received the questionnaire responded (56%). The respondents came from four types of institution (Figure 3.15): five from universities or research institutions, four from government ministries or public authorities, four from consultancy or technical advisory firms, and one independent expert. No respondent came from a non-governmental or civil-society organization or from a basin or river authority.

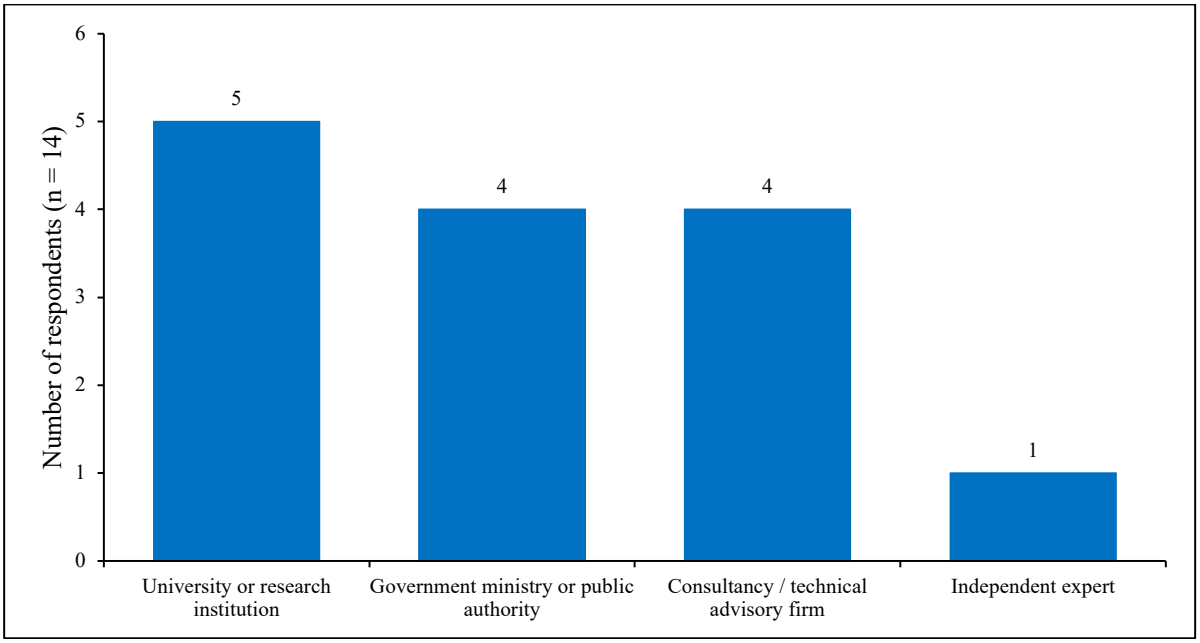


Figure 3.15. Institution type of the 14 questionnaire respondents.

Respondents were asked which types of data their organization collects or maintains for the basin (Figure 3.16). Streamflow records and water demand or socio-economic data were each reported by nine respondents, followed by rainfall and other meteorological data and water quality measurements (seven each), groundwater or baseflow data and abstraction or infrastructure-use records (six each), and ecological data (two). Government and public authorities reported streamflow, water quality, groundwater, abstraction, and demand data but no rainfall or ecological data. Universities and research institutions, and the single independent expert, were the only groups to report all seven types of data, and they accounted for four of the seven rainfall responses. Three respondents, one each from a university, a government body, and a consultancy, collected none of these types of data.

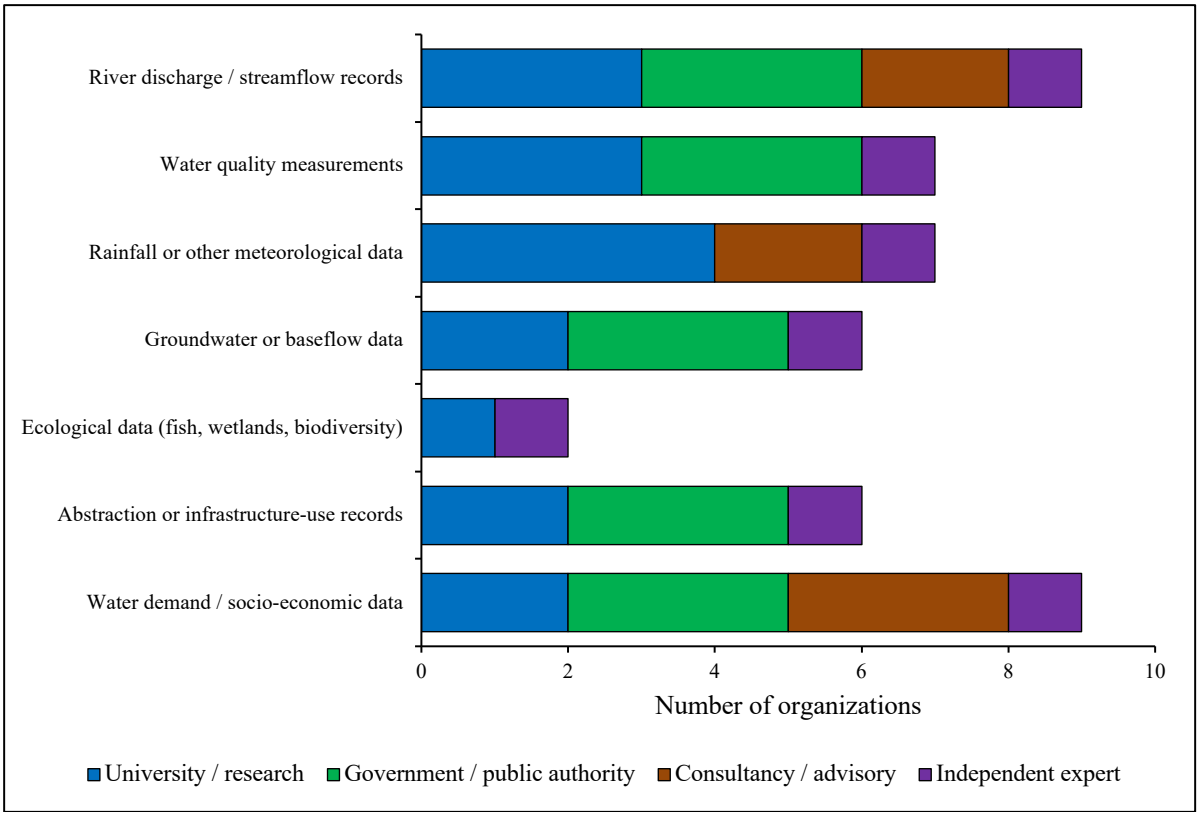


Figure 3.16. Types of data collected or maintained for the Omo-Gibe Basin, by institution type. Respondents could select more than one type.

Respondents rated gauge station coverage as moderate upstream of the Gibe cascade (10 of 14) but poor downstream toward South Omo and Lake Turkana (10 of 14) (Figure 3.17). Only two respondents rated coverage as good in either reach. Water quality monitoring coverage was rated lower overall (Figure 3.18). Upstream, five respondents rated it poor and six moderate, while downstream nine rated it poor and two moderate. One respondent rated water quality coverage as good in each reach, and two were not sure.

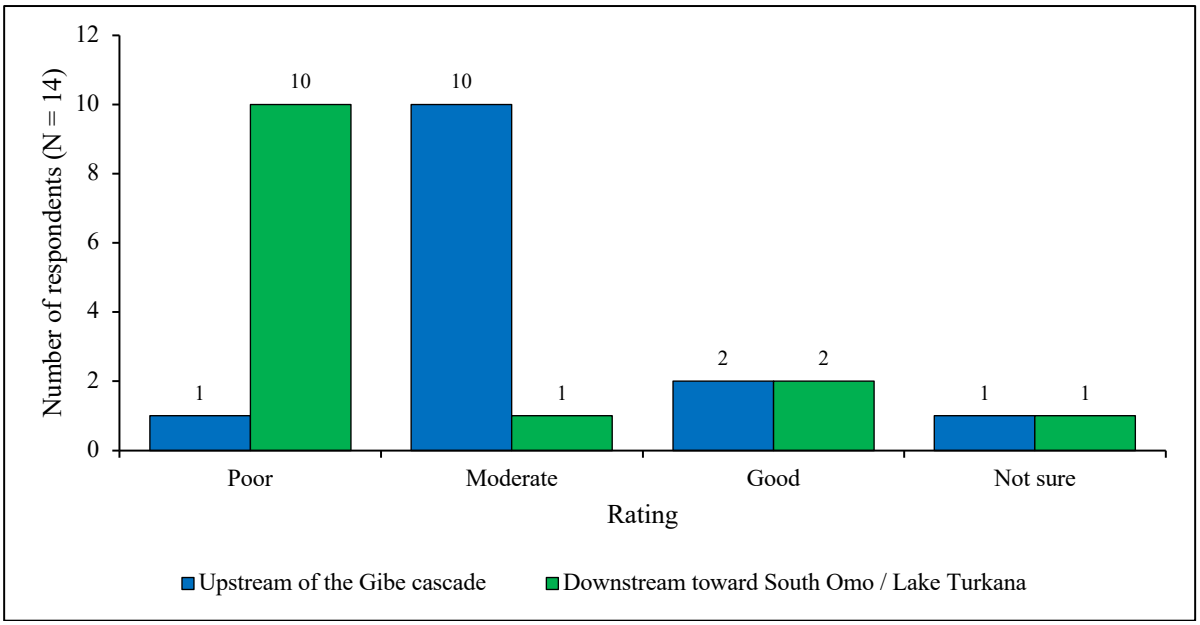


Figure 3.17. Ratings of gauge station coverage upstream and downstream of the Gibe cascade (n = 14).

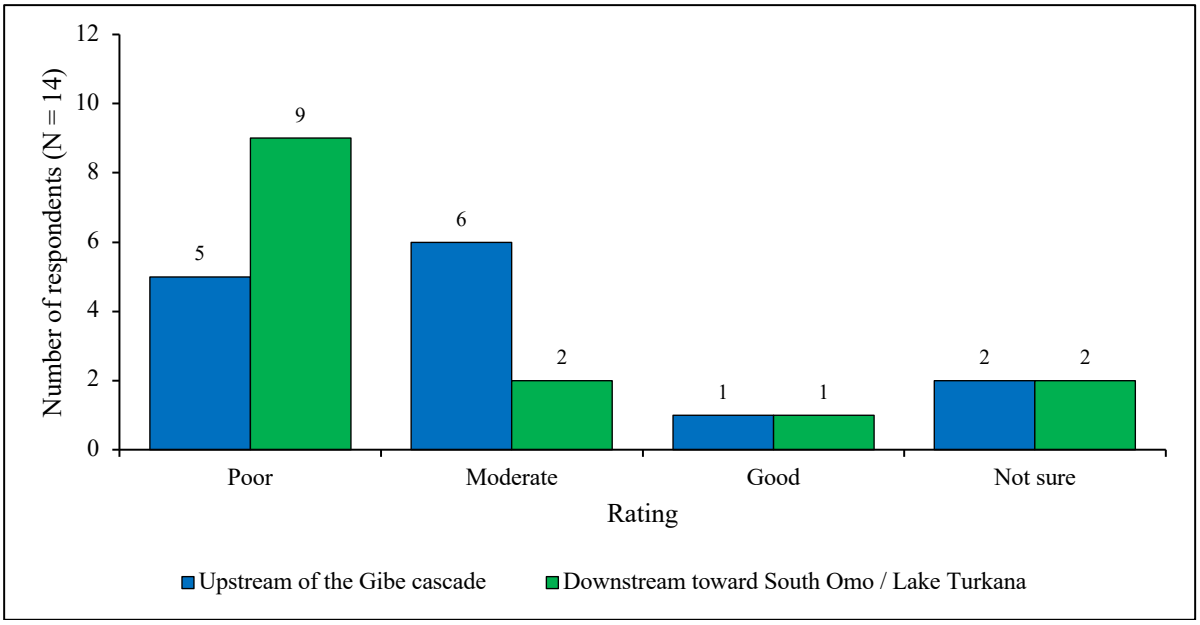


Figure 3.18. Ratings of water quality monitoring coverage upstream and downstream of the Gibe cascade (n = 14).

Respondents also described how easy it is to obtain data when they exist (Figure 3.19). For gauge data, seven respondents said the data were available through a formal request or data sharing agreement, four said the data technically exist but are very hard to obtain, two said they remain within one institution and are rarely shared, and one said they were available to universities on request by letter. For water quality data, five respondents named formal requests, three said the data are rarely shared, two said they are hard to obtain, one said they are available by payment or through publications, and three were not sure.

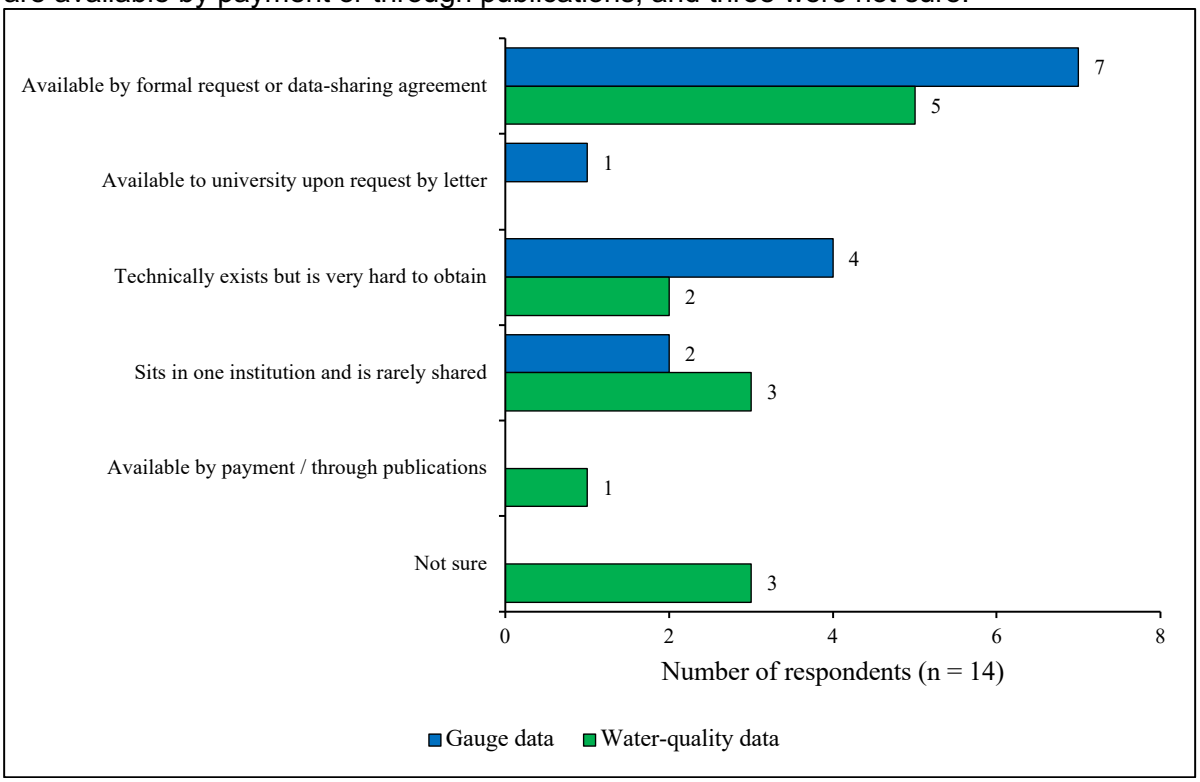


Figure 3.19. Ease of obtaining gauge data and water quality data when they exist (n = 14).

When asked where environmental-flow information should be used in decisions (Figure 3.20; up to three choices were allowed, but three respondents chose four, six, and six, and four

chose fewer than three), respondents most often selected protecting downstream livelihoods (12 of 14), followed by Gibe cascade dam operations and release schedules (10), irrigation abstraction permitting (8), and ecosystem and biodiversity conservation (7). Climate adaptation planning (3) and urban or industrial water planning (2) were selected less often, and one respondent wrote in that there was no water-resources limitation in the basin. This free-text answer is shown in Figure 3.20. The 43 responses in the figure comprise the 42 ticked options and this answer.

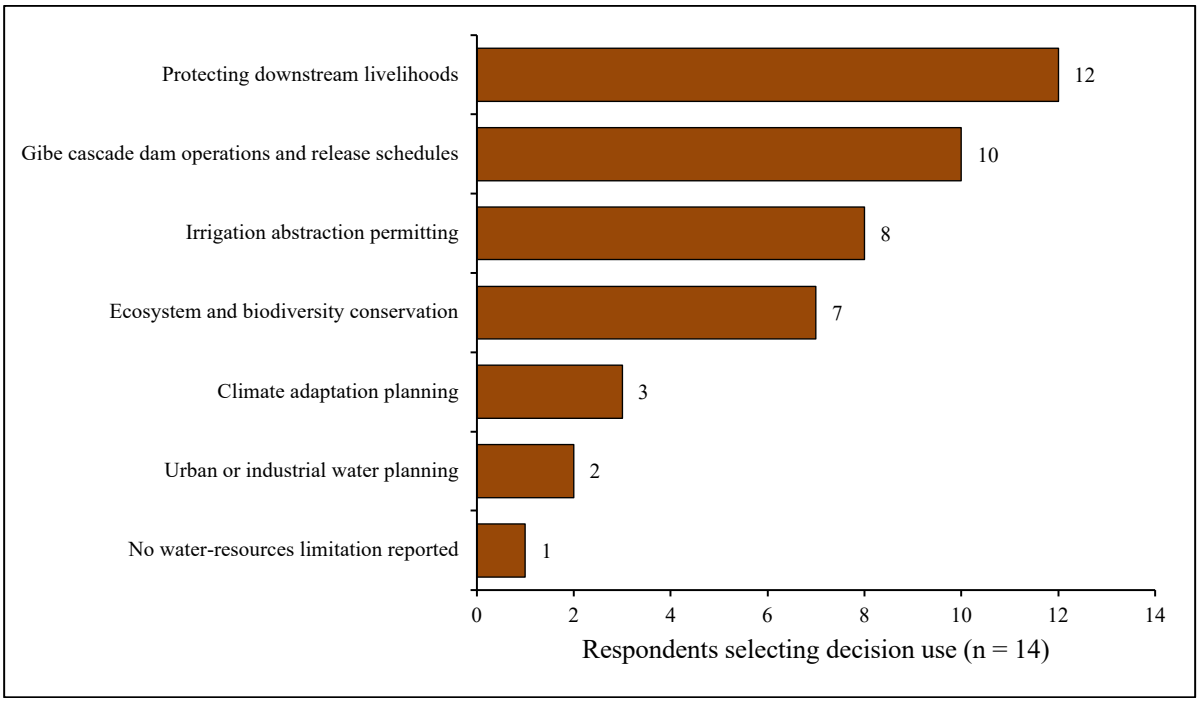


Figure 3.20. Decisions in which respondents thought environmental-flow data should be used (n = 14; up to three choices were allowed, but three respondents chose more).Respondents were asked to choose up to three barriers to implementing e-flows (Figure 3.21). The limit was not enforced by the survey platform: one respondent chose four barriers, one chose five, and one chose six, and all selections were counted. One respondent gave a free-text answer, a recommendation to visit the basin, instead of choosing a barrier, and this is not shown in the figure. The most frequently selected barriers were insufficient continuous gauge monitoring and poor coordination among institutions (nine respondents each), followed by weak regulation or enforcement (eight) and limited hydro-ecological modeling capacity (seven). Competing demands among energy, agriculture, and ecology, and insufficient funding, were each selected by five respondents.

