Guide to Hydrologic Modeling
of Natural Infrastructure

Natural Infrastructure
for Water Security

Authors:
Boris F. Ochoa-Tocachi[1,2]
José Cuadros-Adriazola[3]
Edwing Arapa[3]
Natalia Aste[3]
Eric Ochoa-Tocachi[1]
Vivien Bonnesoeur[3]

Reviewers
Morgane Lalonde-Le Pajolec[2]
Víctor Alarcón[3]
Fabian Drenkhan[2]
Francisco Román[3]
Wouter Buytaert[2]
Gena Gammie[4]

Edited by Forest Trends Association. Av. Ricardo Palma 698, Miraflores
1st edition, june 2022

Production and Editing: Gabriel Rojas, Forest Trends[4]

Mediation: Mercy Sandoval[4] | Guisella Infantes[4] | Gabriel Rojas[4]

Translation: Cheyenne Coxon[4]

Proofreading: Antonio Luya Cierto

Design and layout: Javier Domínguez | Juan José Vásquez

Cover photo: Natalia Aste

This document was made possible through the support of the United States Agency for International Development and the Government of Canada. The views expressed in this document are those of the authors and do not necessarily reflect the views of the U.S. Agency for International Development or the Government of Canada.

[1] ATUK Consultoría Estratégica; Cuenca, Ecuador.
[2] Imperial College London; London, United Kingdom.
[3] CONDESAN; Lima, Perú.
[4] Forest Trends; Washington DC, United States.

Photo: Julio Reaño

Photo: Renny Daniel Díaz Aguilar

Prologue

Every day, we see the evidence of growing dual crises of climate change and water insecurity. Peaks and valleys in the water cycle translate to floods and droughts in homes and businesses – and loss and damage in our communities.

At the same time, we are losing critical allies in this fight: deforestation and ecosystem degradation have risen to unprecedented levels – weakening the very ecosystems that could help us to adapt to the extremes of our new reality.

Nevertheless, we also see unprecedented commitments to act on this crisis. Governments, businesses, and civil society are starting to come together to build innovative policy and financial solutions to address these crises. In the realm of natural infrastructure for water, hundreds of billions of dollars in new finance are on the table in developed and developing countries alike.

Prioritizing those resources and mobilizing investments will require clear, quantified estimates of how we expect specific interventions in natural infrastructure to benefit our communities by addressing water risks. Local stakeholders, water resource planners, project developers, and technical experts need to work together – strategically and swiftly -- to use the best available information and each other's expertise to act with the care and urgency this moment requires.

In Peru, our Natural Infrastructure for Water Security project aims to mobilize and scale action in nature-based solutions for water. Our experience has taught us that hydrological models play a fundamental role in this process – and it has also taught us that, unfortunately, in practice they often fall short of generating information that is useful and utilized to make better decisions.

This publication is our response to a need we found for accessible guidance to quantify the water benefits of natural infrastructure interventions that could be accessible for integrated teams of decision-makers and technical specialists. We hope this guide is a valuable contribution to the design, evaluation, and mobilization of investments that are so necessary in this moment.

Fernando Momiy Hada
Director, Natural Infrastructure for Water Security project
Forest Trends

Gena Gammie
Director, Water Initiative
Forest Trends

TABLE OF CONTENTS

Photo: Ana Castañeda

Introduction

Around the world, governments, businesses, and communities face an increasingly difficult and uncertain future as water crises intensify and become more complex due to factors such as population growth and climate change. Many stakeholders are now considering a new asset class to address water risks: *natural infrastructure*, including forests, grasslands, wetlands, and other features that make up our watershed landscapes and perform critical functions to support reliable water quantity and quality.

While natural infrastructure for water has gained momentum in recent years, moving from aspiration to action requires answering some key questions, such as: *Where are the priority areas for natural infrastructure investments, given the water risks we face and the water infrastructure we depend on? How and how much should we invest in natural and conventional infrastructure to achieve our water goals? How can we choose among proposed natural infrastructure projects?*

Hydrologic models are a fundamental tool to help water managers answer these questions, understand their options, and make good decisions. A hydrologic model is a simplified representation of the water cycle that uses concepts and approximations of the actual processes of the hydrologic system. Through comparisons between representations of reality and alternative scenarios that we can generate in a hydrological model, we can produce quantitative estimates of the water benefits of different courses of action (or inaction) in the landscape.

However, using hydrologic models poses many challenges that may prevent the generation of useful information for decision making, especially on natural infrastructure. Many of these challenges come down to the technical aspects of the analysis. While there are several manuals and guides written by and for technical specialists to support these aspects (e.g., Beven, 2011; Arnold et al., 2021; Pianosi et al., 2015; Rogers et al., 2018), many of these challenges have to do with the scope and focus of the modeling study, the selection of the most appropriate model, and the interpretation of the results—aspects that have to do with both technical considerations and non-technical values and interests. Therefore, the success of the modeling process requires that technical specialists, stakeholders, and decision-makers work together to generate information that ultimately serves to make good decisions.

Using hydrologic models to improve decisions about natural infrastructure requires understanding their strengths as well as their limitations. Hydrologic modeling, in and of itself, should be viewed as a useful but also uncertain tool. As in the evaluation of gray infrastructure, using information generated by models intelligently requires transparency about potential sources of error, critical analysis of the results obtained, and recognition that real values may be outside the confidence interval obtained with modeling.

Also, **modeling is not a replacement for monitoring**. In data-scarce areas, there is often an interest in using hydrologic modeling to estimate the impacts of a course of action—and models can add value in these cases. However, models themselves are most reliable when they use data observed in the system they are trying to represent—or very similar systems. Even as we move forward with the application of models to make decisions in the short term, we must find ways to invest in monitoring eco-hydrometeorological data to improve our capacity for analysis and management in the medium and long term.

This publication aims to guide the assessment of water benefits resulting from interventions on natural infrastructure using hydrological models, considering the particular challenges presented in these exercises. It builds on valuable inputs, such as the *Guide to Selecting Ecosystem Services Models for Decision Making*, published by the World Resources Institute (Bullock & Ding 2018), and practical lessons learned from hydrological modeling exercises, mainly in the Andean region. We hope that this guide will be an accessible, useful resource for both decision makers and technical teams tasked with hydrological modeling, thus helping to bridge the science-policy gap and increasing the availability of relevant and robust information to support specific decisions.

Objectives of the Modeling Guide

The overall objective of this guide is to provide criteria for the development, selection, and use of models to quantify the expected hydrologic benefits of implementing NI projects, and more specifically, to:

- Guide the identification of policy questions and decision-making context of NI interventions to ensure that they can be appropriately modeled;
- Support the selection of appropriate hydrologic models within the decision-making context, based on the prioritized hydrologic ecosystem services, the interventions under consideration, and the availability of data, technical capabilities, and resources;
- Describe the characteristics of hydrologic models and the modeling process for the quantification of hydrologic benefits resulting from NI interventions;
- Guide the implementation of hydrologic models for the evaluation of NI, including descriptions and recommendations on model calibration, validation, sensitivity and uncertainty analyses, and interpretation of results for decision-making.

Who is this guide for?

This guide is intended for professionals who use or need to use hydrologic models, specifically:

1. Hydrologic modeling experts whose models are to be used in NI projects for water security.
2. Experts (with or without training in hydrology) in the design and evaluation of NI projects who need to model future interventions scenarios to evaluate the benefits of certain NI interventions.
3. Decision-makers who are not hydrology or modeling experts, but that require the results of hydrologic models to make and evaluate decisions related to NI projects.

Photo: Chancay-Lambayeque Watershed Resources Council

How is this guide organized?

The Modeling Guide begins with a conceptual framework that describes the basic concepts of NI and hydrologic ecosystem services and introduces hydrologic modeling, its aims, and applications. **Volume I** is targeted at professionals (with or without hydrology training) during the first stage of the modeling process, where the scope and context of modeling is defined, and which concludes with model selection useful for decision-making. **Volume II** presents technical details and is aimed at experts during the second stage of the hydrologic modeling process, where models are implemented, validated, and analyzed. Examples are included for each volume, and conclusions summarized at the end.

The knowledge shared in this guide will allow professionals who are not hydrology experts to engage in and oversee the modeling process and work more closely with modeling experts. It will also help experts see modeling as part of a collaborative problem solving process and to keep policy questions and project scope in mind.