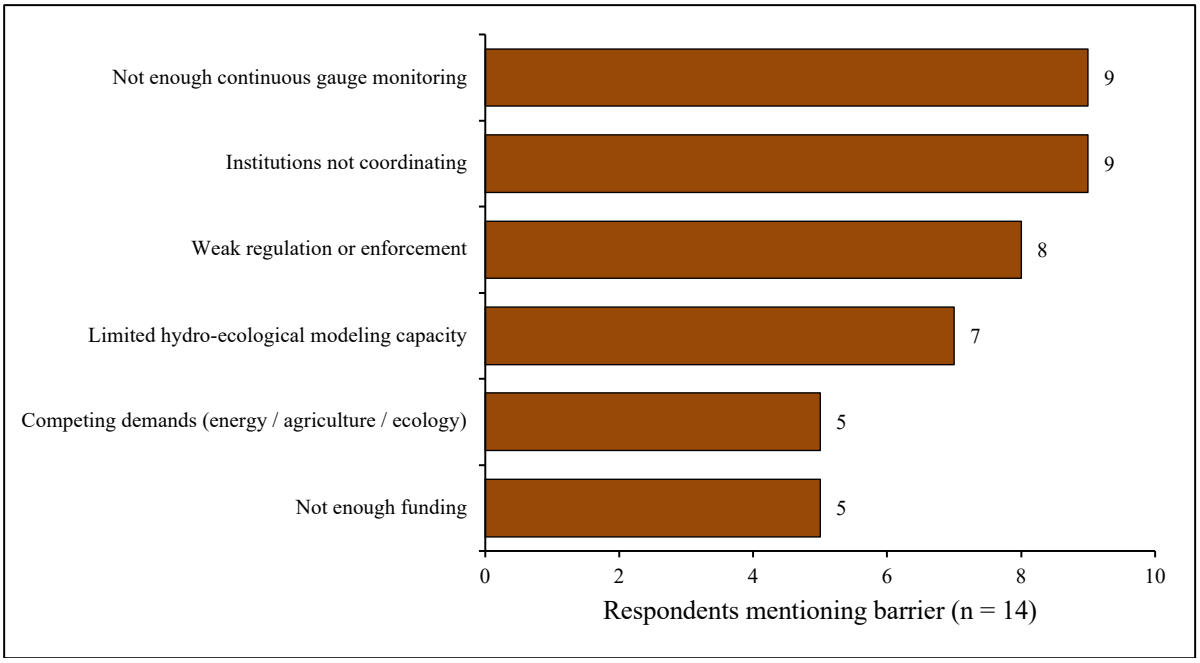


Figure 3.21. Barriers to implementing e-flows selected by respondents (n = 14; up to three choices were allowed, but three respondents chose more).

The open-ended questions gave the following picture. When asked which stretches most urgently need e-flows attention (question 11), nine of the 14 respondents named the lower Omo or the reach below the Gibe III dam toward Lake Turkana, one named the upper and middle basin, and two answered that no attention was needed because the basin has plenty of water or flows naturally. Two respondents gave no usable answer. Asked whom to contact about the lower catchment (question 10), four pointed to the national hydrology office of MoWE, one to a regional irrigation and lowlands office, two named individual contacts, and one said that there are no stations downstream and suggested earth observation data and the investors at Omorate. Six gave no specific lead. Asked what training would help (question 14), nine respondents named hydrological modeling, six named e-flow methods or indicator design, five named telemetry, management information systems, or gauge data management, two named reservoir operation or hydraulic modeling, and one named earth observation analysis. One respondent noted that modeling is not possible without continuous gauge data. In the final question (question 15), respondents asked for work on the operation of the dam cascade and on the effect of the dams on the water balance and the level of Lake Turkana (four), named who should attend the workshop, such as MoWE, basin authorities from Ethiopia and Kenya, universities, and downstream users (four), and asked for better data sharing and a shared open data repository (two). One respondent argued that the main problem is poor use of the available water and land and not a shortage of water. Summaries of all four open-ended questions are in Appendix C.3.

4. Discussion

This study produced a first desktop estimate of environmental flow requirements for the whole Omo-Gibe Basin by combining open-access satellite data, a machine learning model of naturalized flow, the RDRM, and the views of 14 water professionals. Three insights stand out. A simple tree-based model can reproduce monthly flow at gauged stations, but its transfer to ungauged catchments is untested. The total requirement is robust as a share of natural flow, whereas the low-flow requirement depends heavily on doubtful channel inputs. And the main obstacles to applying e-flows are missing continuous monitoring and weak coordination between institutions, not a lack of interest.

4.1 Naturalized flow model

The monthly model reached the very good class of Moriasi et al. (2007) at all three stations, which shows that gridded rainfall and evaporation can support useful monthly simulations in the Omo-Gibe. Large-sample studies such as Kratzert et al. (2019b) work at a daily step and test on unseen basins, so they are not directly comparable. The consistently higher monthly than daily R^2 is expected, because timing errors in a 5 km rainfall product largely cancel over a month. The daily model should therefore not be used to interpret single events, and the monthly step required by the RDRM is the appropriate one.

Excluding Ajancho did not improve the fit at the other stations; it only removed a poorly fitting station from the average. It also removed the only small, low-flow catchment from the training data, although small tributaries such as U_4 were simulated later. The final model is therefore best seen as valid for catchments resembling Megecha, Asendabo, and Abelti. Checking the Ajancho record and refitting with it would be preferable to excluding it.

The model has a systematic bias toward underestimating peaks and volume. Tree-based models average within leaves and rarely exceed the training maximum (Galelli and Castelletti 2013), which explains the underestimated peaks, the low annual coefficient of variation, and the missed 1998 El Niño peak at Asendabo, an event outside the training range (Avery and Tebbs 2018). Land-cover change after the training period (Worku et al. 2014; Anose et al. 2021) may add to the bias, because the model has no land-cover predictor. “Natural” flow here therefore means natural with respect to dams, not with respect to land cover or to abstractions, whose magnitude is unknown.

The key weakness is spatial transfer. All 15 sites are ungauged, yet performance was tested only in time, and the leave-one-station-out test rests on three or four northern stations. The model cannot extrapolate beyond the catchment areas seen in training, so flows at the large lower-basin sites assume that runoff depth does not change with catchment size, which we could not test. The GLDAS-2.1 comparison agreed least in the lower sub-basin, and both products share uncertain rainfall inputs, particularly over mountains (Dinku et al. 2018). The simulated flows are therefore more reliable for the upper and middle sub-basins, and lower-basin volumes are more likely under- than overestimated.

Independent estimates support the simulated flows. MAR at M_4 was 17% below the Gibe III inflow estimate (Velpuri and Senay 2012), consistent with the negative bias, and the D_5 value lies within published estimates of the Omo inflow to Lake Turkana (Avery and Tebbs 2018). The narrow range of the baseflow index suggests that the model distinguishes sites mainly by rainfall, evaporation, and area, not by geology. The model also has no routing or loss terms, so it cannot represent transmission losses, floodplain evaporation, or abstraction in the lower Omo, which may make simulated flows there too high in some months even as the annual bias points the other way.

4.2 Environmental water requirements

The total requirement as a share of MAR was similar across sub-basins despite different reference rivers, and it lies within the ranges reported globally (Pastor et al. 2014) for the Great Ruaha River (Kashaigili et al. 2006), and for the Blue Nile in Ethiopia, where desktop methods gave 21–28% of mean annual flow (McCartney et al. 2009; Reitberger and McCartney 2011), so the results are not outliers. Because the requirement scales almost linearly with MAR, its accuracy in volume depends mainly on the naturalized flow; correcting the -13% bias would raise volumes by about 15% but leave percentages largely unchanged. We therefore recommend passing results to managers first as percentages of natural flow.

The low-flow requirement is the least certain result, because it is very sensitive to the channel inputs. Maximum channel depth cannot be measured by remote sensing, so a single depth was applied to the whole sub-basin. A test in which the maximum depth was changed showed that the low-flow requirement changed substantially, whereas the total requirement changed much less. This is expected, because the total is dominated by the flood component, while the low-flow component depends on the hydraulics of the channel. The uniform depth is therefore the most likely cause of the very low dry-season values, notably at U_5, where the February category A requirement is only 2% of natural flow. Other inputs in Table B1 are also doubtful:: very large channel widths at U_4, M_2, M_3, and D_5 probably reflect floodplains or the Gibe III reservoir, and bankfull discharge is inconsistent between sites with the same catchment. These inputs likely explain the large differences between neighboring sites such as D_4 and D_5 and the high requirement at U_4. Site-level low-flow results should not be used for management until these inputs are checked against dated imagery or field data and the model is rerun.

The transferred South African parameters add further uncertainty. The desktop approach is designed for rapid, low-confidence assessment (Hughes and Hannart 2003), the reference rivers were matched only by channel type, and no sensitivity analysis or ecological data were available to test them. Two artifacts indicate where the model is unreliable: very low dry-season requirements at several sites, far below the 25 m³/s interim release for Gibe III (Agriconsulting S.p.A. and Mid-Day International Consulting Engineers 2009), and a sharp March fall at all lower sites that is absent from the input flow. Both should be disregarded until the runs are repeated with corrected inputs.

Compared with our results, the earlier Gibe III measures (Agriconsulting S.p.A. and Mid-Day International Consulting Engineers 2009) cover the low-flow end of the requirement but not the whole flow regime. The 25 m³/s minimum release is reasonable as a low flow but is only about a fifth of the category A total requirement, and a single 10-day flood cannot reproduce the sustained July to October high flows. A monthly step also cannot represent flood pulses or flood-recession agriculture, so flood requirements need a daily step. The observed 2015–2016 drop in Lake Turkana water levels during Gibe III filling, measured by satellite altimetry (Avery and Tebbs 2018) shows why such requirements matter.

4.3 Scope and missing components

The requirements are river volumes only. They are not a Lake Turkana level target, nor operating rules for the Gibe cascade, and they exclude water quality, sediment, groundwater, social needs, climate change (Woldemariam 2019), Koysha, and irrigation abstraction such as Kuraz. Each needs separate analysis or scenarios.

No Present Ecological State (PES) or Ecological Importance and Sensitivity (EIS) assessment was undertaken. Consequently, an ecological target category cannot be recommended, nor can current flows be compared with regulated flows because recent records are lacking (MoWE 2026). The degree of hydrological alteration, and thus required releases, therefore remain undetermined. . A logical next step would be a targeted study combining PES and EIS

assessments with more detailed hydrological analysis of naturalised and regulated flow regimes. This would provide an evidence base for defining ecological objectives, identifying the extent of hydrological alteration and determining appropriate environmental flow requirements.

4.4 Stakeholder perspectives

The questionnaire indicates that data and coordination limit progress. The poor downstream gauge coverage reported by respondents matches the MoWE network status (MoWE 2026) and creates a circular problem: the model is needed most where it is least testable. Data that exist are also hard to obtain. These barriers match the enabling factors identified by Arthington et al. (2024) and the slow, uneven implementation of e-flows worldwide (Tickner et al. 2020a). Respondents prioritized downstream livelihoods and dam operation, so e-flows are likely to gain support when framed in those terms, while the choice of a target category belongs to stakeholders and the water authority.

The respondent pool was small and biased toward technical specialists who had attended an e-flow workshop, with no civil-society or basin authority voices, so the results are a snapshot of expert opinion and cannot be generalized.

4.5 Limitations and next steps

Overall, the outputs are first-order estimates: no observed flow or ecological data were available at the sites, the climate and land cover are those of 1981 to 2005, and randomly chosen sites miss key decision points. The consistent message is that requirements are best expressed as a share of natural flow, that the flood component carries most of the requirement, and that the weakest evidence concerns low flows and the lower sub-basin.

Priorities for further work are to correct the channel inputs in Table B1 and rerun the RDRM with a sensitivity analysis; to recheck Ajancho and test the flow model with new middle and lower basin gauges; to compare observed regulated flows with naturalized flows; to collect field and ecological data for a Present Ecological State assessment and agree target categories with stakeholders; and to move to a daily step with scenarios for Lake Turkana, the cascade, Koysha, and irrigation. Until then, the results should be treated with caution and provide no more than an indicative guide for planning and the choice of more detailed methods (Tharme 2003; Poff et al. 2010).

4.6 Towards a digital twin of the Omo-Gibe Basin

The limitations above are as much about information infrastructure as about method. The data, models, and stakeholder knowledge needed for e-flow decisions in the Omo-Gibe Basin are scattered across institutions, and much of the information that matters most, such as flow records downstream of the dams, does not yet exist. A digital twin, a virtual representation of the basin that combines data, models, and decision tools and is updated with new observations, is one way to bring these elements together. IWMI has already built such a system for the Limpopo River Basin (Jolaiya et al. 2025), and its architecture provides a benchmark for what a first-generation Omo-Gibe twin would need.

Appendix D compares the outputs of this study with eight minimum requirements derived from the Limpopo twin. The spatial framework, the gridded forcing data, a reproducible flow model, and a first set of environmental water requirements already exist. Several improvements, including correcting the RDRM channel inputs, running the flow model to the present day, moving sites to decision points, and obtaining post-2005 MoWE records, need only modest effort. Other gaps need major investment: gauges and telemetry in the middle and lower basin, field-based ecological assessment, daily flood-pulse and reservoir modeling, the link to Lake

Turkana, and an institution with a mandate to own and maintain the system. This points to a phased pathway, from a planning twin built on existing components, to a monitored twin fed by rehabilitated gauges, and only then to an operational twin able to support decisions on the Gibe cascade (Appendix D).

5. Conclusions

This desktop assessment gives the first basin-wide, first-order estimates of environmental flow requirements for the Omo-Gibe Basin. The requirements are best expressed as a share of natural flow rather than as fixed volumes, and most of the total requirement comes from the high-flow (flood) component, which existing release rules do not cover. Confidence is lowest for low flows and for the lower basin, where the model was transferred beyond the gauged northern catchments and the channel inputs are uncertain. The main barriers are a lack of monitoring and weak coordination between agencies, more than a lack of methods. Before the results are used to set operating rules, the channel inputs should be corrected, gauging in the middle and lower basin restored and extended, ecological condition (PES and EIS) assessed, and target management categories agreed with stakeholders. A daily-step analysis is then needed to set flood-pulse releases below the Gibe III and Koysha dams. Finally, the naturalized flow series, e-flow estimates, and stakeholder insights assembled here could form the first components of a digital twin of the basin.

References

Agriconsulting S.p.A., and Mid-Day International Consulting Engineers. 2009. *Gibe III EIA: Additional Study on Downstream Impacts*. Report 300 ENV R AG 003 B. Ethiopian Electric Power Corporation. https://www.afdb.org/en/documents/document/ethiopia-gibe-iii-hydroelectric-project-esia-additional-study-17268.

Anose, Fikru Abiko, Kassahun Ture Beketie, Tadesse Terefe Zeleke, Desalegn Yayeh Ayal, and Gudina Legese Feyisa. 2021. "Spatio-Temporal Hydro-Climate Variability in Omo-Gibe River Basin, Ethiopia." *Climate Services* 24: 100277. https://doi.org/10.1016/j.cliser.2021.100277.

Arthington, Angela H., Anik Bhaduri, Stuart E. Bunn, et al. 2018. "The Brisbane Declaration and Global Action Agenda on Environmental Flows (2018)." *Frontiers in Environmental Science* 6: 45. https://doi.org/10.3389/fenvs.2018.00045.

Arthington, Angela H., David Tickner, Michael E. McClain, et al. 2024. "Accelerating Environmental Flow Implementation to Bend the Curve of Global Freshwater Biodiversity Loss." *Environmental Reviews* 32 (3): 387–413. https://doi.org/10.1139/er-2022-0126.

Avery, Sean T., and Emma J. Tebbs. 2018. "Lake Turkana, Major Omo River Developments, Associated Hydrological Cycle Change and Consequent Lake Physical and Ecological Change." *Journal of Great Lakes Research* 44 (6): 1164–82. https://doi.org/10.1016/j.jglr.2018.08.014.

Breiman, Leo. 2001. "Random Forests." *Machine Learning* 45 (1): 5–32. https://doi.org/10.1023/A:1010933404324.

Carr, Claire J. 2017. *River Basin Development and Human Rights in Eastern Africa: A Policy Crossroads*. Springer. https://doi.org/10.1007/978-3-319-50469-8.

Chen, Tianqi, and Carlos Guestrin. 2016. "XGBoost: A Scalable Tree Boosting System." In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge*

Discovery and Data Mining. Association for Computing Machinery. https://doi.org/10.1145/2939672.2939785.

Chisanga, Charles B., Kabwe H. Mubanga, Henry Sichigabula, et al. 2022. "Modelling Climatic Trends for the Zambezi and Orange River Basins: Implications on Water Security." *Journal of Water and Climate Change* 13 (3): 1275–96. https://doi.org/10.2166/wcc.2022.308.

Dinku, Tufa, Chris Funk, Pete Peterson, et al. 2018. "Validation of the CHIRPS Satellite Rainfall Estimates over Eastern Africa." *Quarterly Journal of the Royal Meteorological Society* 144 (S1): 292–312. https://doi.org/10.1002/qj.3244.

Dormann, Carsten F., Jane Elith, Sven Bacher, et al. 2013. "Collinearity: A Review of Methods to Deal with It and a Simulation Study Evaluating Their Performance." *Ecography* 36 (1): 27–46. https://doi.org/10.1111/j.1600-0587.2012.07348.x.

Duan, Naihua. 1983. "Smearing Estimate: A Nonparametric Retransformation Method." *Journal of the American Statistical Association* 78 (383): 605–10. https://doi.org/10.1080/01621459.1983.10478017.

Du Plessis, B. J. 2023. "The Development, Implementation and Evaluation of an Integrated Aquatic Vegetation Management Plan for the Control of Submersed Aquatic Vegetation in South African Impoundments." PhD diss., North-West University.

Friedman, Jerome H. 2001. "Greedy Function Approximation: A Gradient Boosting Machine." *Annals of Statistics* 29 (5): 1189–232. https://doi.org/10.1214/aos/1013203451.

Friedman, Jerome H. 2002. "Stochastic Gradient Boosting." *Computational Statistics & Data Analysis* 38 (4): 367–78. https://doi.org/10.1016/S0167-9473(01)00065-2.

Funk, Chris, Pete Peterson, Martin Landsfeld, et al. 2015. "The Climate Hazards Infrared Precipitation with Stations—A New Environmental Record for Monitoring Extremes." *Scientific Data* 2: 150066. https://doi.org/10.1038/sdata.2015.66.

Farr, Tom G., Paul A. Rosen, Edward Caro, et al. 2007. "The Shuttle Radar Topography Mission." *Reviews of Geophysics* 45 (2): RG2004. https://doi.org/10.1029/2005RG000183.

Galelli, Stefano, and Andrea Castelletti. 2013. "Assessing the Predictive Capability of Randomized Tree-Based Ensembles in Streamflow Modelling." *Hydrology and Earth System Sciences* 17 (7): 2669–84. https://doi.org/10.5194/hess-17-2669-2013.

Hughes, Denis A., Anupama Y. Desai, Anthony L. Birkhead, and Delana Louw. 2014. "A New Approach to Rapid, Desktop-Level, Environmental Flow Assessments for Rivers in South Africa." *Hydrological Sciences Journal* 59 (3–4): 673–87. https://doi.org/10.1080/02626667.2013.818220.

Hughes, Denis A., and D. A. Forsyth. 2006. "A Generic Database and Spatial Interface for the Application of Hydrological and Water Resource Models." *Computers & Geosciences* 32 (9): 1389–402. https://doi.org/10.1016/j.cageo.2005.12.013.

Hughes, Denis A., and Peter Hannart. 2003. "A Desktop Model Used to Provide an Initial Estimate of the Ecological Instream Flow Requirements of Rivers in South Africa." *Journal of Hydrology* 270 (3–4): 167–81. https://doi.org/10.1016/S0022-1694(02)00290-1.

Hughes, Denis A., and Delana Louw. 2010. “Integrating Hydrology, Hydraulics and Ecological Response into a Flexible Approach to the Determination of Environmental Water Requirements for Rivers.” *Environmental Modelling & Software* 25 (8): 910–18. https://doi.org/10.1016/j.envsoft.2010.03.004.

Hughes, Denis A., Delana Louw, Anupama Y. Desai, and Anthony L. Birkhead. 2012. *Development of a Revised Desktop Model for the Determination of the Ecological Reserve for Rivers*. WRC Report 1856/1/11. Water Research Commission.

International River Foundation. 2007. “The Brisbane Declaration.” In *Environmental Flows Are Essential for Freshwater Ecosystem Health and Human Well-Being*. Declaration of the 10th International River Symposium and International Environmental Flows Conference, Brisbane, Australia.

Jolaiya, E.A., H. Retief, P. Silva, A. Afham, K. Vickneswaran, and M. Garcia Andarcia. 2025. *Limpopo Digital Twin Architecture Report*. Colombo, Sri Lanka: International Water Management Institute (IWMI). CGIAR Accelerator for Digital Transformation.

Kashaigili, Japhet J., Matthew McCartney, and Henry F. Mahoo. 2006. *Estimation of Environmental Flows in the Great Ruaha River Catchment, Tanzania: Use of the Desktop Reserve Model*. WaterNet/WARFSA/GWP-SA.

Kleynhans, Christa J., and Marisa D. Louw. 2007. “Module A: EcoClassification and EcoStatus Determination.” In *River EcoClassification: Manual for EcoStatus Determination (Version 2)*. WRC Report TT 329/08. Water Research Commission.

Kratzert, Frederik, Daniel Klotz, Mathew Herrnegger, Alden K. Sampson, Sepp Hochreiter, and Grey Nearing. 2019a. “Toward Improved Predictions in Ungauged Basins: Exploiting the Power of Machine Learning.” *Water Resources Research* 55 (12): 11344–54. https://doi.org/10.1029/2019WR026065.

Kratzert, Frederik, Daniel Klotz, Guy Shalev, Günter Klambauer, Sepp Hochreiter, and Grey Nearing. 2019b. “Towards Learning Universal, Regional, and Local Hydrological Behaviors via Machine Learning Applied to Large-Sample Datasets.” *Hydrology and Earth System Sciences* 23 (12): 5089–110. https://doi.org/10.5194/hess-23-5089-2019.

Lehner, Bernhard, and Günther Grill. 2013. “Global River Hydrography and Network Routing: Baseline Data and New Approaches to Study the World’s Large River Systems.” *Hydrological Processes* 27 (15): 2171–86. https://doi.org/10.1002/hyp.9740.

McCartney, Matthew P., Abeyu Shiferaw, and Yilma Seleshi. 2009. “Estimating Environmental Flow Requirements Downstream of the Chara Chara Weir on the Blue Nile River.” *Hydrological Processes* 23 (26): 3751–58. https://doi.org/10.1002/hyp.7254.

Moriasi, Daniel N., Jeffrey G. Arnold, Michael W. Van Liew, Ronald L. Bingner, R. Daren Harmel, and Tamie L. Veith. 2007. “Model Evaluation Guidelines for Systematic Quantification of Accuracy in Watershed Simulations.” *Transactions of the ASABE* 50 (3): 885–900. https://doi.org/10.13031/2013.23153.

MoWE (Ministry of Water and Energy, Ethiopia). 2026. “Hydrological Station Network of the Omo-Gibe Basin and Basin Monitoring Strategy 2026–2035.” Presentation by Surafel Dawit, Senior Hydromet Coordinator. Ministry of Water and Energy.

MoWE (Ministry of Water and Energy, Ethiopia). n.d. Daily discharge records of the Megecha, Ajancho, Gilgel Gibe, and Gibe stations, 1981–2005. Unpublished data. Ministry of Water and Energy.

Muñoz-Sabater, Joaquín, Emanuel Dutra, Anna Agustí-Panareda, et al. 2021. "ERA5-Land: A State-of-the-Art Global Reanalysis Dataset for Land Applications." *Earth System Science Data* 13 (9): 4349–83. https://doi.org/10.5194/essd-13-4349-2021.

Pastor, Amandine V., Fulco Ludwig, Hester Biemans, Holger Hoff, and Pavel Kabat. 2014. "Accounting for Environmental Flow Requirements in Global Water Assessments." *Hydrology and Earth System Sciences* 18 (12): 5041–59. https://doi.org/10.5194/hess-18-5041-2014.

Poff, N. LeRoy, J. David Allan, Mark B. Bain, et al. 1997. "The Natural Flow Regime." *BioScience* 47 (11): 769–84. https://doi.org/10.2307/1313099.

Poff, N. LeRoy, Brian D. Richter, Angela H. Arthington, et al. 2010. "The Ecological Limits of Hydrologic Alteration (ELOHA): A New Framework for Developing Regional Environmental Flow Standards." *Freshwater Biology* 55 (1): 147–70. https://doi.org/10.1111/j.1365-2427.2009.02204.x.

Reitberger, Bianca, and Matthew McCartney. 2011. "Concepts of Environmental Flow Assessment and Challenges in the Blue Nile Basin, Ethiopia." In *Nile River Basin: Hydrology, Climate and Water Use*, edited by Assefa M. Melesse, 337–58. Springer. https://doi.org/10.1007/978-94-007-0689-7_17.

Roberts, David R., Vivienne Bahn, Simone Ciuti, et al. 2017. "Cross-Validation Strategies for Data with Temporal, Spatial, Hierarchical, or Phylogenetic Structure." *Ecography* 40 (8): 913–29. https://doi.org/10.1111/ecog.02881.

Rowntree, Kate M., Rob A. Wadeson, and Jay O'Keeffe. 2000. "The Development of a Geomorphological Classification System for the Longitudinal Zonation of South African Rivers." *South African Geographical Journal* 82 (3): 163–72. https://doi.org/10.1080/03736245.2000.9713710.

Simaika, John P., and Alemseged Tamiru Haile. 2026a. *Workshop Report: Towards Sustaining Healthy Flows and Functions in the Omo-Gibe Basin*. International Water Management Institute (IWMI). https://hdl.handle.net/10568/185290.

Simaika, John P., and Alemseged Tamiru Haile. 2026b. *Towards Sustaining Healthy Flows and Functions in the Omo-Gibe Basin: Second Stakeholder Workshop*. International Water Management Institute (IWMI). https://hdl.handle.net/10568/185913.

Tharme, Rebecca E. 2003. "A Global Perspective on Environmental Flow Assessment: Emerging Trends in the Development and Application of Environmental Flow Methodologies for Rivers." *River Research and Applications* 19 (5–6): 397–441. https://doi.org/10.1002/rra.736.

Van der Waal, Benjamin Wentsel. 2014. "Sediment Connectivity in the Upper Thina Catchment, Eastern Cape, South Africa." PhD diss., Rhodes University. https://hdl.handle.net/10962/d1013589.

Tickner, David, Nitin Kaushal, Robert Speed, and Rebecca Tharme. 2020a. "Editorial: Implementing Environmental Flows; Lessons for Policy and Practice." *Frontiers in Environmental Science* 8: 106. https://doi.org/10.3389/fenvs.2020.00106.

Tickner, David, Jeffrey J. Opperman, Robin Abell, et al. 2020b. “Bending the Curve of Global Freshwater Biodiversity Loss: An Emergency Recovery Plan.” *BioScience* 70 (4): 330–42. https://doi.org/10.1093/biosci/biaa002.

Tickner, David, Laurenne Schiller, Steven J. Cooke, et al. 2025. “Addressing the Policy and Business Drivers of Global Freshwater Biodiversity Loss.” *Environmental Reviews* 33: 1–20. https://doi.org/10.1139/er-2024-0139.

UNEP (United Nations Environment Programme). 2023. “Largest River and Wetland Restoration Initiative in History Launched at UN Water Conference.” Press release, March 23. https://www.unep.org/news-and-stories/press-release/largest-river-and-wetland-restoration-initiative-history-launched-un.

Velpuri, Naga Manohar, and Gabriel B. Senay. 2012. “Assessing the Potential Hydrological Impact of the Gibe III Dam on Lake Turkana Water Level Using Multi-Source Satellite Data.” *Hydrology and Earth System Sciences* 16 (10): 3561–78. https://doi.org/10.5194/hess-16-3561-2012.

Woldemariam, Tesfaye Tadesse. 2019. “Impact Assessment of Climate Change and Anthropogenic Actions on Streamflow Characteristics, Hydropower and Environmental Flow of Omo-Gibe River Basin in Ethiopia.” PhD diss., Indian Institute of Technology Delhi.

Worku, Fekadu F., Michel Werner, Nick Wright, Pieter van der Zaag, and Sileshi S. Demissie. 2014. “Flow Regime Change in an Endorheic Basin in Southern Ethiopia.” *Hydrology and Earth System Sciences* 18 (9): 3837–53. https://doi.org/10.5194/hess-18-3837-2014.

Appendix A. Naturalized streamflow model

This appendix gives supporting information for the naturalized streamflow model: the candidate predictors, the performance at the gauged stations, and a comparison of the simulated flows at the assessment sites with GLDAS-2.1 runoff.

A.1 Candidate predictors

Table A1 lists the 16 candidate predictors and shows which ones were kept after the collinearity check. Figure A1 shows the Spearman correlation matrix used for that check.

Table A1. Candidate predictors of the naturalized flow model and whether they were retained.

Group	Variable name	Description	Retained
Physical	Log catchment area	Natural logarithm of upstream catchment area	Yes
Seasonality	Day of year (sine)	Sine of day of year	Yes
Seasonality	Day of year (cosine)	Cosine of day of year	No
Solar forcing	Solar radiation	Daily mean surface solar radiation	No
Solar forcing	Shortwave radiation flux	Downward shortwave radiation flux	No
Evapotranspiration	Potential evapotranspiration	Daily potential evapotranspiration	No
Evapotranspiration	PET, 7-day rolling mean	7-day rolling mean of potential evapotranspiration	Yes
Evapotranspiration	PET, lagged	Lagged potential evapotranspiration	No
Soil moisture	Soil moisture	Volumetric soil moisture (0-10 cm)	No
Soil moisture	Soil moisture, lagged	Lagged soil moisture	No
Soil moisture	Soil moisture, anomaly	Departure of soil moisture from the long-term mean	No
Precipitation	Precipitation	Daily precipitation	No
Precipitation	Precipitation, 3-day lag	3-day lagged daily precipitation	Yes
Precipitation	Precipitation, 60-day sum	60-day cumulative precipitation	Yes
Precipitation	Precipitation, 7-day sum	7-day cumulative precipitation	No
Precipitation	Precipitation, anomaly	Departure of precipitation from the long-term mean	No

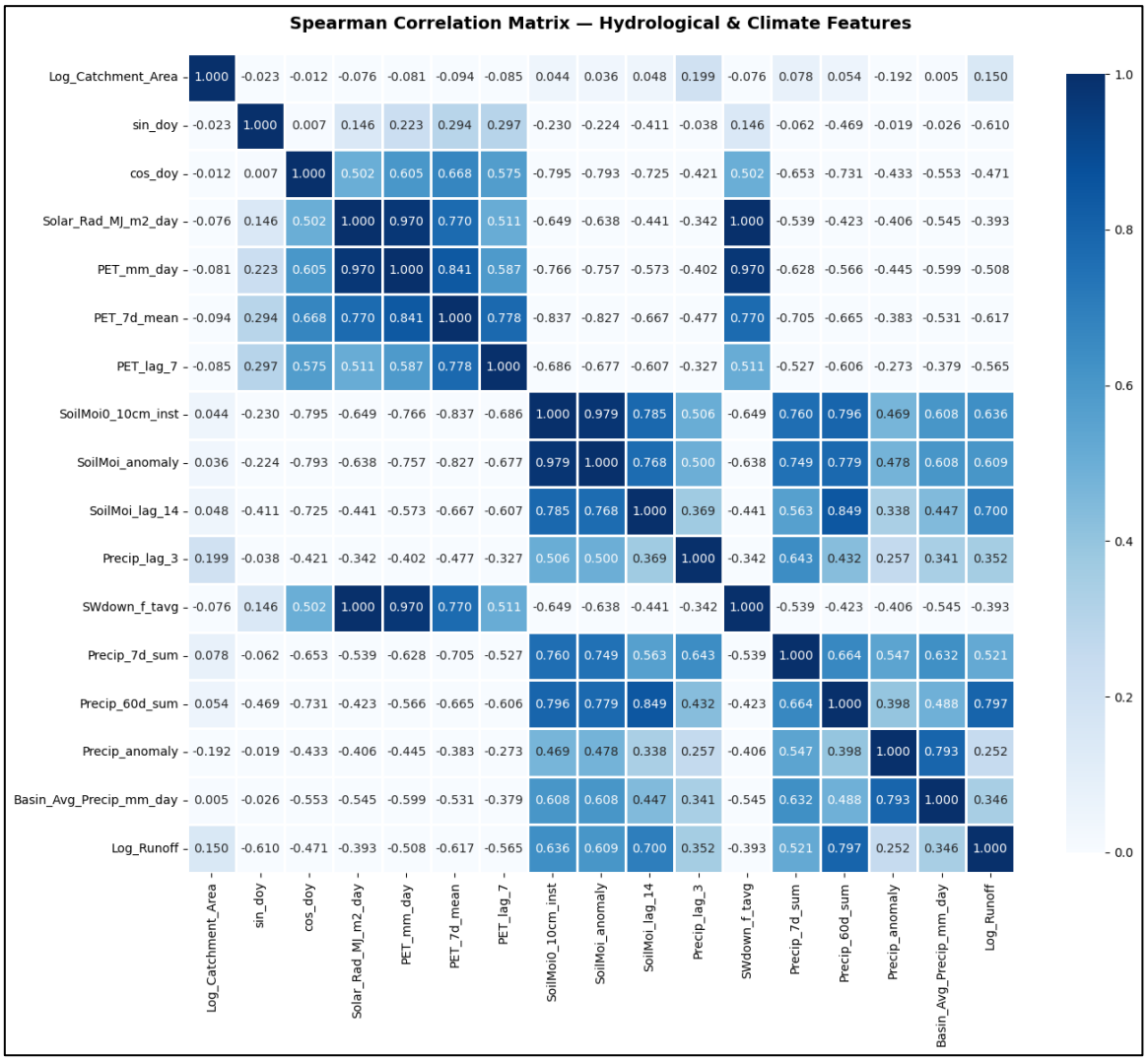


Figure A1. Spearman rank correlation matrix of the 16 candidate predictors and log-transformed runoff (Log_Runoff).

A.2 Performance at the gauged stations

Table A2 gives the coefficient of determination (R^2) at each gauging station for the four model configurations, and the test-period Nash-Sutcliffe efficiency (NSE) of the selected model. The selected model is the monthly model without the Ajancho station.

Table A2. Coefficient of determination (R^2) at the gauging stations for the models with and without the Ajancho station, and test-period NSE of the selected model (monthly, without Ajancho).

Station	With Ajancho, daily R^2	With Ajancho, monthly R^2	Without Ajancho, daily R^2	Without Ajancho, monthly R^2	Selected model, NSE
Megecha	0.53	0.84	0.47	0.79	0.78
Asendabo	0.65	0.79	0.65	0.84	0.82
Ajancho	0.01	0.13	–	–	–
Abelti	0.79	0.88	0.79	0.90	0.89
Mean	**0.66**	**0.84**	**0.64**	**0.84**	**0.83**

A.3 Comparison with GLDAS-2.1

The assessment sites have no observed flow records, so the simulated monthly flows for 2000 to 2005 were compared with runoff from the GLDAS-2.1 Noah land surface. Specifically, surface runoff (Qs_acc) and subsurface baseflow (Qsb_acc) accumulation variables from GLDAS were summed, spatially averaged over the catchment polygon, and converted from water depth to volumetric discharge (m^3/s) using the total upstream catchment area. This is a consistency check against another model and not a validation against observations. It covers only 2000 to 2005, and both products depend on gridded estimates of rainfall and temperature, so they are not independent of each other's input errors. The comparison was made at 15 sites (Table A3; Figures A2 to A16).

The agreement was good at all sites. NSE ranged from 0.774 (D_4) to 0.849 (M_3 and M_4), with a mean of 0.82, and R^2 ranged from 0.800 to 0.867. The two sources show the same seasonal pattern and the same differences between years, including the weaker wet season of 2002. Most of the differences are in the peaks. At U_1, U_3, and U_5, the simulated peaks were lower than the GLDAS peaks in most years, by up to about a third, and they were also somewhat lower at M_1, M_3, M_4, and M_5. At the lower sites the peaks were of similar size, but the simulated flows tended to peak a little later than GLDAS and to recede more slowly, and the 2002 peak was higher in the simulated flow than in GLDAS. GLDAS gives the runoff generated within each grid cell and does not route it through the river network, which may explain part of the difference in timing.