The main content of this document is divided into two volumes, each with five steps:

Volume I. **Selection of the most appropriate hydrologic model for the context in question.**

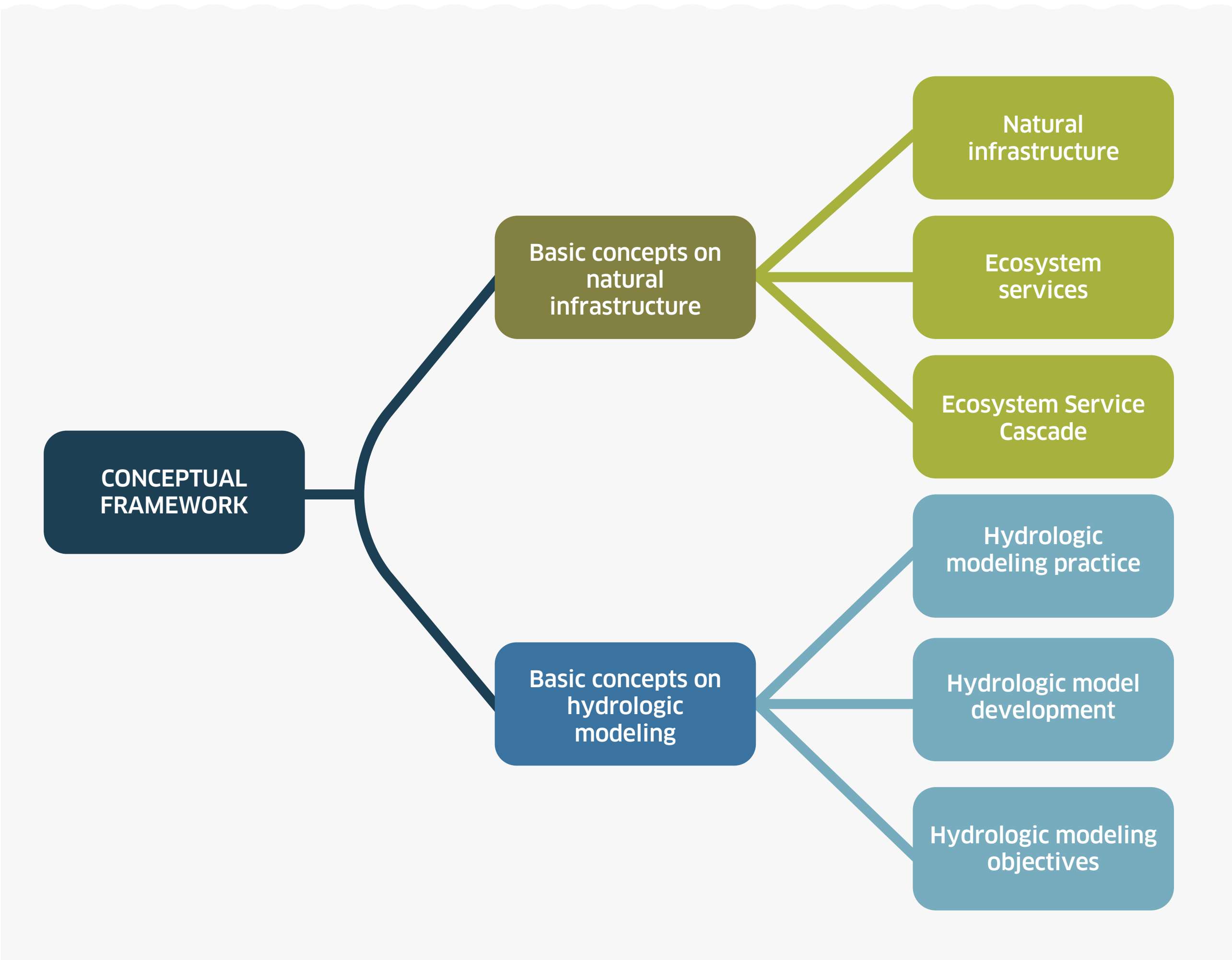
Step 1.	**Identify** policy questions and define their scope.
Step 2.	**Identify** priority hydrologic ecosystem services.
Step 3.	**Define** the NI interventions to be modeled.
Step 4.	**Recognize** the decision-making context, available resources, data, and capabilities.
Step 5.	**Evaluate** hydrologic models in this context and select the most appropriate ones.

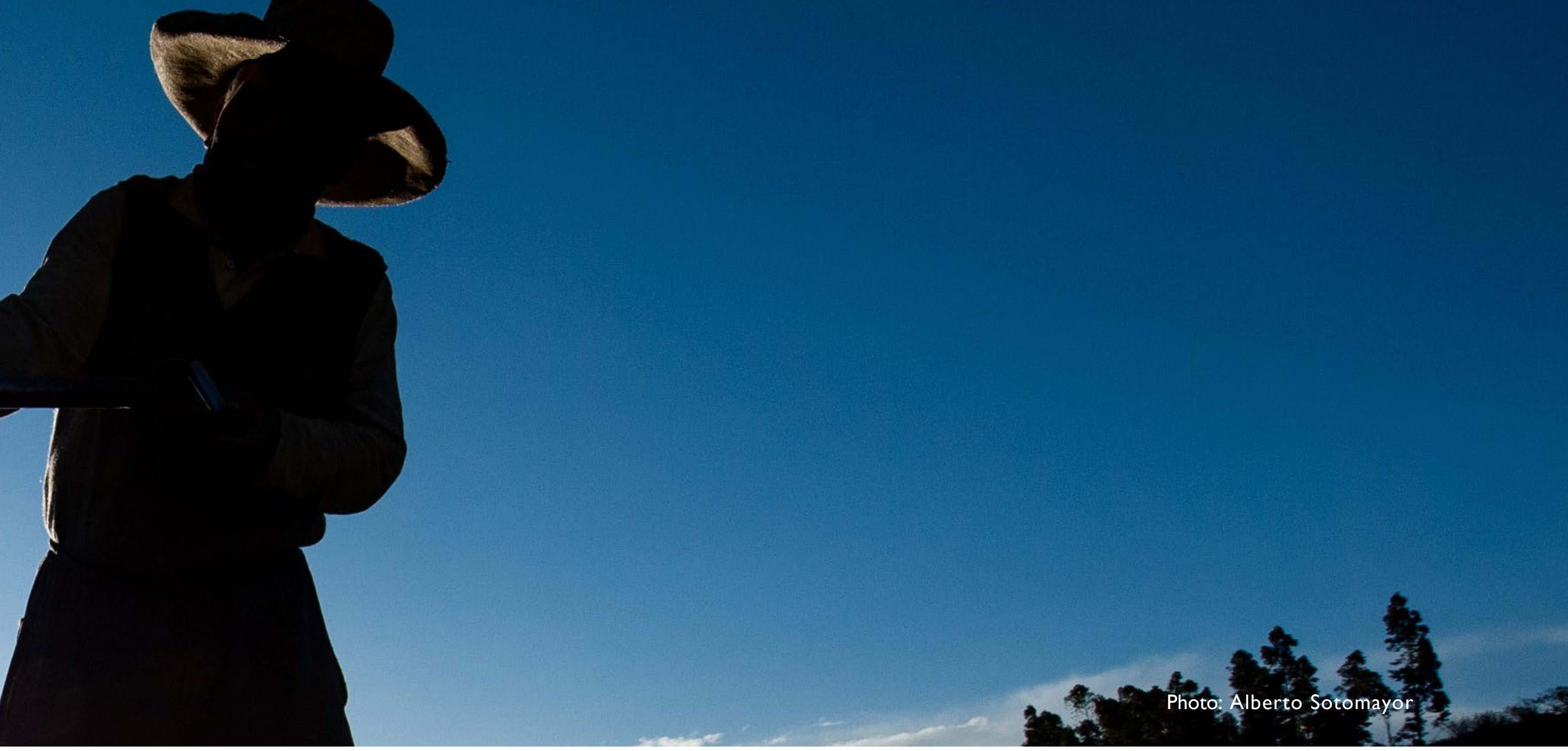
Volume II. **Details of model implementation, with guidelines and recommendations.**

Step 6.	**Generate** primary data or compile secondary data.
Step 7.	**Develop** modeling scenarios for the evaluation of future or alternative NI interventions.
Step 8.	**Calibrate and validate** the hydrologic model in a robust and ideally automated manner.
Step 9.	**Evaluate** model calculations and performance through sensitivity and uncertainty analyses.
Step 10.	**Interpret** the results with decision context in mind to generate relevant indicators for decision-making.

Photo: Ana Castañeda

I. Conceptual Framework





1.1 Natural Infrastructure

Natural Infrastructure (NI) is the network of natural spaces (water, soils, vegetation, subsoil, biodiversity, etc.) that preserves the values and functions of ecosystems and provide ecosystem services (e.g., climate and hydrologic regulation, carbon sequestration, air purification, etc.) (MINAM, 2019a). A network of natural spaces can be defined as a set of ecosystems recovered and conserved, either through tangible actions (ditches, canals, dikes, fences, etc.) or intangible actions (protection, exclusion, training, zoning, etc.). Its components may encompass natural ecosystems with varying degrees of disturbance, managed landscapes, and other multiuse spaces.