Table A3. Catchment area, NSE, and R^2 of the simulated monthly flow compared with GLDAS-2.1 runoff, 2000-2005. The areas are those printed on the comparison figures.

Site	Catchment area (km^2)	NSE	R^2
U_1	6,763	0.778	0.867
U_2	5,172	0.838	0.859
U_3	21,429	0.821	0.854
U_4	3,123	0.846	0.858
U_5	21,429	0.821	0.854
M_1	31,193	0.842	0.863
M_2	33,302	0.846	0.859
M_3	34,473	0.849	0.860
M_4	34,473	0.849	0.860
M_5	45,529	0.844	0.849
D_1	49,440	0.835	0.843
D_2	56,788	0.788	0.810
D_3	60,852	0.783	0.804
D_4	69,096	0.774	0.800
D_5	70,567	0.779	0.800

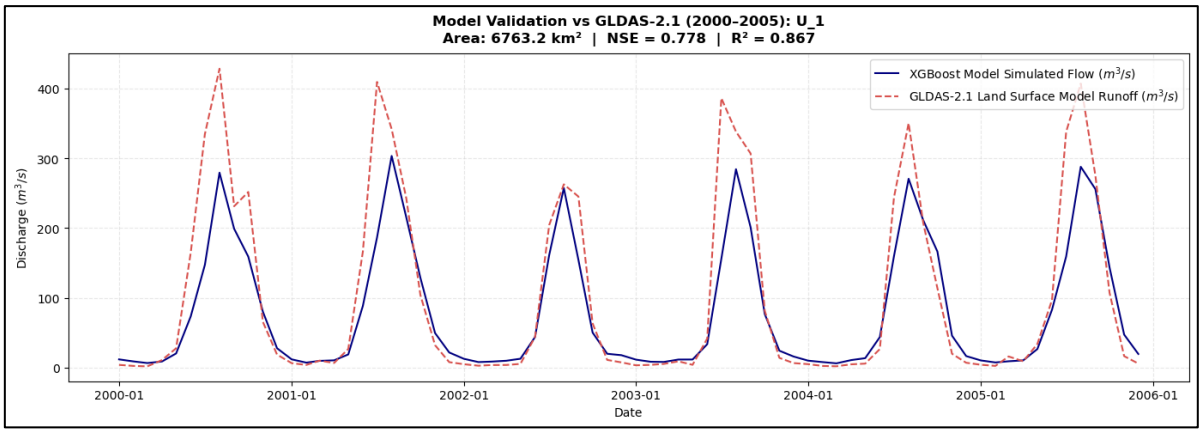


Figure A2. Simulated monthly flow at U_1 compared with GLDAS-2.1 runoff, 2000-2005.

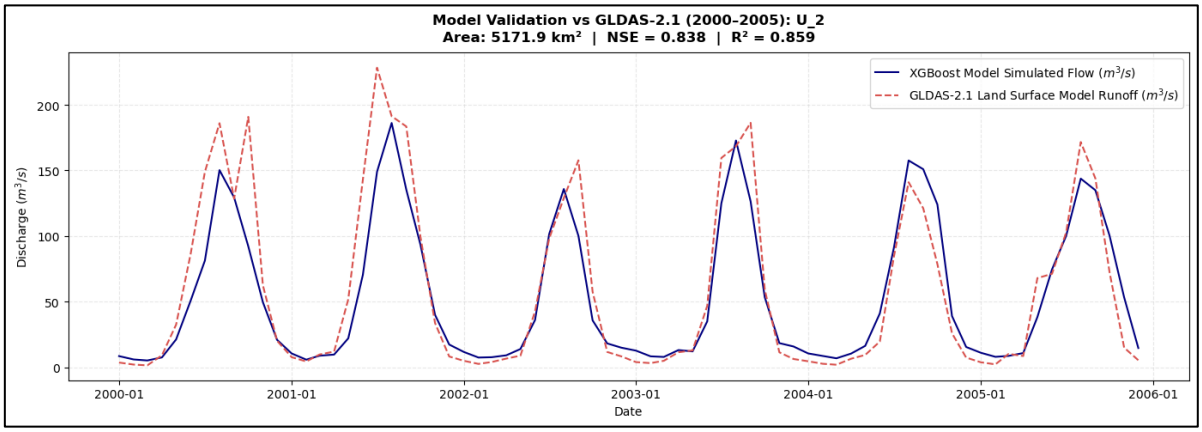


Figure A3. Simulated monthly flow at U_2 compared with GLDAS-2.1 runoff, 2000-2005.

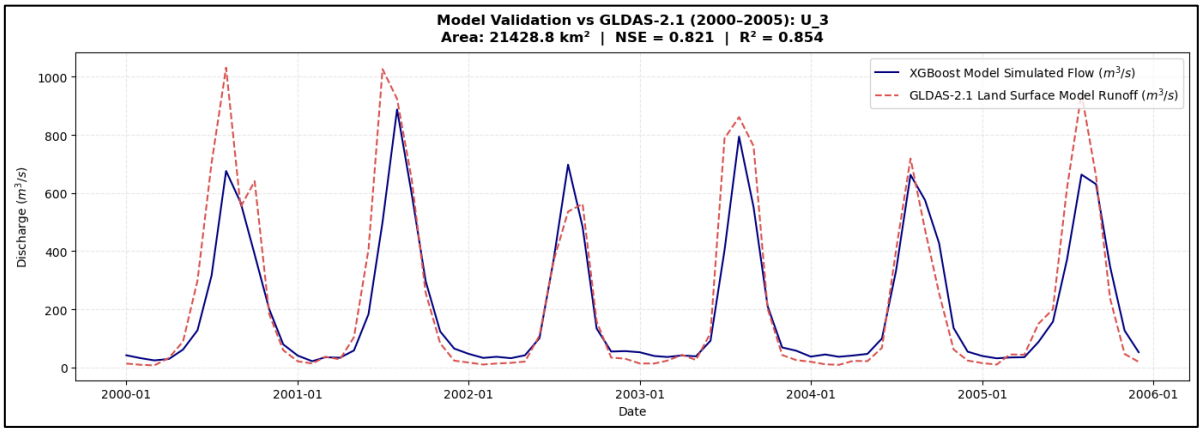


Figure A4. Simulated monthly flow at U_3 compared with GLDAS-2.1 runoff, 2000-2005.

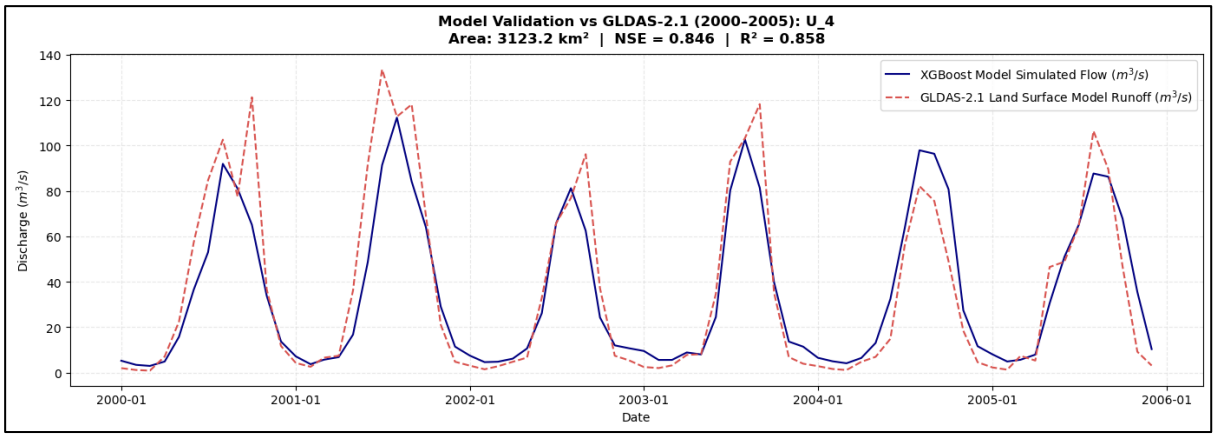


Figure A5. Simulated monthly flow at U_4 compared with GLDAS-2.1 runoff, 2000-2005.

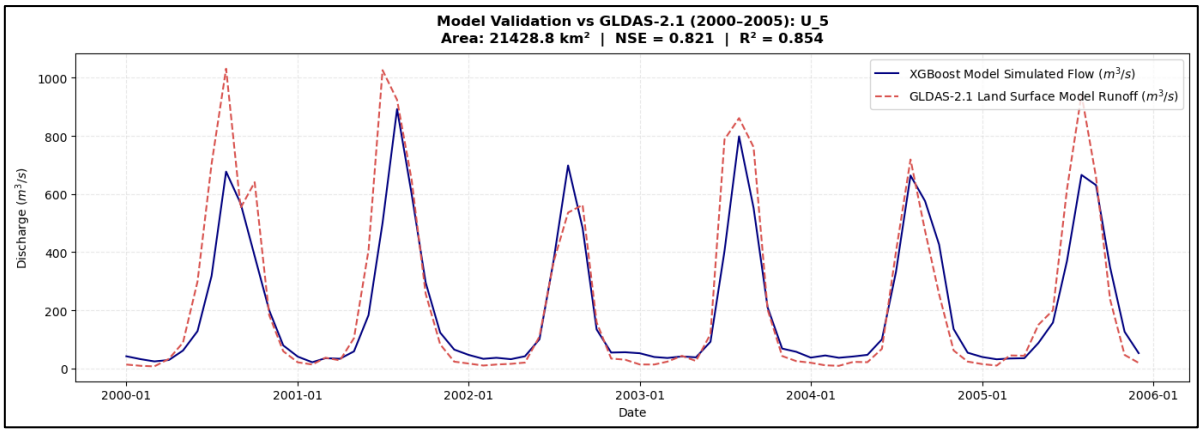


Figure A6. Simulated monthly flow at U_5 compared with GLDAS-2.1 runoff, 2000-2005.

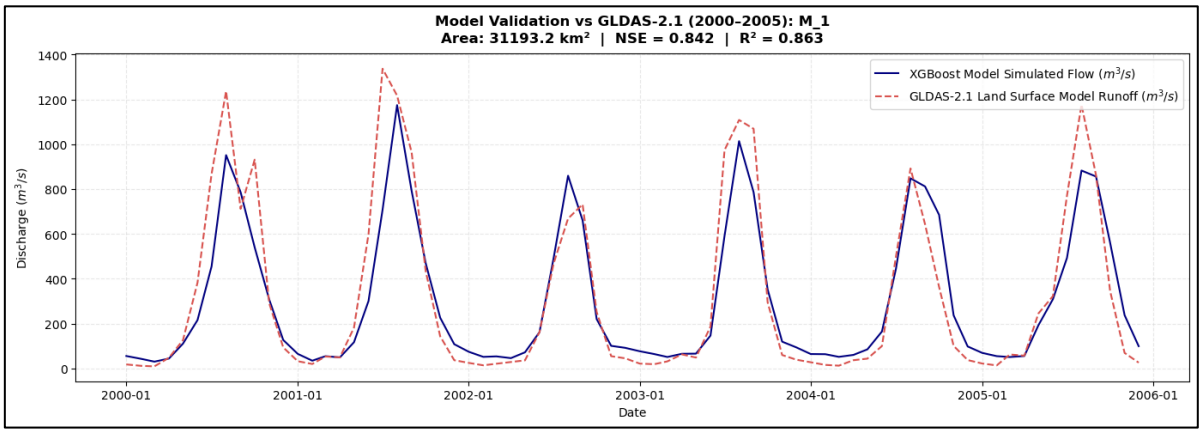


Figure A7. Simulated monthly flow at M_1 compared with GLDAS-2.1 runoff, 2000-2005.

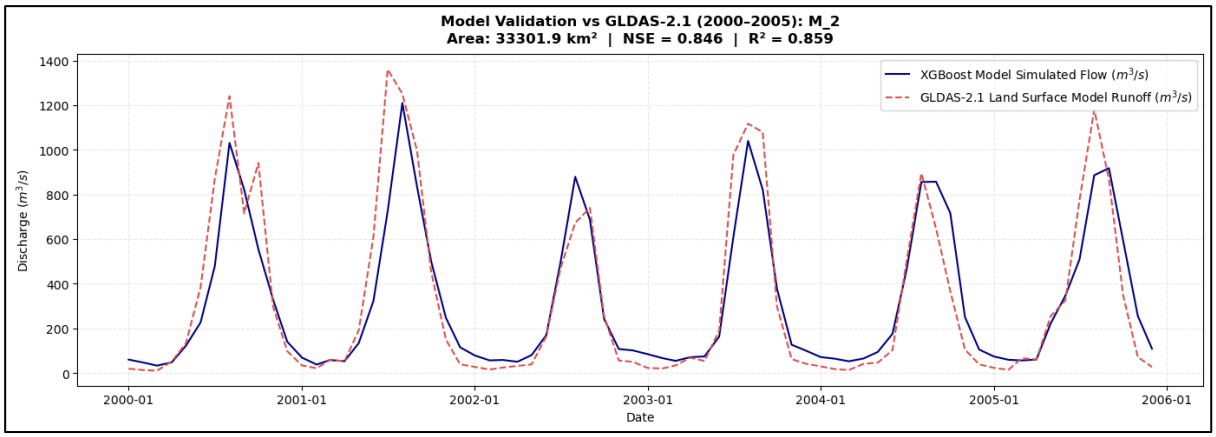


Figure A8. Simulated monthly flow at M_2 compared with GLDAS-2.1 runoff, 2000-2005.

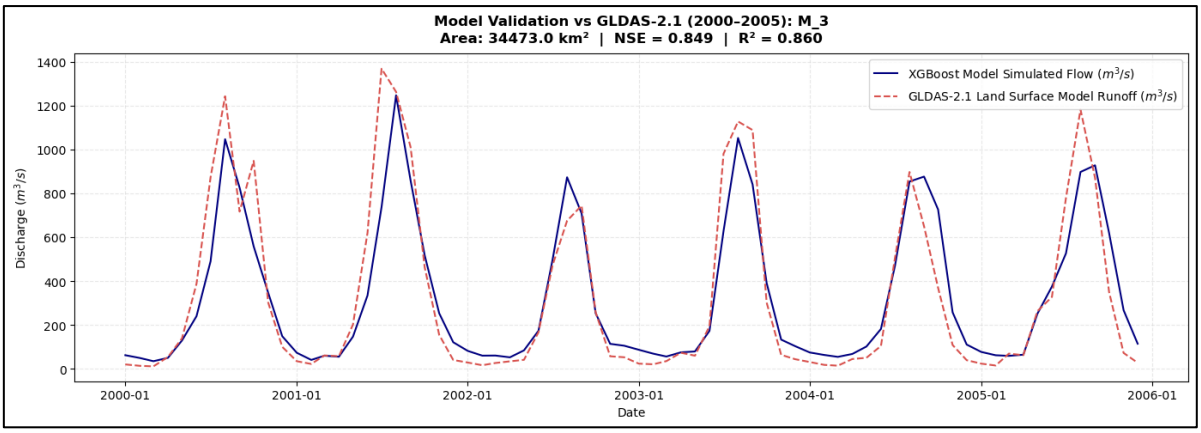


Figure A9. Simulated monthly flow at M_3 compared with GLDAS-2.1 runoff, 2000-2005.

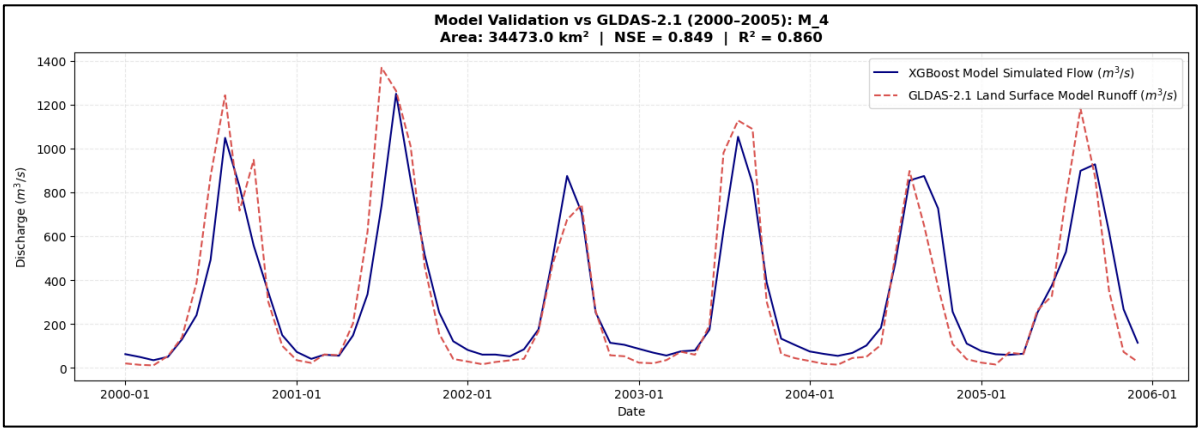


Figure A10. Simulated monthly flow at M_4 compared with GLDAS-2.1 runoff, 2000-2005.

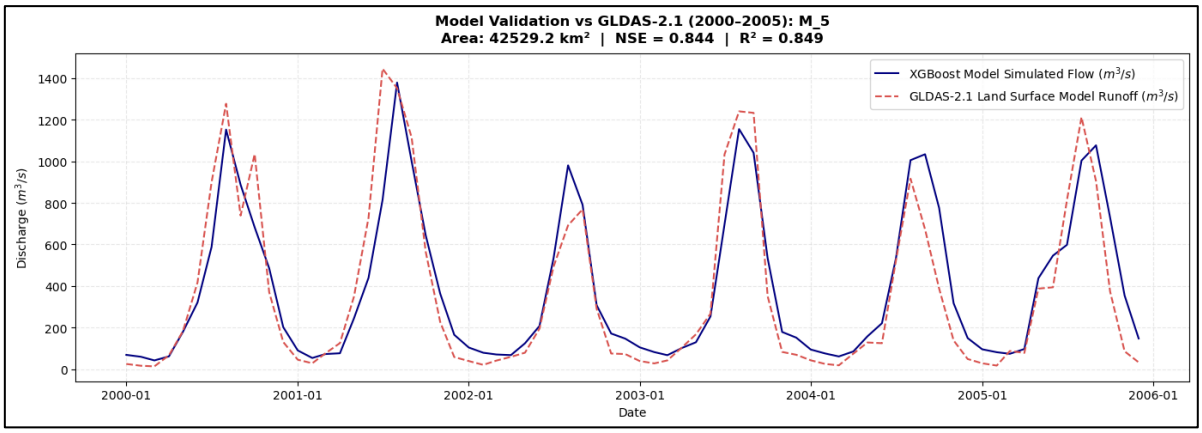


Figure A11. Simulated monthly flow at M_5 compared with GLDAS-2.1 runoff, 2000-2005.

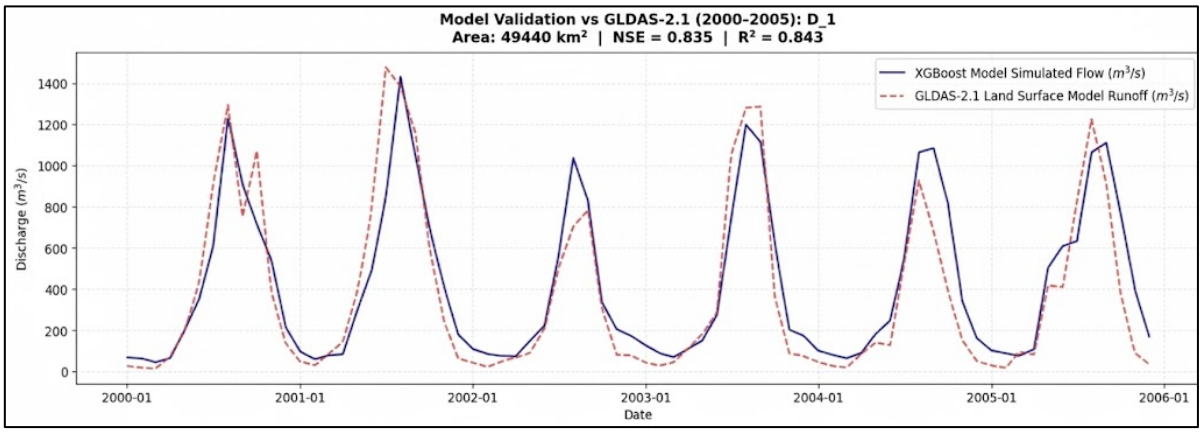


Figure A12. Simulated monthly flow at D_1 compared with GLDAS-2.1 runoff, 2000-2005.

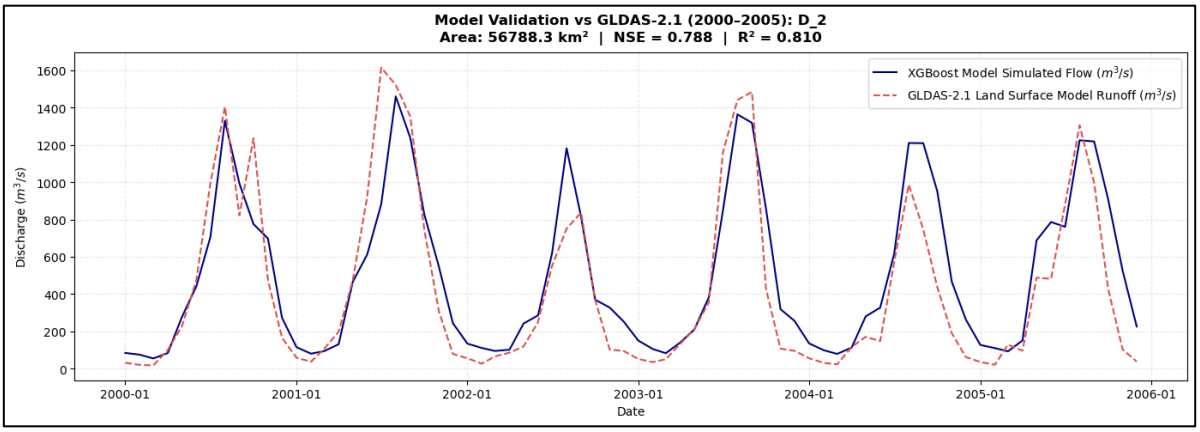


Figure A13. Simulated monthly flow at D_2 compared with GLDAS-2.1 runoff, 2000-2005.

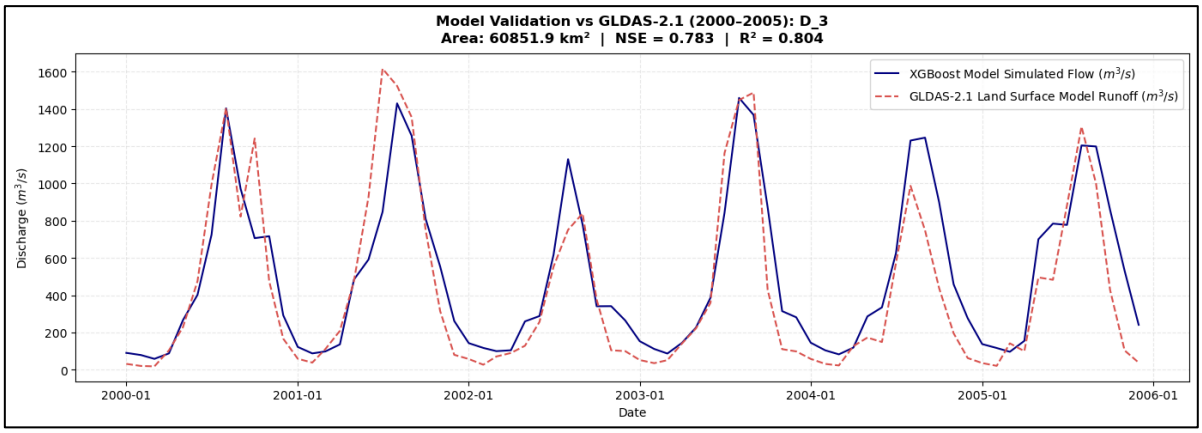


Figure A14. Simulated monthly flow at D_3 compared with GLDAS-2.1 runoff, 2000-2005.

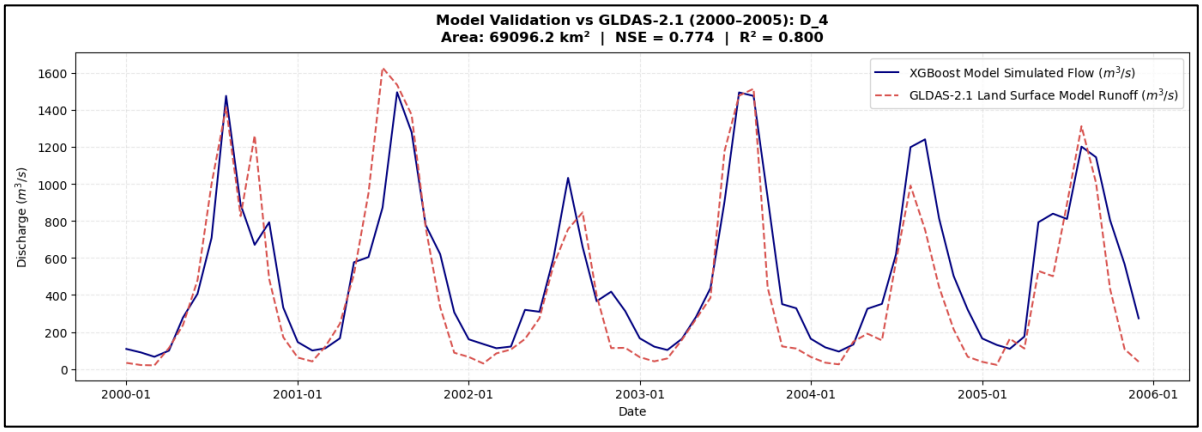


Figure A15. Simulated monthly flow at D_4 compared with GLDAS-2.1 runoff, 2000-2005.

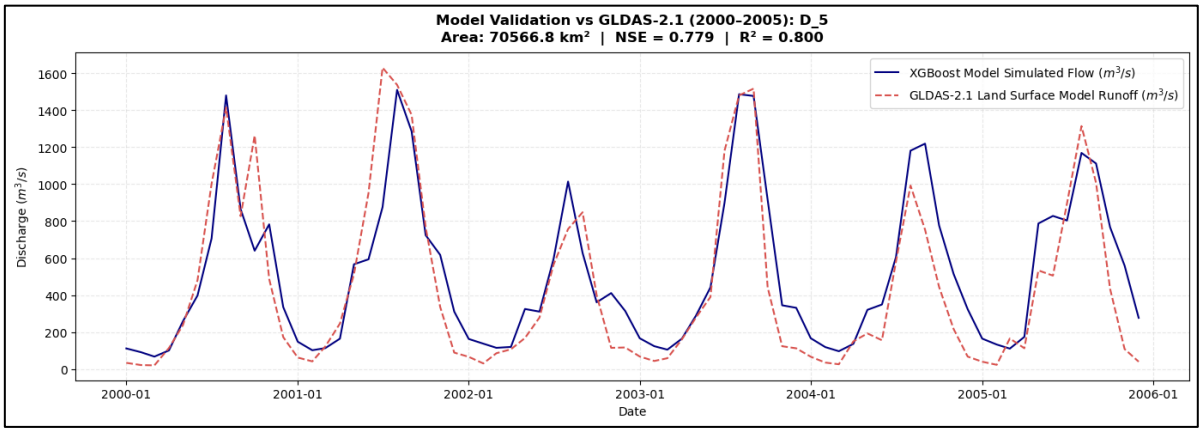


Figure A16. Simulated monthly flow at D_5 compared with GLDAS-2.1 runoff, 2000-2005.

Appendix B. Environmental flow assessment

This appendix gives the RDRM inputs, the simulated natural flow, the results for all ecological categories, and the monthly requirements at all 15 sites. The results for ecological categories A to D at each site are in Tables 3.2 and 3.3 of the Results.

B.1 RDRM inputs

Hydraulic parameters such as maximum width and slope were extracted from Google Earth, and the remaining parameters were set using the reference river of each sub-basin: Thina-2

for the upper, Asseg-2 for the middle, and Augr-3 for the lower sub-basin (section 2.3.3). Table B1 lists the main inputs. Several of these inputs are doubtful, as explained in section 4.

Table B1. Main RDRM inputs at the 15 sites: reference river, geomorphic zone, valley slope, maximum channel width, and bankfull discharge.

Site	Reference river	Geomorphic zone	Valley slope	Maximum channel width (m)	Bankfull discharge (m^3/s)
U_1	Thina-2	3	0.02	46.6	520
U_2	Thina-2	4	0.01	21.4	69
U_3	Thina-2	3	0.02	32.4	241
U_4	Thina-2	4	0.007	6,766.0	85
U_5	Thina-2	2	0.04	108.7	3,318
M_1	Asseg-2	3	0.027	541.4	8,735
M_2	Asseg-2	3	0.03	2,317.8	7,394
M_3	Asseg-2	2	0.07	2,508.4	7,337
M_4	Asseg-2	2	0.04	88.9	7,443
M_5	Asseg-2	4	0.009	100.8	8,476
D_1	Augr-3	4	0.01	111.3	3,615
D_2	Augr-3	4	0.007	145.4	4,762
D_3	Augr-3	4	0.007	238.8	8,069
D_4	Augr-3	4	0.008	99.9	3,276
D_5	Augr-3	4	0.004	912.4	13,897

B.2 Natural flow

Table B2. Mean monthly natural flow simulated at the 15 sites (Mm³ per month), by month of the hydrological year (October to September).

Site	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep
U_1	348	135	56	30	20	23	29	69	204	450	748	605
U_2	249	113	50	28	19	21	32	79	165	310	458	380
U_3	933	372	180	108	78	86	102	205	439	1,192	2,149	1,669
U_4	163	78	34	18	12	13	23	58	112	194	281	235
U_5	931	371	179	108	78	86	102	205	440	1,196	2,152	1,670
M_1	1,412	655	303	166	121	128	164	392	726	1,633	2,732	2,229
M_2	1,483	702	328	177	128	136	177	435	771	1,672	2,795	2,325
M_3	1,521	729	343	182	133	141	186	466	807	1,700	2,829	2,366
M_4	1,518	727	343	182	133	141	186	467	808	1,703	2,832	2,366
M_5	1,826	991	471	230	167	175	258	727	1,082	1,926	3,150	2,680
D_1	1,957	1,106	522	249	179	186	288	841	1,194	2,025	3,284	2,798
D_2	2,319	1,506	747	320	227	231	403	1,212	1,545	2,320	3,698	3,126
D_3	2,254	1,535	805	340	241	243	422	1,259	1,525	2,306	3,686	3,098
D_4	2,275	1,736	954	393	271	274	489	1,430	1,596	2,370	3,741	3,070
D_5	2,199	1,724	968	400	276	278	491	1,428	1,574	2,357	3,715	3,019

B.3 Requirements for all ecological categories

The RDRM also gives the requirements for the intermediate categories A/B, B/C, and C/D. Tables B3 and B4 give the total-flow and low-flow requirements for all seven categories.

Table B3. The total-flow environmental water requirement at the 15 sites for all ecological categories, in Mm³/yr, with the percentage of natural MAR in brackets.

Site	A	A/B	B	B/C	C	C/D	D
U_1	1,227 (45.1)	1,104 (40.6)	1,006 (37.0)	921 (33.9)	826 (30.4)	731 (26.9)	660 (24.3)
U_2	714 (37.5)	655 (34.4)	611 (32.0)	559 (29.3)	503 (26.4)	424 (22.2)	372 (19.5)
U_3	2,387 (31.8)	2,128 (28.3)	1,939 (25.8)	1,782 (23.7)	1,627 (21.7)	1,447 (19.3)	1,310 (17.4)
U_4	617 (50.5)	556 (45.5)	509 (41.6)	453 (37.1)	400 (32.7)	345 (28.3)	303 (24.8)
U_5	2,707 (36.0)	2,378 (31.6)	2,125 (28.3)	1,932 (25.7)	1,778 (23.6)	1,582 (21.0)	1,450 (19.3)
M_1	3,086 (28.9)	2,690 (25.2)	2,488 (23.3)	2,264 (21.2)	2,085 (19.6)	1,881 (17.6)	1,673 (15.7)
M_2	3,893 (35.0)	3,430 (30.8)	3,041 (27.3)	2,751 (24.7)	2,519 (22.6)	2,278 (20.5)	2,090 (18.8)
M_3	3,907 (34.3)	3,447 (30.2)	3,038 (26.6)	2,732 (24.0)	2,490 (21.8)	2,275 (19.9)	2,073 (18.2)
M_4	4,294 (37.6)	3,861 (33.9)	3,449 (30.2)	3,122 (27.4)	2,869 (25.2)	2,582 (22.6)	2,346 (20.6)
M_5	4,903 (35.8)	4,315 (31.5)	3,907 (28.6)	3,550 (25.9)	3,293 (24.1)	2,929 (21.4)	2,638 (19.3)
D_1	5,067 (34.6)	4,452 (30.4)	3,838 (26.2)	3,422 (23.4)	3,113 (21.3)	2,747 (18.8)	2,429 (16.6)
D_2	5,623 (31.9)	4,980 (28.2)	4,378 (24.8)	3,892 (22.0)	3,536 (20.0)	3,159 (17.9)	2,829 (16.0)
D_3	5,883 (33.2)	5,166 (29.2)	4,541 (25.6)	4,065 (22.9)	3,652 (20.6)	3,243 (18.3)	2,863 (16.2)
D_4	5,436 (29.2)	4,780 (25.7)	4,265 (22.9)	3,805 (20.5)	3,446 (18.5)	3,115 (16.7)	2,812 (15.1)
D_5	6,944 (37.7)	6,108 (33.1)	5,383 (29.2)	4,768 (25.9)	4,289 (23.3)	3,800 (20.6)	3,324 (18.0)

Table B4. The low-flow environmental water requirement at the 15 sites for all ecological categories, in Mm³/yr, with the percentage of natural MAR in brackets.

Site	A	A/B	B	B/C	C	C/D	D
U_1	185 (6.8)	143 (5.3)	119 (4.4)	92 (3.4)	72 (2.6)	52 (1.9)	37 (1.4)
U_2	164 (8.6)	142 (7.5)	130 (6.8)	114 (6.0)	100 (5.2)	54 (2.8)	33 (1.8)
U_3	240 (3.2)	150 (2.0)	119 (1.6)	92 (1.2)	72 (1.0)	35 (0.5)	19 (0.3)
U_4	155 (12.6)	130 (10.6)	108 (8.8)	80 (6.5)	59 (4.8)	32 (2.6)	20 (1.6)
U_5	317 (4.2)	204 (2.7)	140 (1.9)	79 (1.1)	47 (0.6)	18 (0.2)	8 (0.1)
M_1	696 (6.5)	517 (4.8)	457 (4.3)	387 (3.6)	331 (3.1)	269 (2.5)	219 (2.1)
M_2	415 (3.7)	217 (1.9)	149 (1.3)	92 (0.8)	61 (0.5)	37 (0.3)	23 (0.2)
M_3	440 (3.9)	226 (2.0)	156 (1.4)	94 (0.8)	62 (0.5)	37 (0.3)	24 (0.2)
M_4	804 (7.1)	619 (5.4)	551 (4.8)	470 (4.1)	405 (3.5)	332 (2.9)	272 (2.4)
M_5	906 (6.6)	682 (5.0)	606 (4.4)	516 (3.8)	442 (3.2)	360 (2.6)	292 (2.1)
D_1	1,374 (9.4)	977 (6.7)	710 (4.9)	512 (3.5)	396 (2.7)	292 (2.0)	219 (1.5)
D_2	1,411 (8.0)	1,012 (5.7)	758 (4.3)	556 (3.2)	424 (2.4)	306 (1.7)	223 (1.3)
D_3	1,758 (9.9)	1,360 (7.7)	1,088 (6.1)	822 (4.6)	637 (3.6)	465 (2.6)	343 (1.9)
D_4	1,262 (6.8)	848 (4.6)	595 (3.2)	421 (2.3)	320 (1.7)	231 (1.2)	168 (0.9)
D_5	2,576 (14.0)	2,107 (11.4)	1,750 (9.5)	1,375 (7.5)	1,100 (6.0)	830 (4.5)	626 (3.4)

B.4 Monthly requirements at the sites

Figures B1 to B15 show the median (50 percent exceedance) monthly natural flow and the total-flow requirement for ecological categories A to D at each site. Figures B5, B10, and B15 correspond to the outlet sites shown in the Results (Figures 3.12 to 3.14).