NI can be thought of as an actively managed natural system, generally focused on maintaining and/or enhancing benefits to the environment and human well-being, such as climate resilience for communities, improved water quality, flood control, among other co-benefits (IISD, 2018). Thus, management actions and interventions are implemented based on the priorities identified by communities and decision-makers and not randomly. It is important to highlight that the conservation and protection of NI is also a form of management.

In contrast, conventional ***gray infrastructure*** such as dams, tunnels, factories, and roads, are entirely designed and built by humans **(Figure 1)**. NI is closely related to ***green infrastructure***, a term that also includes systems that generate positive environmental outcomes, such as sustainable urban drainage or renewable energy (IISD, 2018). In many cases, NI can be more cost-effective and sustainable than gray infrastructure because it can serve multiple functions, provide co-benefits, and is more resilient to climate and environmental changes. The makeup, structure, and function of NI assets in water systems, and how they interact with gray infrastructure, determine the primary services and co-benefits produced.

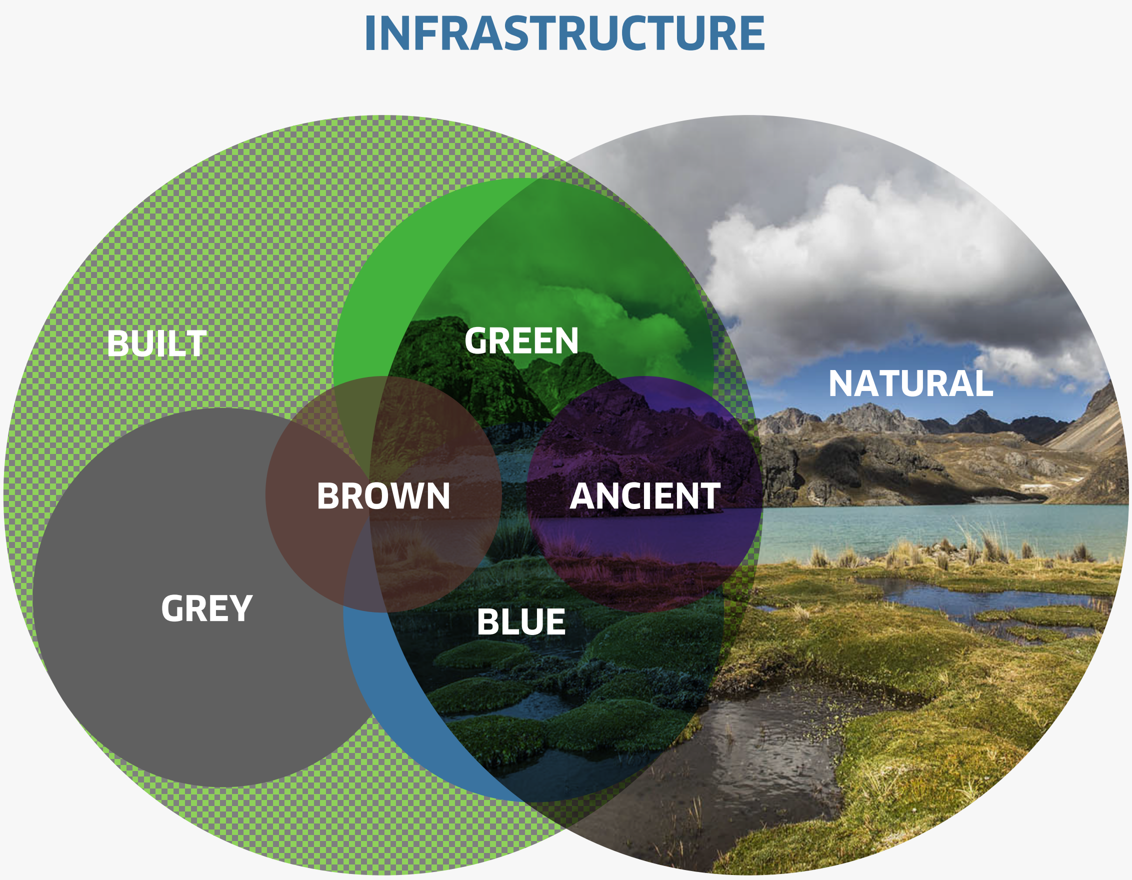


Figure 1. *Infrastructure for water security. Gray and green infrastructure are not antagonic and involve elements of one and the other. Natural infrastructure is a complement to human-built infrastructure. Source: Prepared by the authors.*

Another popular and related term is ***nature-based solutions*** (NBS) (IUCN, 2020a). According to the International Union for the Conservation of Nature (IUCN), NBS are actions to sustainably protect, manage, and restore natural or modified ecosystems that effectively and adaptively address societal challenges, while simultaneously providing benefits for human well-being and biodiversity (Cohen-Shacham et al., 2016). Nature-based solutions, particularly for water, have received special attention from UNESCO's partner agency for water (UN-Water). In 2018, World Water Development management defined NBS for water as a set of actions that are inspired and supported by nature and that use or mimic natural processes that contribute to improved water management (WWAP, 2018). While one area of conservation focuses on safeguarding biodiversity because of its irreplaceable value, NBS focuses on safeguarding society by following conservation rules and principles. While there is some overlap, not all conservation interventions are NBS (IUCN, 2020a, 2020b). Managing NI for water security is considered a NBS.

Nature-based solutions include interventions on NI interventions, such as wastewater purification and flood control (e.g., healthy wetlands); forest landscape restoration to reduce the impacts of extreme events; slope stabilization and water filtration; connection of rivers and aquifers to floodplains; preservation and protection of water resources (e.g., protected areas); establishment of flood diversions to reduce downstream impacts; slope cultivation to reduce erosion and water loss; location of riparian buffer zones to maintain water quality and reduce erosion; protection and restoration of mangroves, coastal wetlands, and dunes; conservation and restoration of wetlands; protection and restoration of reefs for coastal protection and habitat; hybrid solutions that interact with natural features to enhance water-related ecosystem services; rainwater harvesting with green roofs; infiltration enhancement for urban runoff reduction (e.g., permeable pavements); enhanced seepage and bioretention (e.g., urban green spaces); among other interventions.

"Managing natural infrastructure for water security is a nature-based solution."

Nature inherently provides humans with specific benefits, many of which are quantifiable and valuable in economic terms. These benefits are known as *ecosystem services*.

1.2 Hydrologic Ecosystem Services

Ecosystem services* (Figure 2)** are benefits derived by humans from ecosystems (MA, 2003; Mace, 2008). ***Hydrologic ecosystem services, are benefits produced by terrestrial ecosystems to water resources for humans (Brauman, 2007), including regulation of the hydrologic cycle and water yield, maintenance of water quality and aquifer recharge (i.e., ground water), among others. These services are distinct from aquatic ecosystem services, which are specifically produced by freshwater (aquatic) ecosystems (Brauman, 2015) and from marine ecosystem services, which are produced by saline water (marine) ecosystems and estuaries (Palumbi et al., 2008).

At the basin level, ecosystems influence water behavior through local climatic interactions, consumptive use by vegetation, modification of the land surface, modification of water quality, among other factors depicted in **Figure 3a** (Brauman et al., 2007). The hydrologic cycle is driven by solar energy and influenced by natural and built infrastructure. Water evaporated from the oceans, the surface of water bodies, or from the transpiration of vegetation in larger forests forms clouds, which precipitate as rain, hail, snow, or fog on the landscape and in the oceans. On land, water infiltrates into the soil or flows over the surface. Both surface water and subsurface water eventually discharge into the oceans. Evaporation from surface water and the oceans into the atmosphere completes the cycle.

The concept of hydrologic ecosystem services makes it possible to organize hydrologic processes in a basin (**Figure 3a**) according to their impacts on water for people (**Figure 3b**), such as provision of drinking water or recreational resources. Additionally, the ecosystem services conceptual framework can account for the impacts of the same ecosystem on a variety of other ecosystem services of interest, such as timber production or air purification (Brauman, 2015).

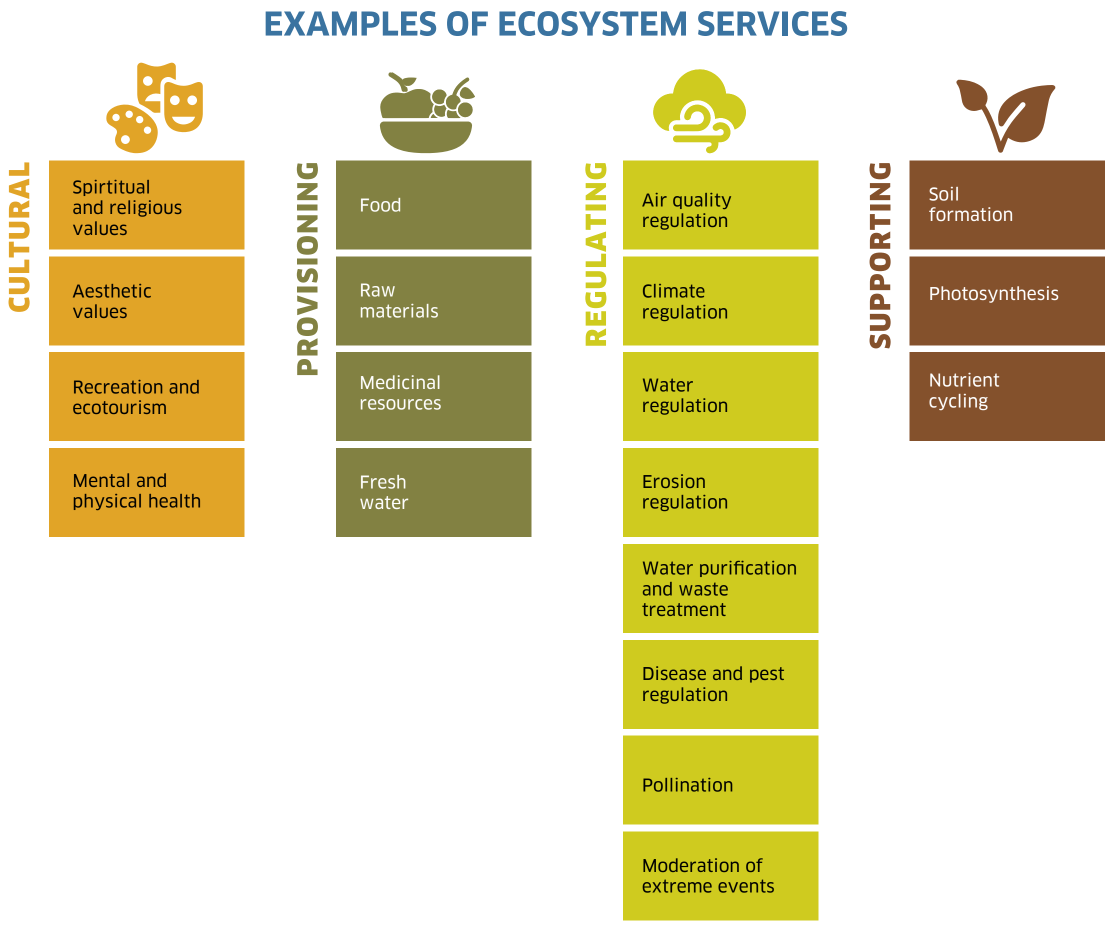


Figure 2. *Examples of ecosystem services. Adapted from WWF.*

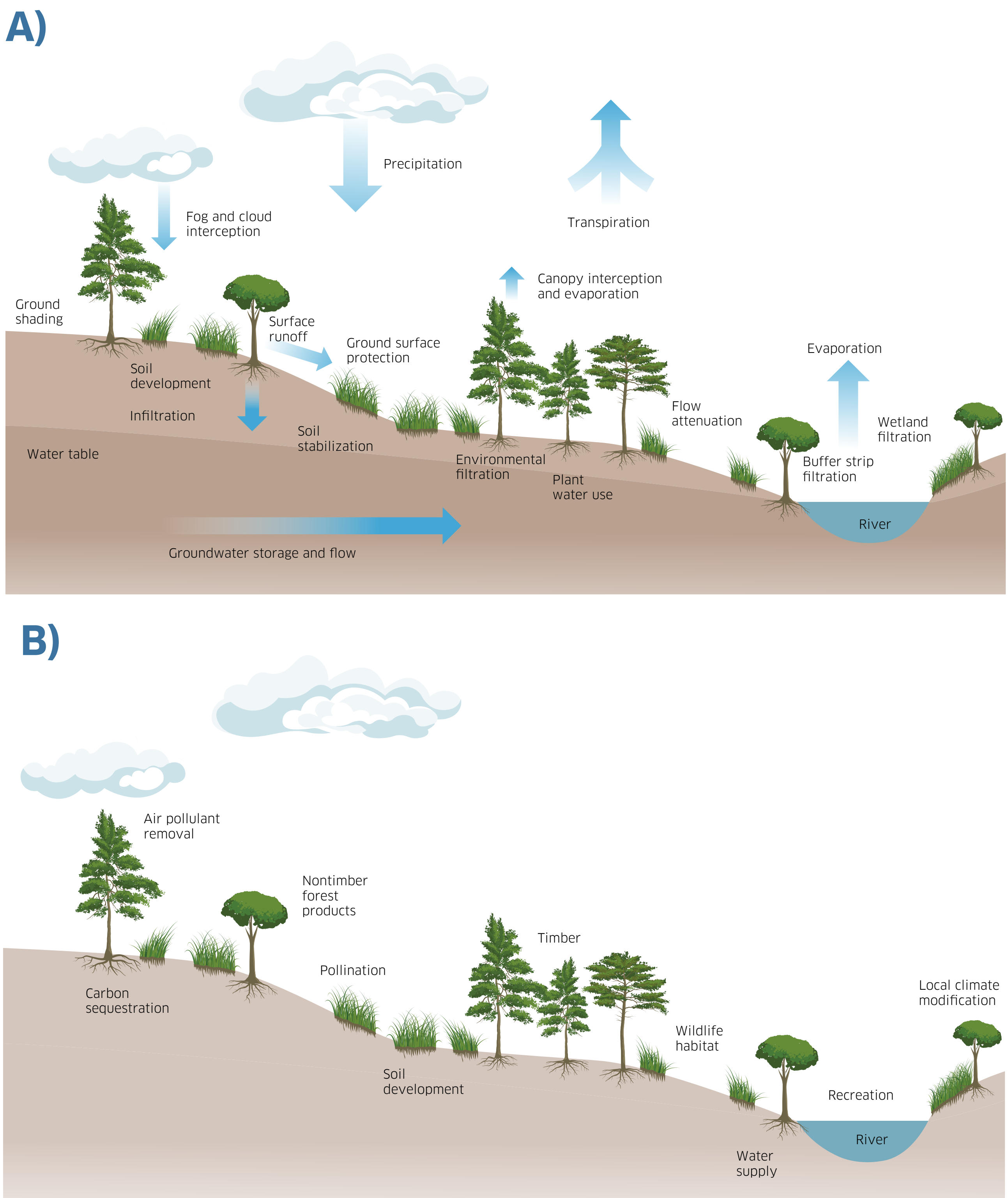


Figure 3. *a) Hydrologic processes; and b) ecosystem services in a watershed. Arrows indicate hydrologic processes and flows, with relative magnitude proportional to arrow width. Source: Brauman et al. (2007).*

Ecosystem services are usually divided into provisioning, regulating, cultural, and people-supporting services (MA, 2005). Water is commonly treated and discussed as a provisioning service, possibly because water is generally perceived as the product of a basin. However, ecosystems do not create water, they move and modify water flows. Therefore, it is more useful for research and water management to considering hydrologic ecosystem services as regulating services (Brauman, 2015).

> ***"Ecosystems do not create water, they move and modify water flows."***

There is a wide range of key ecosystem services, including:

Water quantity and quality	**Carbon capture and storage**
Mitigation of hydrologic risks (e.g., flood, drought)	**Food provision**
Erosion control	**Sustaining biodiversity**
Hydrologic regulation	**Scenic beauty, landscape or aesthetic value**

1.3 Ecosystem Service Cascade

It is essential to identify the beneficiaries of these services to transition from understanding the biophysical structure and processes (ecological, hydrologic, biological, abiotic, etc.) to identifying and quantifying the benefits generated by ecosystem services. Few studies satisfactorily develop this last requirement (Harrison-Atlas et al., 2016). The concept of the "***ecosystem services cascade"*** (CSE, **Figure 4**), is used in this guide. It connects ecosystem structures and processes with elements that impact human well-being to form a production chain or value chain (Haines-Young & Potschin, 2010; Potschin & Haines-Young, 2011). CSE comprises five steps organized into two domains: i) the environment; and ii) the socio-economic system. The first demonstrates that biophysical structures and processes of an ecosystem (e.g., forest habitat; basin where hydrologic processes occur) are required to fulfill ecosystem functions (e.g., slowing water flow; buffering flows), which in turn produce ecosystem services for the second domain (e.g., flood control; dry season water supply). Ecosystem services become the connection between the natural environment and the socio-economic system and generate goods and benefits for people (e.g., contribution to feeding the population through irrigation water; security of human settlements against floods), which have a tangible value (e.g., willingness to pay for ecosystem protection; retribution mechanisms for basin interventions). The CSE principle shows that to obtain a continuous flow of ecosystem services, people must protect, conserve, and restore the NI that sustains them by implementing policies that address anthropogenic threats.