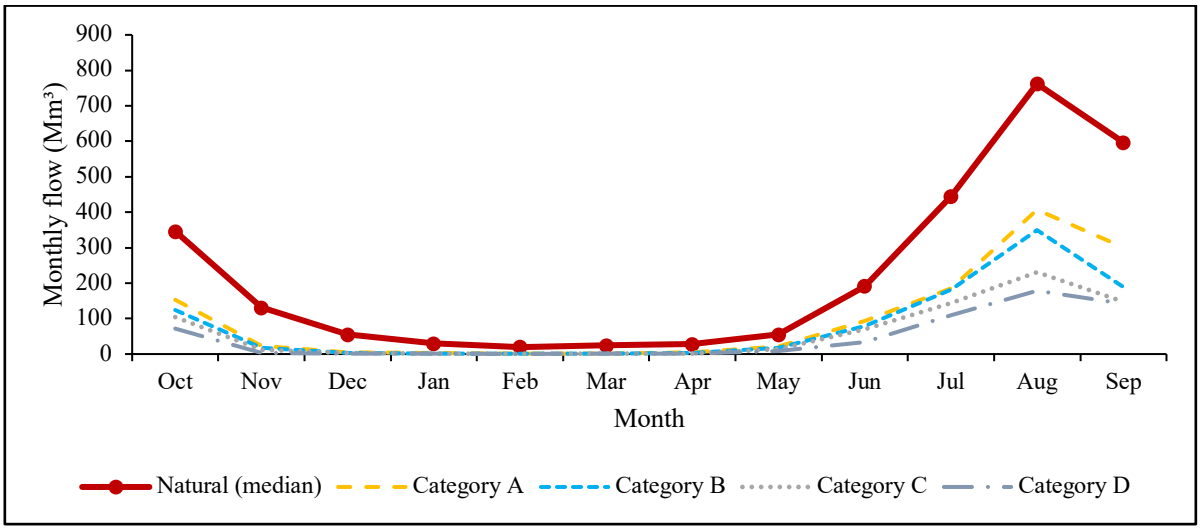


Figure B1. Median monthly natural flow and total-flow requirement for ecological categories A to D at U_1.

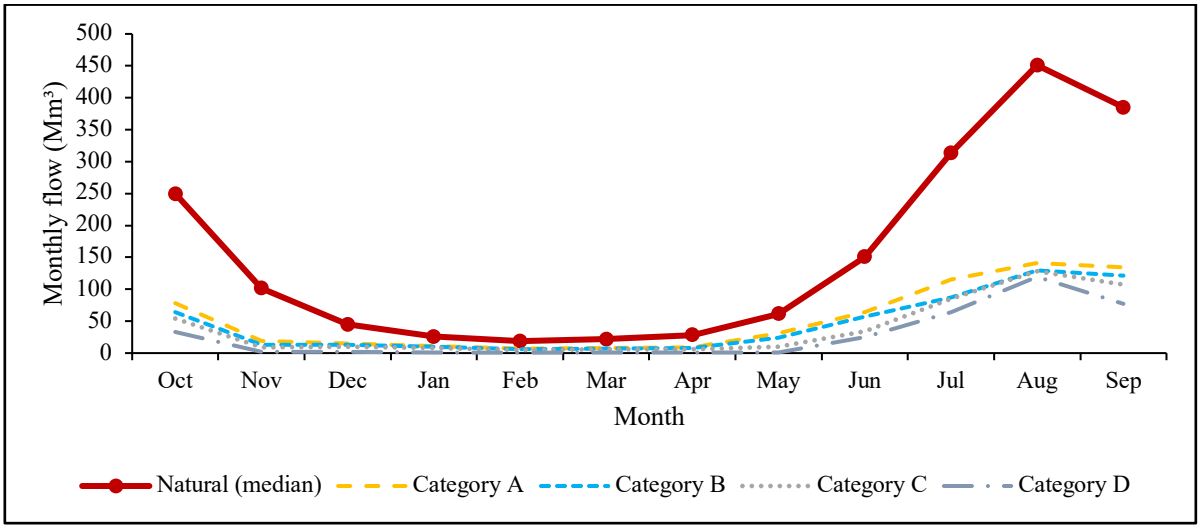


Figure B2. Median monthly natural flow and total-flow requirement for ecological categories A to D at U_2.

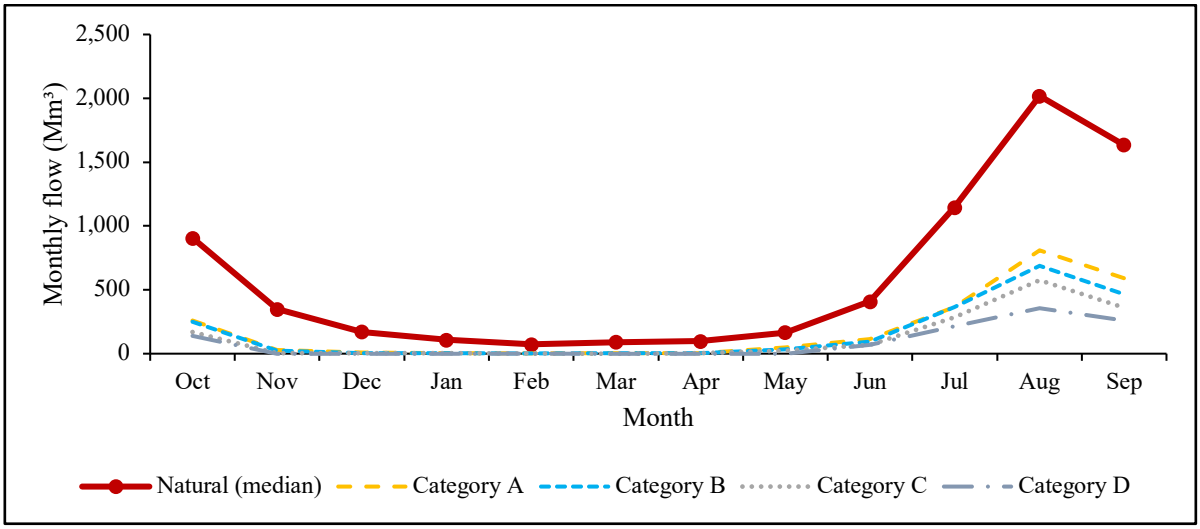


Figure B3. Median monthly natural flow and total-flow requirement for ecological categories A to D at U_3.

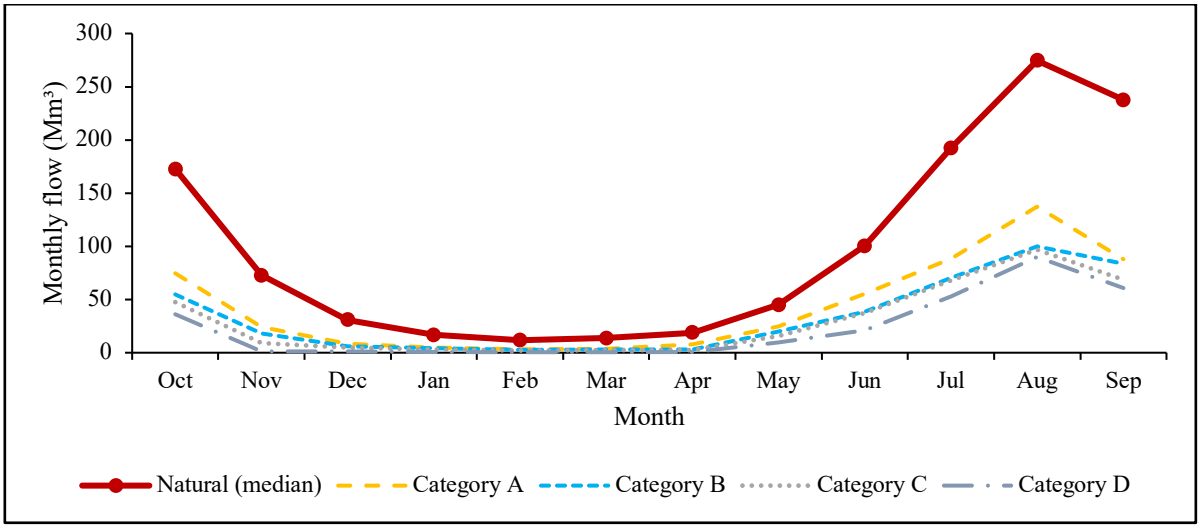


Figure B4. Median monthly natural flow and total-flow requirement for ecological categories A to D at U_4.

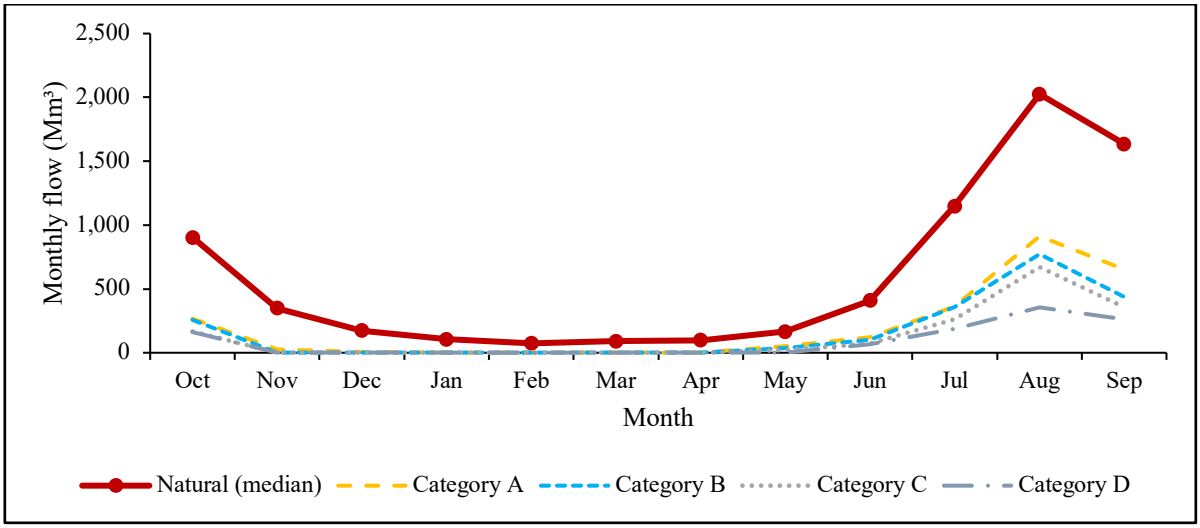


Figure B5. Median monthly natural flow and total-flow requirement for ecological categories A to D at U_5.

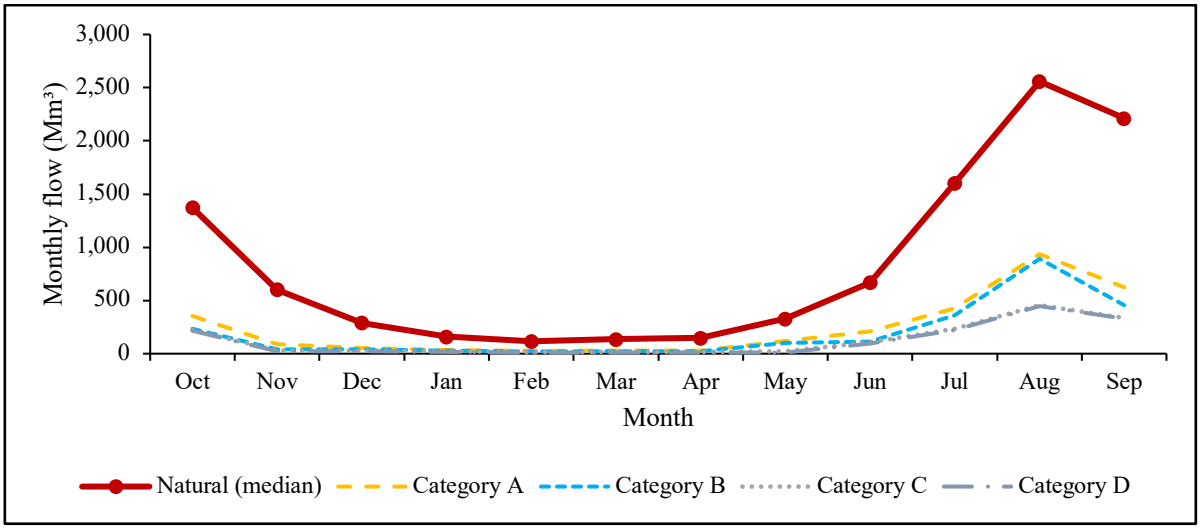


Figure B6. Median monthly natural flow and total-flow requirement for ecological categories A to D at M_1.

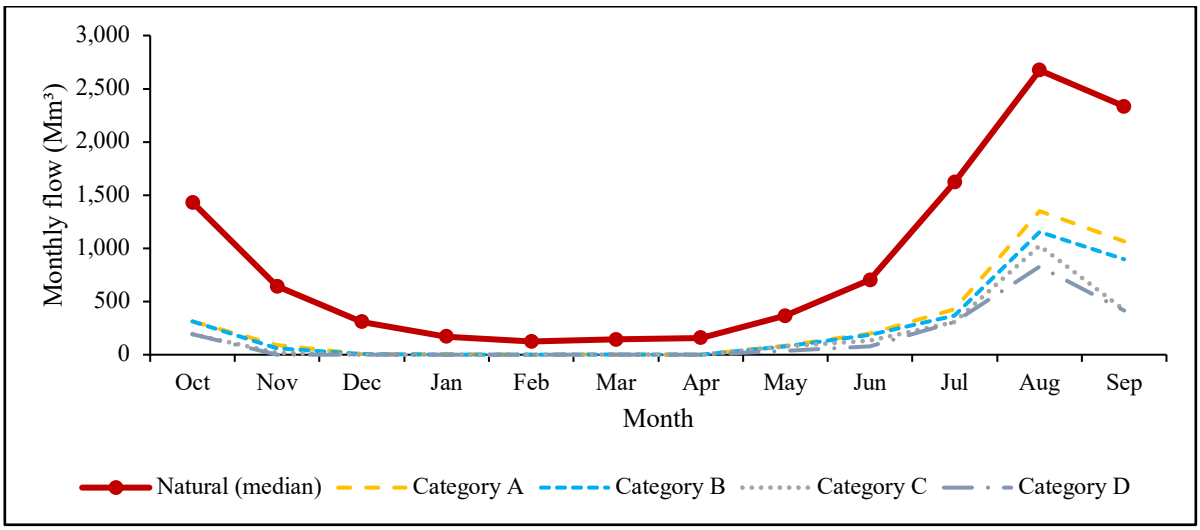


Figure B7. Median monthly natural flow and total-flow requirement for ecological categories A to D at M_2.

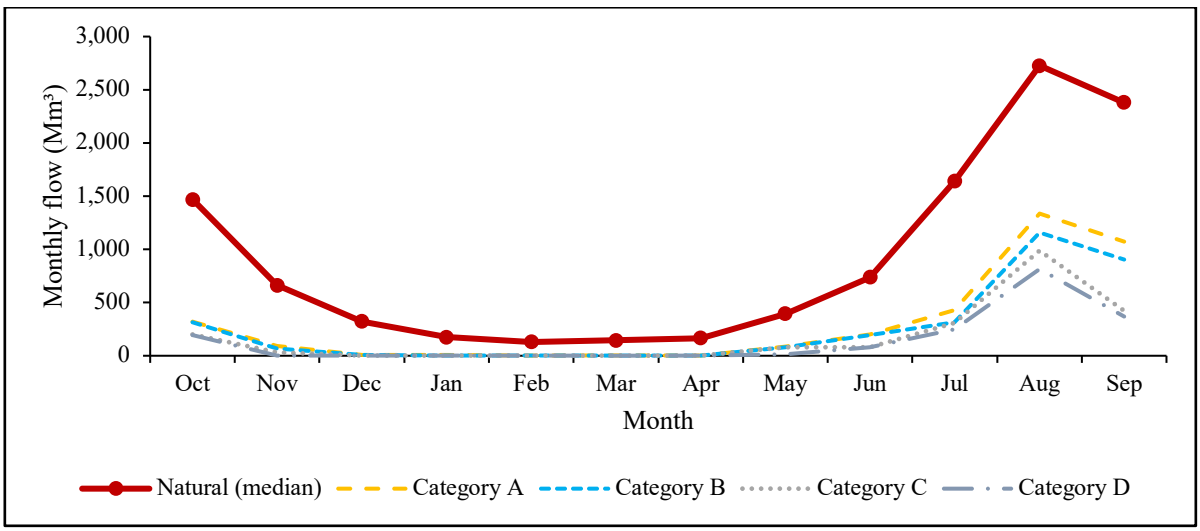


Figure B8. Median monthly natural flow and total-flow requirement for ecological categories A to D at M_3.

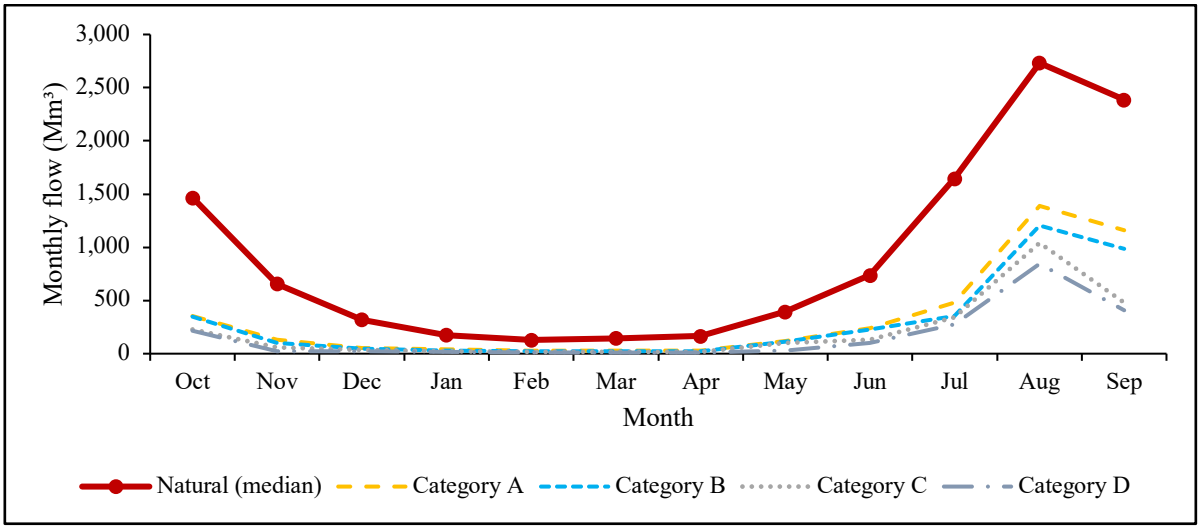


Figure B9. Median monthly natural flow and total-flow requirement for ecological categories A to D at M_4.