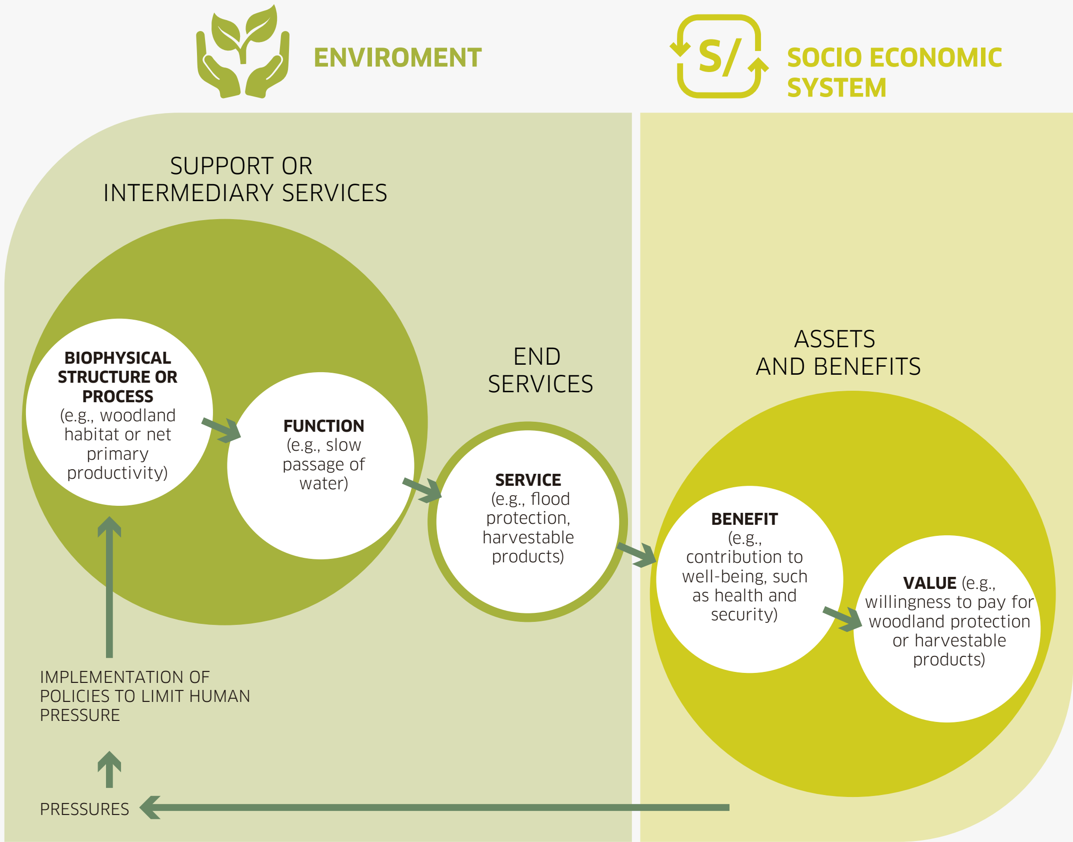


Figure 4. *Cascade of ecosystem services. Source: Potschin & Haines-Young (2011); Haines-Young & Potschin (2012). Free translation by the Ministry of Environment of Chile.*

Depending on the beneficiaries and their characteristics, ecosystem services will be used and valued differently. For example, for a company that provides drinking water, raw water is its raw material (the good produced by nature). For such a beneficiary, ecosystem services, such as total water yield and the maintenance of high base flows during the dry season, are a priority because they represent a greater quantity of raw material; similarly, services such as water quality regulation are important because they reduce drinking water treatment costs and, therefore, optimize the cost-efficiency of production. Other associated services, such as the maintenance of biodiversity or carbon sequestration, are desirable co-benefits, but are not necessarily priority services for a drinking water provider. For a forestry organization, timber production and carbon sequestration through increased plant biomass are priorities, although it is possible that increased tree cover may result in reduced water production (due to increased plant consumption). This impact on the ecosystem service of water yield becomes an externality but is not necessarily a priority for the forestry organization. The balance between interests, benefits, and externalities among different beneficiaries across a given geographic area must be considered in any analysis of ecosystem services.

The management of hydrologic ecosystem services generated by NI requires an understanding of the state and dynamics of the natural processes that control those services. These processes are sensitive to the effects of human interventions on NI, which include physical actions, policies, strategies, and practices linked to management decisions. Natural infrastructure interventions modify the ecosystem services they generate, positively or negatively, and hydrologic models can be used to simulate and analyze this effect.

1.4 The Practice of Hydrologic Modeling

Hydrologic modeling is a tool for representing the water cycle through simplifications and approximations of the actual hydrologic system. Models can be physically built (real) to represent a basin, like the Saywite stone or hydraulic models made in laboratories in academic and high investment settings. This guide focuses on simulation or computational models. As a model approximates the real system, its inputs and outputs are hydrologic variables, and equations with some physical or conceptual basis are used to represent processes and relationships between variables. Evaluating the hydrologic benefits of NI through modeling involves understanding and analyzing how a hydrologic system functions and forecasting its response to various management alternatives or scenarios.

While there are resources available on hydrologic modeling, they tend to focus on technical use of models, rather than a comprehensive look at how models can be applied to resource management and decision-making. This Guide to Hydrologic Modeling builds on the Guide to Selecting Ecosystem Service Models for Decision Making, by the World Resources Institute (Bullock & Ding, 2018), and takes it a step further by walking readers through a comprehensive process to use modeling to support policy- and decision-making and to evaluate and manage NI interventions.

Hydrological modeling **(Figure 5)** includes aspects from asking key policy questions and defining the decision-making context; identifying which types of hydrologic ecosystem services are relevant to the questions asked; deciding which NI interventions should be modeled; learning which hydrologic models are available to answer those questions and quantify the desired ecosystem services; selecting a hydrologic model that best responds to the needs and capacities identified for the decision-making context; collecting primary information or compiling secondary information; developing modeling scenarios, such as varying NI intervention options, using available information and knowledge; discussing difficulties and problems that may arise during the modeling process; to critically analyzing and objectively interpreting the model results and sources of uncertainty to inform decision-making related to NI interventions.

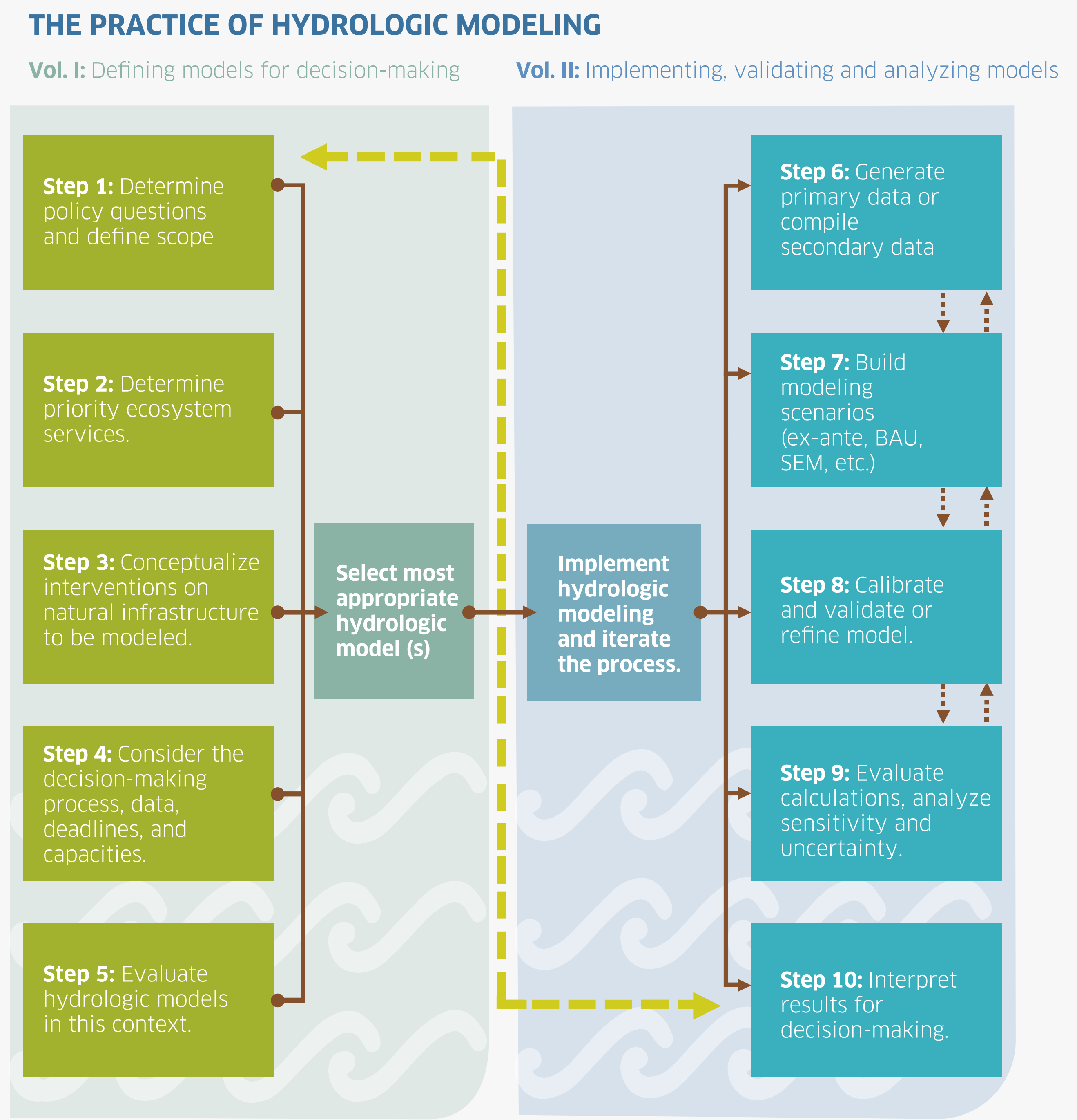


Figure 5. *Conceptual framework for hydrologic model selection and use. Adapted from Bullock & Ding (2018), World Resources Institute.*

"Greater complexity in a hydrologic model does not necessarily translate to better or more useful results for decision-making."