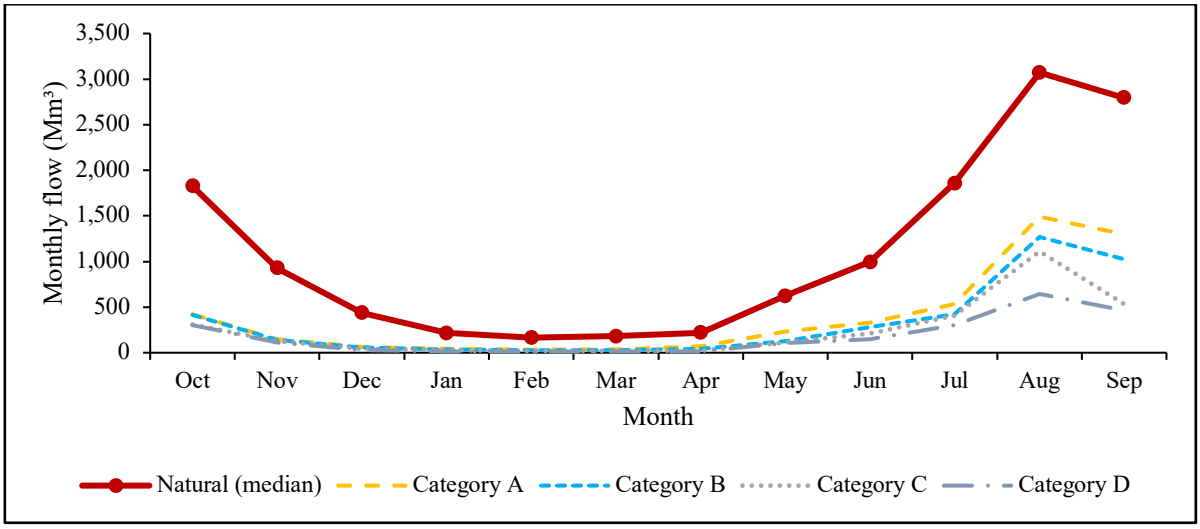


Figure B10. Median monthly natural flow and total-flow requirement for ecological categories A to D at M_5.

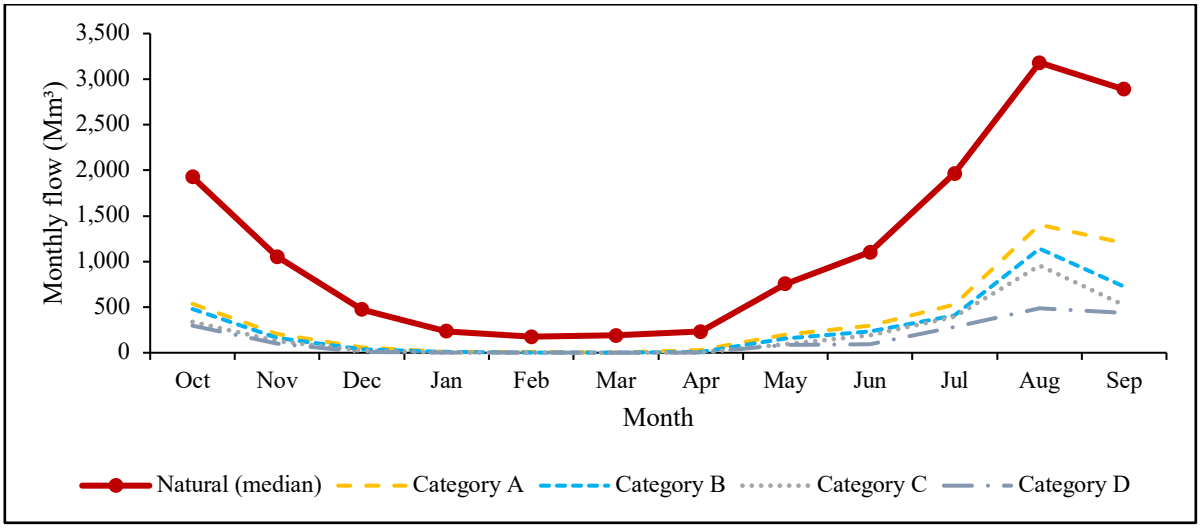


Figure B11. Median monthly natural flow and total-flow requirement for ecological categories A to D at D_1.

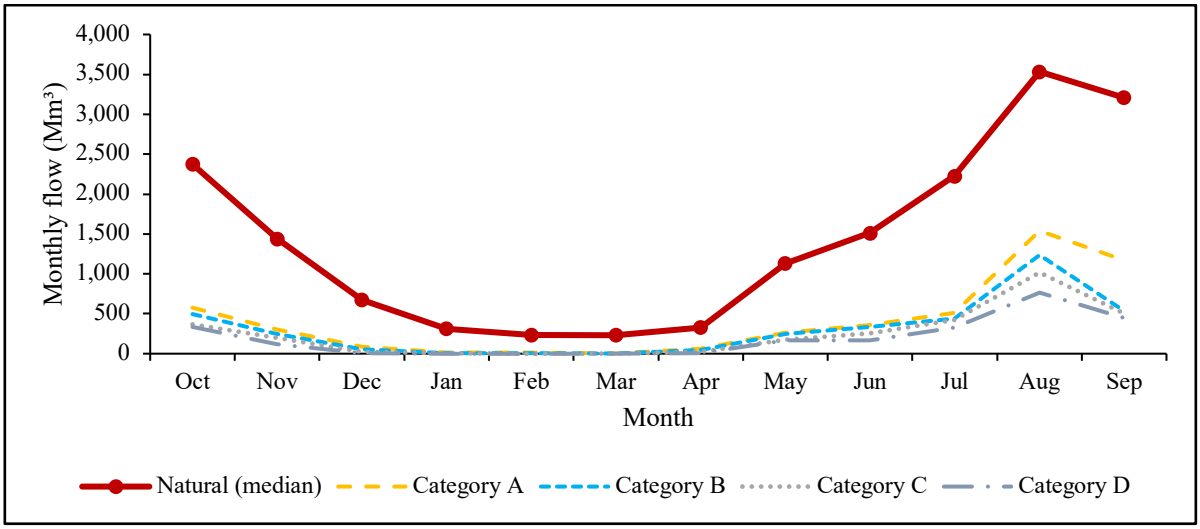


Figure B12. Median monthly natural flow and total-flow requirement for ecological categories A to D at D_2.

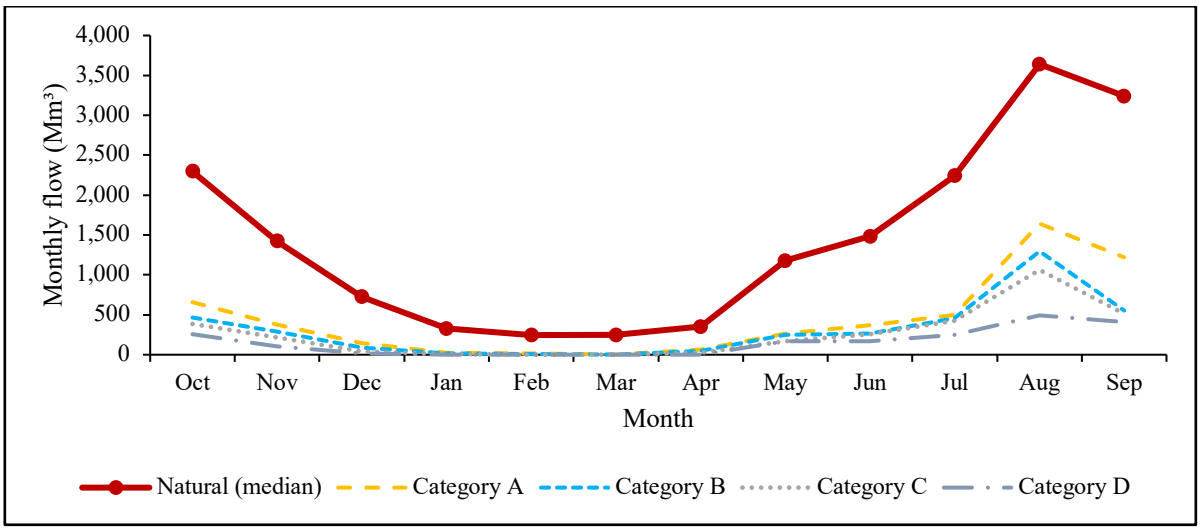


Figure B13. Median monthly natural flow and total-flow requirement for ecological categories A to D at D_3.

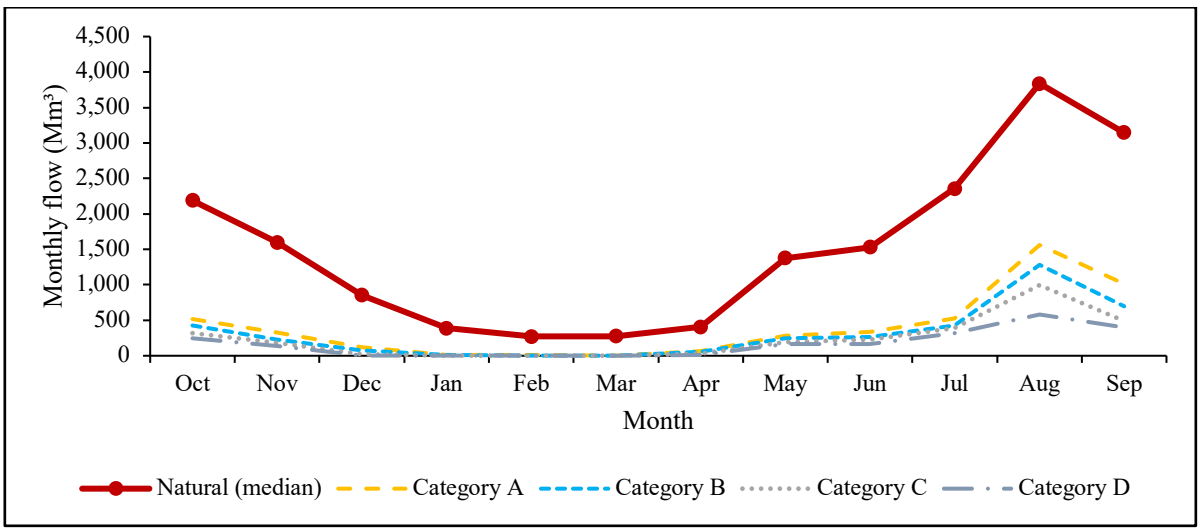


Figure B14. Median monthly natural flow and total-flow requirement for ecological categories A to D at D_4.

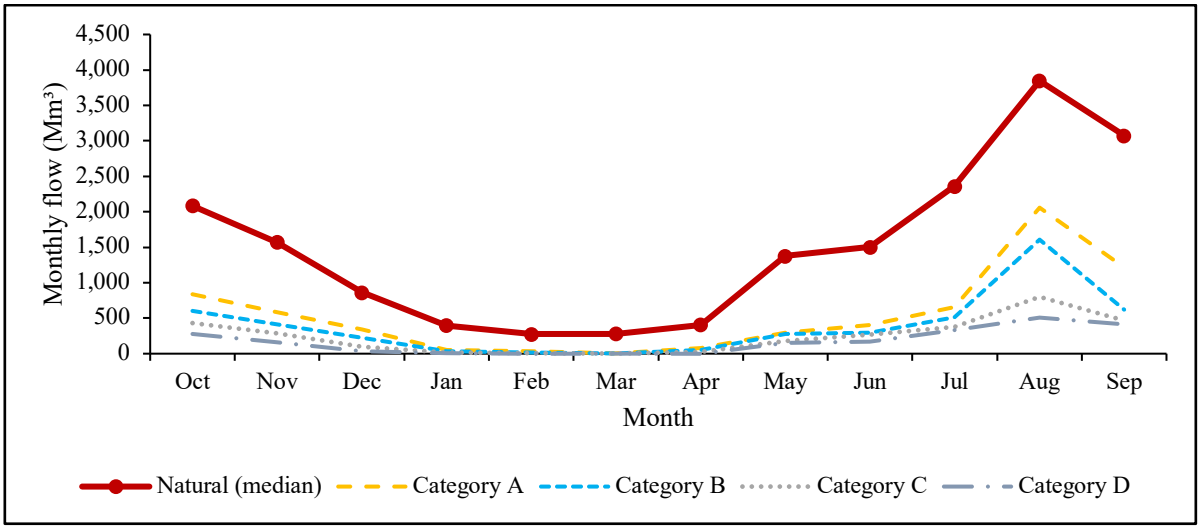


Figure B15. Median monthly natural flow and total-flow requirement for ecological categories A to D at D_5.

Appendix C. Stakeholder questionnaire

The questionnaire was sent to 25 participants of the May workshop, and 14 responded. This appendix lists the questions reported in the Results, gives the numbers behind Figures 3.15 to 3.21, and summarizes the answers to the open-ended questions.

C.1 Questions reported

Omo-Gibe Environmental Flows

Stakeholder input questionnaire

This questionnaire has been prepared following the 14 May workshop on environmental flows in the Omo-Gibe Basin. Its purpose is to support a clearer understanding of available data, key data gaps, and the basin-management decisions that would benefit most from environmental flows evidence prior to the next stakeholder engagement this September.

Participation in this survey is completely voluntary, and you may discontinue at any time. The information you provide will be used to identify available data, key data gaps, priority river stretches, decision-making needs, and capacity-development priorities for environmental flows work in the Omo-Gibe Basin. The findings will help shape discussions and next steps at the planned stakeholder workshop in Addis Ababa. Individual responses will not be shared; results will be presented only in aggregated form.

The questionnaire takes approximately **10 minutes** to complete.

Section A. Respondent information

1. Name

[answer]

2. Email

[answer]

3. Which category best describes your institution?
(Please select one)

- ☐ Government ministry or public authority
- ☐ University or research institution
- ☐ NGO or civil-society organization
- ☐ Consultancy / technical advisory firm
- ☐ Basin or river authority
- ☐ Independent expert
- ☐ Other: _____

4. Name of Organization and department
(Optional; this information will help avoid duplication in institutional representation.)

[answer]

Section B. Data availability and accessibility

5. What types of data does your organization collect or maintain for the Omo-Gibe Basin?
(tick everything that applies)

- ☐ River discharge / streamflow records
- ☐ Water quality measurements
- ☐ Rainfall or other meteorological data
- ☐ Groundwater or baseflow data
- ☐ Ecological data - fish, wetlands, biodiversity
- ☐ Abstraction or infrastructure-use records
- ☐ Water demand / socio-economic data
- ☐ None that I know of
- ☐ Other: _____

6. How would you rate gauge station coverage across the basin right now?
One row for the stretch above the Gibe I–IV cascade, one for downstream toward South Omo

Stretch	Good	Moderate	Poor	Not sure
Upstream of the Gibe cascade (Gibe I–V)				
Downstream toward South Omo				

7. How would you rate water quality data coverage across the basin right now?
One row for the stretch above the Gibe I–IV cascade, one for downstream toward South Omo

Stretch	Good	Moderate	Poor	Not sure
Upstream of the Gibe cascade (Gibe I–V)				
Downstream toward South Omo				

8. When the gauge data does exist, how easy is it to get hold of it?
(pick the one closest to your experience)

☐ Open - anyone can access it
☐ Available if you formally request it or sign a data-sharing agreement
☐ Technically exists but is very hard to obtain in practice
☐ Sits in one institution and rarely gets shared across ministries
☐ Not sure
Additional comments: _____

9. When the water quality data does exist, how easy is it to get hold of it?
(pick the one closest to your experience)

☐ Open - anyone can access it
☐ Available if you formally request it or sign a data-sharing agreement
☐ Technically exists but is very hard to obtain in practice
☐ Sits in one institution and rarely gets shared across ministries
☐ Not sure
Additional comments: _____

10. Any specific gauge stations, datasets, reports, or people we should be talking to for the lower catchment?
Names, contacts, whatever comes to mind - we'll follow up.
[answer]

C. Where the ~~priorities are~~

11. In your opinion, which stretches of the Omo-Gibe River most urgently need e-flows attention?
Could be a specific reach, or a stretch you'd nominate as a pilot site for follow-up work.
[answer]

12. In your opinion, where do environmental flows data need to plug into a decision?
(choose up to 3)

☐ Gibe cascade dam operations and release schedules
☐ Irrigation abstraction permitting
☐ Climate adaptation planning
☐ Ecosystem and biodiversity conservation
☐ Protecting downstream livelihoods (fisheries, pastoralism, floodplain farming)
☐ Urban or industrial water planning
Additional comments: _____

D. Capacity and governance

13. What's standing in the way of implementing environmental flows here?
(up to 3)

☐ Not enough continuous streamflow / gauge monitoring
☐ Limited in-house capacity for hydro-ecological modeling
☐ Institutions not coordinating with each other
☐ Weak regulatory or enforcement mechanisms
☐ Competing demands - energy vs. agriculture vs. ecology
☐ Not enough funding
Additional comments: _____

14. If we ran a training or capacity-building session alongside the follow-up workshop, what would be useful to you or your team?
e.g. hydrological modeling, telemetry/MIS systems, environmental flow indicator design - whatever's a real gap.

[answer]

15. Anything else - recommendations, context we're missing, or people who should be in the room for the September workshop?

[answer]

C.2 Numbers of responses

Table C1 gives the number of respondents (n = 14) who chose each answer. Questions 5, 12, and 13 allowed more than one choice, and questions 12 and 13 allowed up to three. The survey platform did not enforce the limit of three. In question 12 three respondents chose four, six, and six options, and in question 13 three respondents chose four, five, and six. Answers typed in instead of a listed option are marked as free text.

Table C1. Number of respondents choosing each answer to the closed questions.

Question	Answer	Respondents
3. Institution	University or research institution	5
	Government ministry or public authority	4
	Consultancy or technical advisory firm	4
	Independent expert	1
	NGO or civil-society organization	0
	Basin or river authority	0
5. Data held	Streamflow records	9
	Water demand or socio-economic data	9
	Water quality measurements	7
	Rainfall or other meteorological data	7
	Groundwater or baseflow data	6
	Abstraction or infrastructure-use records	6
	Ecological data	2
	None that I know of	3
6. Gauge coverage	Upstream: good / moderate / poor / not sure	2 / 10 / 1 / 1
	Downstream: good / moderate / poor / not sure	2 / 1 / 10 / 1
7. Water quality coverage	Upstream: good / moderate / poor / not sure	1 / 6 / 5 / 2
	Downstream: good / moderate / poor / not sure	1 / 2 / 9 / 2
8. Ease of getting gauge data	Formal request or data-sharing agreement	7
	Exists but very hard to obtain	4
	Sits in one institution, rarely shared	2
	Open to anyone	0
	Not sure	0
	Free text: available to universities by letter	1
9. Ease of getting water quality data	Formal request or data-sharing agreement	5
	Exists but very hard to obtain	2

Question	Answer	Respondents
	Sits in one institution, rarely shared	3
	Open to anyone	0
	Not sure	3
	Free text: only by payment and through publications	1
12. Decisions for e-flow data	Protecting downstream livelihoods	12
	Gibe cascade operations and releases	10
	Irrigation abstraction permitting	8
	Ecosystem and biodiversity conservation	7
	Climate adaptation planning	3
	Urban or industrial water planning	2
	Free text: no water-resources limitation in the basin	1
13. Barriers	Not enough continuous gauge monitoring	9
	Institutions not coordinating	9
	Weak regulation or enforcement	8
	Limited hydro-ecological modeling capacity	7
	Competing demands (energy, agriculture, ecology)	5
	Not enough funding	5
	Free text: recommendation to visit the basin	1

C.3 Open-ended answers

Question 10 (contacts for the lower catchment). Four respondents pointed to the national hydrology office of MoWE, including its surface water desk and the lead officer for hydrology and basin information. One named a regional irrigation and lowlands office, two named individual contacts, and one said that there are no stations downstream and suggested earth observation data and the investors at Omorate. Six respondents did not know of a specific contact or said none. Names and telephone numbers are not reproduced.