There are many hydrologic models available, and developers constantly update them to improve performance. Level of complexity, characteristics, and process representation vary and determine how useful a model will be to answering decision-making questions. For example, some models group an entire river basin into a single element that functions together and receives water inflows (e.g., rainfall), and represents processes using simple mathematical concepts (e.g., a linear relationship between reservoirs that discharge flow as a function of the current water volume in those reservoirs). Other more complex models use physics-based concepts to represent hydrologic processes and their controllers. For example, the effect of climatic conditions, water consumption by vegetation, how water moves across and into the landscape, etc. Some models group a unit of analysis into a single aggregate output, others generate results scattered over the study area. However, greater complexity in a hydrologic model does not necessarily translate to better or more useful results for decision-making

Selection of a hydrologic model should be guided by the nature of the questions posed, availability of data and information, technical capability to operate the model, and the relevance of outputs and their ability to effectively inform the decision-making context. It is necessary to recognize that hydrologic modeling does not provide all the answers to all existing policy questions. Hydrologic modeling for decision-making requires policy questions that are able to be answered or informed by a hydrologic model; it also requires recognizing which questions cannot be answered. This guide is focused on this process whereby decision-makers can select appropriate models, develop and analyze scenarios, and define and interpret the results. This guide will enable decision-makers to work with consulting teams or technical specialists to make the practice of hydrologic modeling more relevant to NI interventions for water security.


Photo: Carlos Alberto Vergara Manrique de Lara

1.5 Hydrologic Model Development

Different methodological approaches for developing models can be found in the literature but one of the most influential is presented in **Figure 6** (Beven, 2011). Here, the modeling process is divided into five steps to create a workflow that can be iterative, depending on the success of the final modeling step.

Step one is to define a ***perceptual model*** of hydrologic processes, which is understanding of a specific system, including the flows, complexity, and interactions between processes, and how it responds under certain conditions. This is a subjective process, as perceptions vary among practitioners. During this stage, observations play an important role because they inform the current state of processes in a specific area. Similarly, a team's experience and knowledge will play an important role and influence how well the modeller can conceptualize each of the processes. Someone with more experience across different branches of hydrology is likely to have a greater capacity for designing this step than someone with little experience. The experience gained from developing models in different locations and intimate knowledge of the study area are also invaluable, especially in complex systems where surface water and groundwater interactions are dominant. These insights are not yet mathematically formalized and need not necessarily be written down. The perceptual modeling step makes it possible to decide which processes will be the most important to represent in a model, which ones can be left out, and what level of detail is sufficient to simulate the dynamics of a system.

"The perceptual modeling step makes it possible to decide which processes will be very important to represent in a model, which ones can be left out, and what level of detail is sufficient to simulate the dynamics of a system."

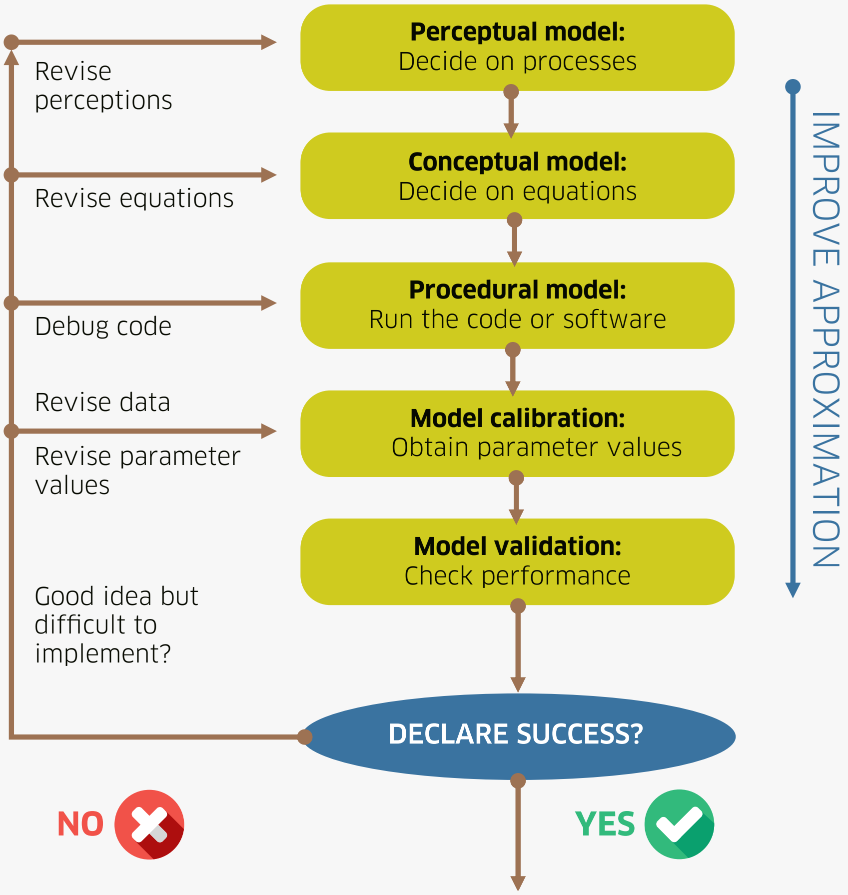


Figure 6. *Development of a hydrologic model. Adapted from Beven (2011).*

Insights from the perceptual model are then formalized in a ***conceptual model***, a representation of the relevant hydrologic processes in the form of equations. As such, these equations are a simplification of the perceptions and carry implicit assumptions about how the system works. The choice of equations must also be informed by the objectives of a particular study. For example, a bofedal (high Andean wetland) can be fed by surface water (runoff) and groundwater (groundwater discharge). To simulate only the buffer function of a bofedal, equations that simulate land and groundwater dynamics would be ignored in favor of those representing the storage capacity of the microtopography and the water holding capacity of the soil. However, to see how the existence of bofedales influences baseflow provision, components that simulate the interaction between the bofedal and groundwater would need to be added. Model complexity is also decided at this stage: the spatial disaggregation, the temporal resolution, and the level of complexity of the hydrologic processes. This will affect the computational cost and the predictive capability of the system. Complex models will require more simulation time, which, as will be discussed below, does not necessarily imply that they will produce a better prediction.

The next stage ***procedural modeling***, translates the equations of the conceptual model into code that can solve them. This stage involves developing the equation-solving code and running it to produce numerical results. Conceptual models often include a set of partial differential equations, often nonlinear, which require the use of numerical methods to solve them. In other words, the procedural model operationalizes the mathematical equations formulated in the previous stage. This stage requires great care to minimize errors and computational costs associated with the execution of the developed code.

After having ***code*** that implements the desired equations, model parameters must be ***calibrated,*** an adjustment process to check model outputs against observed data for each parameter. ***Parameters*** describe and summarize the properties of the area or flows in a watershed or hydrologic system, such as the values for hydraulic conductivity of land in a spatially distributed model, or a linear reservoir coefficient to simulate the water buffering capacity of a wetland in a conceptual model. Based on available observations, a priori estimates of the value (or range of values) of these parameters can be made. However, given the quantitative or qualitative uncertainties associated with these estimates, parameters must be adjusted to improve the performance of the model with respect to the observed data.

Evaluating the performance of a calibrated model is called ***validation***. With the calibrated model, one or more simulations are performed and compared to precipitation and flow observations. Based on the performance statistics, the system simulation can be rated as good, fair, or poor.

It should be noted that these validations are conditional validations (Beven & Young, 2013) because models are evaluated using observed data corresponding to a certain range of values, making it possible to know whether a model performs acceptably for a range of values, even for values outside the range. For example, a hydrologic model that incorporates the kinematic wave equation for surface water flow may be suitable for simulating slow flows and yield good validation results for a certain range of precipitation and flow values, being conditionally valid for these ranges. However, the same model may yield poor results for simulating peak flows because the kinematic wave equation is not a good representation of how extreme events behave in a system. A priori, we could suggest that shallow water equations are used instead of a kinematic wave approximation, and assume that the model validated for certain ranges will not be suitable for extreme events.