Question 11 (reaches most in need of e-flows attention). Nine respondents named the lower Omo, the reach below the Gibe III dam, or the stretch toward Lake Turkana and the delta. Some gave reasons, such as changes in the size, timing, and seasonality of flow and their

effects on sediment, floodplain processes, fisheries, riparian habitats, and livelihoods, and one suggested the reach just below Gibe III as a pilot site. One respondent named the upper and middle basin. Two respondents did not think that e-flows need attention, because the basin has plenty of water or flows naturally, and one of them added that the main problem is the management and use of water and land on the lower Omo. Two respondents gave no usable answer.

Question 14 (training that would be useful). Nine respondents named hydrological modeling, six named e-flow estimation methods or indicator design, five named telemetry, management information systems, or the management of gauge data, two named reservoir operation or hydraulic modeling, and one named cloud-based earth observation analysis. One respondent wrote that modeling is not possible without continuous gauge data, so gauges should be installed or reinstated first and technicians trained.

Question 15 (anything else). Four respondents raised the operation of the Gibe cascade, the effect of the dams on the water balance and the level of Lake Turkana, or floods around South Omo below the new dams. Four suggested who should attend the September workshop, including MoWE, the basin development offices, the meteorological institute, universities, basin authorities from Ethiopia and Kenya, and users of the lower Omo and Lake Turkana. Two asked for better data sharing and a shared open repository of pre-processed data. One respondent argued that environmental flow is not the main challenge in the basin and that the problem is the use of land and water, citing unused irrigable land, the large area held in parks, and a sugar scheme working below capacity. Four respondents gave no substantive answer.

Appendix D. Pathway to a digital twin of the Omo-Gibe Basin

This study gives a first-order desktop assessment, and its aim was to estimate indicative environmental water requirements for a basin where observed flow records, ecological data, and information on water use are all scarce. The study met that aim, but it also showed how fragmented the information base of the basin is. The current evidence is still far from supporting operational decisions on releases from the Gibe cascade, permits for irrigation abstraction, or protection of the lower Omo. This appendix treats the results as the starting point for a digital twin of the basin (section 4.6).

IWMI has already developed such a system for the Limpopo River Basin. The Limpopo Digital Twin supports data sharing, scenario analysis, and decision support for the four member states of the basin, bringing together Earth observation data, hydrological models, geospatial services, and prediction tools in a cloud environment (Jolaiya et al. 2025). It is organized in four layers (data, processing and services, applications, and DevOps and operations), and users interact with it mainly through the TerriaMap web map and the Water Copilot AI assistant. Its data-generating services cover hydrological modeling with SWAT+, water quality monitoring from Internet of Things (IoT) sensors, an Open Data Cube that indexes Earth observation products, and integration of external climate and evapotranspiration datasets. These feed an analytics service, the Water Copilot, and the TerriaMap dashboard (Jolaiya et al. 2025).

The Limpopo system is a mature production platform designed for real-time monitoring, prediction, and scenario testing. The Omo-Gibe Basin does not need to replicate it from the outset. The more useful questions are what a minimum viable twin for the Omo-Gibe would require, and how much of it already exists. Table D1 sets out eight minimum requirements (MR) derived from the Limpopo architecture, which serve as the reference point for the rest of this appendix.

Table D1. Minimum requirements for a first-generation digital twin of the Omo-Gibe Basin and corresponding components of the Limpopo Digital Twin.

Code	Minimum requirement	Corresponding Limpopo Digital Twin component
MR1	Spatial framework: basin, river network, infrastructure, and decision points	Physiographic layers (elevation, channel network, basin and sub-basin boundaries) indexed in the Open Data Cube and displayed in TerriaMap
MR2	Hydroclimatic forcing data that can be updated	Open Data Cube service and external data integration service (e.g., CHIRPS rainfall and evapotranspiration products)
MR3	Observed flow and water quality data for calibration, validation, and updating	Hydrological observations and water quality monitoring service (IoT sensors)
MR4	Hydrological model estimating flows at decision points	Hydrological modeling service (SWAT+)
MR5	Decision layer comparing environmental water requirements with actual or regulated flows	Analytics service
MR6	Scenario capability for dam operation, abstraction, and climate	SWAT+ scenario assessments; scenario exploration in TerriaMap

MR7	Platform for storage, access, visualization, and deployment	Containerized microservices, cloud storage, single sign-on with role-based access, staging and production environments, automated deployment
MR8	Governance: data sharing, ownership, and maintenance	Support from the Limpopo Watercourse Commission (LIMCOM); decision support for the four member states

Source: Authors' synthesis; Limpopo components from Jolaiya et al. (2025). The grouping into eight minimum requirements is the authors' own.

D.1 Ready components

Several building blocks developed in this study could form the core of a first-generation twin (Table D2). The spatial framework (MR1) is largely complete. The basin, sub-basins, and river network are defined from HydroBASINS, HydroRIVERS, and HydroSHEDS; the upper, middle, and lower sub-basins are delimited at the Gibe II scheme and the Koysha dam; and the 15 assessment sites are georeferenced together with the hydropower schemes, reservoirs, and the 52-station network of MoWE (section 1; Figure 2.3). Forcing data (MR2) are also in place: CHIRPS rainfall and ERA5-Land radiation, evapotranspiration, and soil moisture were extracted for each assessment catchment through a reproducible workflow, and GLDAS-2.1 runoff was used as a consistency check. These are the same kind of open gridded products already ingested by the Limpopo twin, whose data layer includes CHIRPS (Jolaiya et al. 2025).

The study also produced a reproducible hydrological model (MR4) and a set of environmental water requirements (MR5), summarized in Table D2 together with their main limitations. Although the monthly model was selected, a daily configuration was also built and tested, with lower performance (section 3.1.2); it provides a starting point for the daily modeling discussed below. Observed flows (MR3), however, cover only 1981 to 2005, come from four northern stations, and end before Gibe II, Gibe III, and Koysha were commissioned.

Table D2. Components currently available for an Omo-Gibe digital twin, by minimum requirement.

Requirement	Available component	Main limitation
MR1	HydroBASINS, HydroRIVERS, and HydroSHEDS layers; three sub-basins; 15 georeferenced sites; hydropower schemes, reservoirs, and MoWE stations; SRTM terrain	Sites selected at random and not at decision points; two pairs share a catchment, giving 13 distinct catchments
MR2	CHIRPS daily rainfall (0.05°); ERA5-Land radiation, potential evapotranspiration, and soil moisture (0.1°); GLDAS-2.1 runoff; reproducible extraction workflow	Not yet indexed on a platform
MR3	Daily discharge, 1981 to 2005, at Megecha, Ajancho, Asendabo, and Abelti	Northern basin only; ends before the cascade; Ajancho record questionable; no abstraction records
MR4	XGBoost model of naturalized runoff with fixed hyperparameters and seed; monthly test-period NSE 0.78 to 0.89 and PBIAS +6.0% to -	Poor leave-one-station-out performance; no routing, losses, or abstraction; land cover fixed at 1981 to 2005 conditions

	13.3%; daily configuration also tested	
MR5	Environmental water requirements for ecological categories A to D at 15 sites (RDRM); total requirement at sub-basin outlets about 36% to 38% of natural MAR (category A) and 18% to 19% (category D)	Static outputs; low-flow requirements and very low dry-season values unreliable
MR6	Benchmarks: Gibe III interim minimum release (25 m³/s) and 10-day controlled flood; Gibe III inflow estimate; Lake Turkana inflow and level estimates	No scenario model
MR7	IWMI Limpopo Digital Twin, designed to integrate new datasets, models, and applications without redesigning existing services	Multi-basin hosting not documented
MR8	Stakeholder questionnaire (14 respondents); May and September 2026 workshops; MoWE Basin Monitoring Strategy 2026 to 2035	No basin or river authority or civil-society respondent

Source: This study; benchmarks from Agriconsulting S.p.A. and Mid-Day International Consulting Engineers (2009), Velpuri and Senay (2012), and Avery and Tebbs (2018); platform information from Jolaiya et al. (2025). NSE = Nash-Sutcliffe efficiency; PBIAS = percent bias; MAR = mean annual runoff; RDRM = Revised Desktop Reserve Model.

D.2 Near-ready components

Several improvements require mainly desk work, data requests, or reuse of existing tools and not new infrastructure (Table D3). In the Limpopo data architecture, raster products are stored as Cloud Optimized GeoTIFFs with standard metadata and indexed in the Open Data Cube (Jolaiya et al. 2025). The Omo-Gibe forcing datasets could be indexed in the same way. The flow model could then be run from 2005 to the present and packaged as a containerized service communicating through REST interfaces, as each Limpopo microservice does; the resulting flows would still reflect 1981 to 2005 land cover, and this would need to be stated clearly. Because map catalogues can be updated without redeploying code, and new datasets can be tested in a separate staging environment, the cost of adding the Omo-Gibe as a new set of layers is relatively low.

The e-flow assessment can also be strengthened at low cost. Section 4.5 identifies correction of the RDRM channel inputs, followed by a sensitivity analysis, as a priority: several inputs are implausible, including maximum channel widths of about 6,766 m at U_4 and 912 m at D_5 (Table B1), and bankfull discharge is inconsistent between sites that share a catchment. Re-examining the Ajancho record and refitting the model would restore the only small, low-flow catchment in the training data.

Two further gains depend on access to data that already exist. As some of these data are likely to be restricted, the role-based access control of the Limpopo platform and its practice of keeping sensitive datasets in private storage until they are cleared for publication (Jolaiya et al. 2025) would allow MoWE data to be used without being made public.

Table D3. Improvements achievable with modest effort, and the minimum requirements they address.

Action	Requirement	Basis	Expected gain

Index CHIRPS, ERA5-Land, and GLDAS for the basin in an Open Data Cube	MR2, MR7	Same data pipeline as the Limpopo twin	Forcing data available on the platform
Run the model from 2005 to the present and deploy it as a containerized service	MR4	Fixed hyperparameters and seed; forcing products continuously updated	First live hydrological service
Correct RDRM channel inputs and conduct a sensitivity analysis	MR5	Priority recommendation of this study (section 4.5)	More credible low-flow requirements
Re-examine the Ajancho record and refit the model	MR3, MR4	Existing data	Restores the only small, low-flow training catchment
Relocate assessment sites to decision points (Gibe III and Koysha outlets, Kuraz, lower floodplain, delta)	MR1, MR5	Rerun of the existing workflow	Outputs aligned with management needs
Request post-2005 records, including from the four telemetry stations	MR3	Data request to MoWE	First comparison of natural and regulated flows
Request existing water quality data	MR3	Data request to MoWE	Historical water quality layer
Add satellite-derived Lake Turkana levels as an indicator	MR5	Approach applied by Avery and Tebbs (2018)	Downstream outcome indicator
Add Omo-Gibe layers to a staging environment with role-based access	MR7, MR8	Existing Limpopo platform features	Restricted MoWE data shared without publication
Data-sharing memorandum of understanding with MoWE, with targeted training	MR8	Respondent requests for shared data and training (section 3.3)	Institutional basis for the twin

Source: Authors' synthesis based on this study and Jolaiya et al. (2025).

These steps also respond directly to stakeholder requests. Two respondents called for better data sharing and a shared open repository of pre-processed data, which is essentially what an Open Data Cube provides. Training needs identified by respondents included hydrological modeling (9 respondents), e-flow methods or indicator design (6), telemetry, management information systems, or gauge data management (5), and cloud-based Earth observation analysis (1) (Appendix C). Training of this kind would build the national capacity needed to use and maintain the system.

D.3 Components requiring investment

The remaining gaps require new infrastructure, sustained fieldwork, or institutional agreements (Table D4). The most serious constraint is monitoring. Twelve of the 52 MoWE stations are non-functional or closed, only four have telemetry, and the Gilgel Gibe I reservoir requires seven gauges to measure its inflows and outflow but has only two in operation (section 1; MoWE 2026). Ten of 14 respondents rated downstream gauge coverage as poor; one noted

that modeling is not possible without continuous gauge data, and another, observing that there are no stations downstream, suggested Earth observation as an alternative. As noted in section 4.4, this creates a circular problem: the model is most needed where it can least be tested. The MoWE Basin Monitoring Strategy proposes 32 new stations, which would bring the basin total to 74. Delivering these stations, adding telemetry, and securing long-term funding for operation and maintenance are preconditions for a twin that is updated from observations and not from models alone.

The ecological basis of the assessment requires equally long-term work. No PES or EIS assessment exists, so no target category can yet be recommended, and the RDRM parameters were transferred from South African rivers matched only by channel type (section 2.3.3). A monthly time step also cannot represent flood pulses or flood-recession agriculture in the lower Omo, and a single 10-day flood cannot reproduce the sustained July to October high flows. The decline in Lake Turkana levels during the filling of Gibe III in 2015 and 2016 (Avery and Tebbs 2018) shows why the link to the lake matters, and respondents themselves called for the involvement of basin authorities in both Ethiopia and Kenya.

Table D4. Gaps requiring significant investment, by minimum requirement.

Requirement	Gap	Work required
MR3 Observations	12 of 52 stations non-functional or closed; 4 with telemetry; poor downstream coverage	32 new stations under the MoWE strategy; telemetry; long-term funding for operation and maintenance
MR3 Water quality	Downstream coverage rated poor by 9 of 14 respondents	New monitoring, potentially following the Limpopo IoT sensor model
MR4 Hydrology	Model not shown to transfer beyond the gauged northern catchments; no routing, losses, or abstraction	Gauges in the middle and lower basin; routing, transmission losses, floodplain evaporation, land-cover change, and abstraction; daily time step; choice between SWAT+ and a hybrid model
MR5 Ecology	No PES or EIS; transferred South African parameters; no agreed target category	Field surveys of fish, macroinvertebrates, and riparian and floodplain condition; local calibration; targets agreed with stakeholders
MR5 and MR6 Flood pulse	Monthly results cannot represent flood pulses or flood-recession agriculture	Daily hydraulic and floodplain modeling below Gibe III and Koysha
MR6 Scenarios	Cascade operations, Koysha, and irrigation abstraction (e.g., Kuraz) outside scope; no abstraction records	Operating rules and release data from dam operators; reservoir simulation
MR6 Transboundary	Lake Turkana levels outside scope	Lake water-balance model; data sharing between Ethiopia and Kenya
MR3 and MR5 Other components	Sediment, groundwater, social needs, and climate change excluded	New monitoring and modeling components

MR7 Platform	No Infrastructure as Code; backups held in a single cloud region (Cape Town)	Automated provisioning; multi-region redundancy; decisions on hosting and data sovereignty
MR8 Governance	Poor inter-institutional coordination (9 of 14) and weak enforcement (8 of 14); no basin authority among respondents	A mandated institutional owner; long-term funding

Source: Omo-Gibe gaps from this study; platform gaps from the limitations section of Jolaiya et al. (2025). The choice between SWAT+ and a hybrid model and the hosting and data-sovereignty decisions are the authors' assessment.

Several of these gaps are as much institutional as technical. Scenarios of cascade and irrigation operation depend on access to operating rules held by dam operators, and a twin will only be sustained if an institution with a clear mandate owns it and uses it in decision-making. The Limpopo platform also has limitations of its own: it does not yet implement Infrastructure as Code, and its disaster recovery relies on backups held within a single cloud region (Jolaiya et al. 2025). An Omo-Gibe twin built on the same platform would inherit both.

D.4 A phased pathway

Tables D2 to D4 point to a phased approach, summarized in Table D5 as readiness against each minimum requirement. In the first phase, a planning twin would be built from the components already available and the improvements listed in Table D3, providing scenarios of natural flow and environmental water requirements on the existing IWMI platform. In the second phase, data from rehabilitated and telemetered gauges would turn this into a monitored twin, able to compare regulated and natural flows and track compliance with e-flow targets. Only in the third phase, once the investments in Table D4 have added cascade operations, daily flood-pulse requirements, an ecological response layer, and the link to Lake Turkana, would the system support operational decisions.

Table D5. Readiness of the Omo-Gibe Basin against the minimum digital twin requirements.

Requirement	Available now	Achievable with modest effort	Requires significant investment
MR1 Spatial	Yes	Relocate sites to decision points	None required
MR2 Forcing	Yes (data)	Index on platform; extend to present	None required
MR3 Observations	1981 to 2005, northern basin only	Post-2005 MoWE records; water quality archives	Downstream gauges, telemetry, water quality monitoring, operation and maintenance
MR4 Hydrology	Prototype validated in time only; weaker daily version	Deploy as a service; refit with Ajancho	Basin-wide daily model with routing
MR5 Decision layer	Static RDRM outputs	Corrected inputs; sensitivity analysis; Lake Turkana level indicator	PES and EIS; local calibration; agreed targets
MR6 Scenarios	Benchmarks only	None identified	Cascade, Koysha, abstraction, flood pulse, climate, Lake Turkana

MR7 Platform	IWMI Limpopo Digital Twin	Omo-Gibe layers in staging with role-based access	Infrastructure as Code, redundancy, hosting decision
MR8 Governance	Questionnaire, workshops, monitoring strategy	Memorandum of understanding; training	Mandated owner; data sharing between Ethiopia and Kenya

Source: Authors' synthesis of Tables D2 to D4, based on this study and Jolaiya et al. (2025).

Progress through these phases will depend on five research questions: how the flow model can be transferred to the ungauged middle and lower basin; whether desktop e-flow parameters are valid for Ethiopian rivers; how flood-pulse requirements should be set below the dams; how river and floodplain ecosystems respond to flow alteration; and how Omo flows translate into Lake Turkana levels. The answers will determine both what the twin can do and how far decision-makers can rely on it.



CGIAR MULTIFUNCTIONAL LANDSCAPES IWMI International Water Management Institute