"A hydrologic model may be suitable for simulating slow flows and for a certain range of precipitation and flow values. However, the same model may yield poor results for simulating peak flows."

It is important to recognize that not all teams will have the resources to develop a model from scratch. Oftentimes, modeling will be limited to conceptualizing the scope and objectives and selecting an existing model that meets the needs of the study. In many cases, choosing a commonly used model can provide greater reliability to users and reviewers and lend credibility to a study.

Customized models for certain applications are essential to better approximate reality, but this is only a sound choice for those with the expertise and resources. This guide is therefore focused on the process of selecting existing hydrologic models that could meet the specific needs of a project at hand and on best practices for hydrologic modeling, regardless of the type selected.

1.6 Hydrologic Modeling Aims

In the context of NI management and intervention, hydrologic modeling is used for three purposes: **1)** to understand the processes of a hydrologic system; **2)** to make hydrologic predictions based on defined scenarios; and **3)** to support and strengthen the generation of data and information. Hydrologic models are used to test hypotheses regarding the functioning of the system, by trying to extrapolate a set of measurements and observe whether the hydrologic response mirrors a specific model structure (e.g. Buytaert & Beven, 2010; Beven 2019). They also simulate how the system would react to specific scenarios, such as climate change, land-use change, or the implementation of NI interventions; lastly, field data and observations are required for calibration and validation. The modeling exercise can also be used as a mechanism to optimize the design and operation of monitoring systems over the long term to generate more relevant data, and to develop more representative simulations based on observed data.

Understanding the processes of a hydrologic system through modeling has positioned models as ***"working hypotheses".*** Several factors determine how water moves through the hydrologic cycle. While knowledge and understanding of hydrology have advanced by leaps and bounds in recent decades, there are still uncertainties that cloud the understanding of those factors. The rise of computation and data science has facilitated the development of hydrologic simulation models that allow us to test various assumptions about how this works and to clarify many of the uncertainties by answering the question, "How does the system work?" This process goes hand in hand with on-site monitoring and robust data generation that require, to some extent or another, the use of models and equations to be assimilated.

Similarly, modeling is used for the analysis of scenarios that represent alternative realities based on data on a system's current state. Future conditions studied using models are commonly referred to as "projections" (Beven & Young, 2013), which are simulations produced for varying scenarios with hydrologic models that answer the question: What if? A scenario may include different modifications or alternatives in the hydrologic system; for example, changes in precipitation and temperature to reflect climate change, or changes in the internal properties of the system, such as a reduction in water infiltration rates into the surface to reflect the impacts of land-use change. For NI interventions, projections are used to predict and quantify potential hydrologic benefits before modifications are made to the actual system. This evaluation of impacts, before a specific intervention occurs, is often referred to as ***ex ante*** evaluación (e.g., CEC, 2009). which means before the fact – it is a term commonly used in economics and finance, where the results of a particular action or series of actions are anticipated. The opposite is ***ex post***, which means after the fact.

Ex ante scenario projections or simulations are useful for evaluating hydrologic behavior before changes occur. Observing actual system responses involves handling different input variables, such as climate or land conservation status. These types of changes are impossible to observe by monitoring because it would require control over the input variables (i.e., it's impossible to control how climate changes). Even if experiments can be carried out in the laboratory or field, extrapolating results to fit an active policy setting would not be sound or appropriate.

Photo: Natural Infrastructure for Water Security

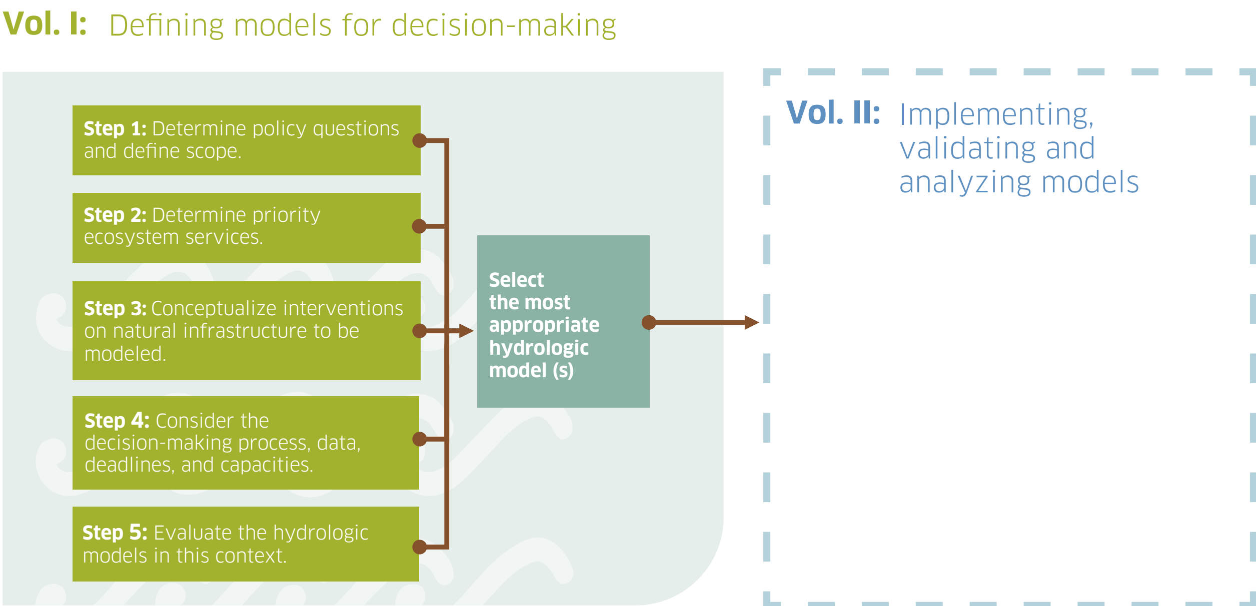
"Ex ante scenario projections or simulations are useful for evaluating hydrologic behavior before changes occur."

Lastly, hydrologic models allow improved quantifications of the hydrologic cycle by filling information gaps that are not reasonably measurable in time and space. Field experiments (in situ) suffer from spatial and temporal limitations. For example, it is not financially or physically possible to measure the flow at the outlets of all the basins in a system. Similarly, some hydrologic processes cannot be measured directly, requiring proxy variables to be selected. A clear example is groundwater movement, which is estimated by combining measurements of land hydraulic energy and conductivity in simple equations, such as Darcy's Law. This allows researchers to estimate the movement of water across a porous medium – in this case, the land surface of a basin – without directly measuring the flow in the field. Another example is generating estimated values for evapotranspiration. Direct measurements of evapotranspiration are expensive but models can be used to quantify this variable using other more easily measured meteorological variables, such as temperature, air humidity, solar radiation, and wind speed. Modeling makes it possible to support and strengthen ecological and hydrologic monitoring to optimize the use of limited resources, both during NI project design and monitoring for ex post evaluation.

VOLUME I

Photo: Will Espinoza

DEFINE MODELS FOR DECISION-MAKING



This section is aimed at NI project experts, as well as at decision-makers who are not specialized in hydrology or hydrologic modeling, but who require the results of hydrologic models to make and evaluate decisions on NI. It also calls upon hydrologic modeling experts to begin to see modeling as part of an integrated process and to participate in this process from designing policy questions to defining the scope of modeling within a given decision-making context. The content of this section draws heavily on the *Guide to Selecting Ecosystem Service Models for Decision Making: Lessons Learned from Sub-Saharan Africa*, produced by the World Resources Institute (Bullock & Ding, 2018).

This phase consists of five steps with twenty guiding questions. **Table I** shows the summary to be completed at the end of Volume I. The content of this volume guides the work team on the meaning and interpretation of each guiding question and the criteria to answer each question clearly and effectively. This volume concludes with hydrologic model selection or a modeling toolbox. The modeling toolbox concept emphasizes the possibility that a single hydrologic model will not answer all questions posed by a project, and that a set of hydrologic models might need to be considered.

> ***"The modeling toolbox concept emphasizes the possibility that a single hydrologic model will not answer all questions posed by a project, and that a set of hydrologic models might need to be considered."***

Table 1. *Guide for Steps 1-5 for the Selection of Hydrologic Models for Natural Infrastructure. Source: Prepared by the authors.*

STEPS	GUIDING QUESTIONS
STEP 1. Determine the policy questions and define their scope.	**Q1.** What is the policy question we want to address using the hydrologic model?
	Q2. How can we define the scope of this question so that we get relevant outputs from the model?
	Q3. What formats are required of the model outputs? (e.g., flow time series, maps of biophysical variables)
	Q4. What levels of accuracy and precision are needed to make a decision on this question?
STEP 2. Determine priority ecosystem services.	**Q5.** What ecosystem services are important to our decision making and at what geographic scale?
	Q6. What hydrologic processes and functions control priority hydrologic ecosystem services?
	P7. What hydrologic indicators can be used to obtain representative representations for these services?

STEPS	GUIDING QUESTIONS
STEP 3. Conceptualize the interventions in the natural infrastructure to be modeled.	**Q8.** What NI is important for our purpose?
	Q9. Which NI interventions are to be modeled?
	Q10. How will the functioning of the modeled interventions be conceptualized?
	Q11. How can the functioning of NI interventions be represented in hydrologic models?
STEP 4. Consider the decision-making context, data, and capabilities.	**Q12.** What is the intended use of the hydrologic model results?
	Q13. What is the timeframe to inform the decision in question?
	Q14. What data are available to address the policy questions?

STEPS	GUIDING QUESTIONS
STEP 4. Consider the decision-making context, data, and capabilities	**Q15.** Is the necessary data quality available for the preferred hydrologic models?
	Q16. Is there in-house technical capacity and prior experience to operate potentially relevant hydrologic models?
	Q17. If necessary, are there resources to finance external technical capacity to operate a selected model?
STEP 5. Evaluate the hydrologic model(s) in this context and select the most appropriate	**Q18.** Which hydrologic model meets the most criteria positively and how?
	Q19. What are the limitations and disadvantages of the selected hydrologic model?
	Q20. What other models are feasible alternatives?


Photo: Ana Castañeda

STEP 1. POLICY QUESTIONS AND THEIR SCOPE

The aim of this step is to determine the policy question(s) to define a scope for the modeling process. It is also necessary at this stage to define the geographic scale and identify the type of outputs desired (e.g., maps, data, indicators, targets), as well as the expected levels of accuracy and precision. The intended outcome of this step is to better understand the policy questions, translate them in a way that can be addressed by the models, and obtain the desired format, precision, and accuracy of model outputs to inform decision making. Some guiding questions for this step include:

- **Q1.** What is the policy question we want to address using the hydrologic model?
- **Q2.** How can we define the scope of this question so that we get relevant outputs from the model?
- **Q3.** What formats are required of the model outputs (e.g., flow time series, maps of biophysical variables)?
- **Q4.** What levels of accuracy and precision are needed to make a decision on this question?

Answering the first guiding question will help summarize the problem at hand and determine whether hydrologic modeling is necessary. The ***policy question***, is different from the hydrologic or modeling question because it has a broader scope than just the hydrologic domain. Policy, as such, is a deliberate system of guidelines to guide decisions and achieve rational outcomes on issues that impact a society or a country. The scope of this question could include elements of socioeconomic development and governance, environmental preservation, private company business, etc.

Some policy questions that may arise for decision-makers in this step include:

- Where should we direct NI restoration funds to generate the greatest impact?
- Will investing in NI preservation help those living in poverty?
- Can we save money on our drinking water supply by restoring a certain wetland?

Defining the scope of the question of interest is the intersection between policy and hydrologic modeling; as noted above, the policy domain is much broader than the hydrologic aspects of a project. By considering both, social and institutional interest can be translated into a modeling question or hypothesis, making it possible to use the model as a tool to support decision-making. Likewise, defining that scope clarifies the modeling objectives, sheds light on which model to select, and makes it possible to determine the hydrologic variables that can be modeled and addressed by the simulation tools.

The policy questions outlined above can be further refined into modeling questions as follows:

- What are the most effective areas for land cover or land-use changes that enhance specific ecosystem services (e.g., sediment control, flood control)?
- What are the economic benefits of maintaining the conservation status of ecosystems for those who depend on them for their livelihoods (e.g., subsistence agriculture, livestock, clean water)?
- What is the reduction in water treatment costs of cleaner water produced by the restored wetland?

"Defining the scope of the question of interest is the intersection between policy and hydrologic modeling."

A hydrologic model generally produces time series or GIS products (geographic information system, i.e., maps) of hydrologic variables as outputs, such as flow, baseflow, land moisture, sediment load, sediment concentration, chemical element concentrations, and pollutant loads, among others. These results can be considered ***"raw"*** and should be further organized and analyzed to obtain by-products that are usable to decision-making and communicating results. It is also necessary to reflect on the following: i) decision-maker interest and expertise, ii) whether products of biophysical variables or economic variables are required, iii) whether it is necessary to generate maps to visualize model outputs over an area of interest, and iv) whether model outputs need to be converted into another format for further use. Defining expected outputs clarifies the types of models and knowledge necessary to answer the policy question. For example, it may be necessary to select a model that works with GIS information and generates GIS outputs, or it may be necessary to have the technical capability to translate model outputs into maps. The format of the outputs is determined by the type of question defined for modeling and the team's ability to generate the outputs they will need (or if this technical work needs to be outsourced). For example, for the questions identified above, outputs are required in the form of maps to define the most effective locations for intervention; or outputs in economic terms, to determine the benefits to the population dependent on hydrologic ecosystem services, or outputs in biophysical terms, to determine the quantity and quality of water entering a treatment plant.

Lastly, it is necessary to define what is acceptable in terms of the accuracy and precision of the results. There is a fundamental difference between these two concepts that is addressed later in the paper **(Volume II, Step 6).** Generally, there is a trade-off between the degree of accuracy and precision and the complexity of the model and the time available to produce results. While more complex models tend to be more accurate, they require more data, resources, technical capacity, and time. Not all questions demand pinpoint accuracy, just as there are questions that are best answered not with one value, but with a range of possibilities. Moreover, it may require considerable resources to refine the estimates of the results. This may not be a good investment of time and resources if the benefits of such refinement are only marginal. It must also be recognized that models are approximations and inherently uncertain, making precision and accuracy relative. It is therefore important that practitioners and decision-makers take care not to let seemingly high precision in a modeling result lead to a false sense of certainty or exaggerated confidence in a decision without continuing to assess the realities in the field.

Some guiding questions include:

- Are very precise estimates of ecosystem services required to represent the current state of NI?
- Are very precise estimates required to compare and decide between different alternatives?
- Are moderately precise estimates required to estimate the future state or economic value of NI?
- Are moderately precise estimates required for the purpose of locating prioritized intervention areas for the generation of specific services?
- Are estimates of general trends, rather than exact values, required to guide policy development or public communication and awareness?

"Seemingly high precision in a modeling result can lead to a false sense of certainty or exaggerated confidence in a decision."


Photo: Natural Infrastructure for Water Security

STEP 2. PRIORITY ECOSYSTEM SERVICES

The purpose of this step is to define which ecosystem services are of interest in the relevant policy question and which should be prioritized during modeling. Identifying the ecosystem services of interest makes it easier to select an appropriate hydrologic model. This goes hand in hand with the delimitation of the geographic study area that will allow focused data collection efforts and support the precision of the expected model outputs. For example, the systematic review by Harrison-Atlas et al. (2016) includes a table of criteria for quantifying hydrologic ecosystem services for decision-making. The guiding questions are:

- **Q5.** What ecosystem services are important to our decision making and at what geographic scale?
- **Q6.** What hydrologic processes and functions control priority hydrologic ecosystem services?
- **Q7.** What hydrologic indicators can be used to obtain representative responses for these services?

The key hydrologic ecosystem services (**Table 2** to **Table 6**) for NI management are (Bonnesoeur et al., 2019):

Photo: Will Espinoza

Table 2. *Water Yield Ecosystem Service. Source: Prepared by the authors.*

1. WATER YIELD	
What is it?	The ability of a catchment to convert water inputs into water outputs. Water yield can be seen as the "factory" that produces water.
Who is it important to?	The total amount of water available in a catchment in rivers, streams, or wells throughout the year, without considering such aspects as regularity or seasonality, is important for water users that require large volumes and storage capacity (e.g., hydroelectric power plants, reservoirs for agricultural, domestic, or industrial use).
What elements or processes are involved?	High water yield depends on the following characteristics: • High precipitation or water inputs in the catchment • Vegetation with the capacity to capture mist and fog • Low water consumption by vegetation (low evapotranspiration)

Table 3. *Hydrologic Regulation Ecosystem Service. Source: Prepared by the authors..*

2. HYDROLOGIC REGULATION	
What is it?	The ability of a catchment to buffer the variability of water inputs for the purpose of generating a more homogeneous output. For example, attenuating flood flows and maintaining base flows during dry periods. Hydrologic regulation acts as a "sponge" in the catchment.
Who is it important for?	The preservation of baseflow is key for all water users to cope with water scarcity resulting from seasonality, climate variability, and climate change. This is particularly significant for those users who lack artificial storage capacity, such as small-scale farmers.
What elements or processes influence it?	Sound hydrologic regulation depends on the following characteristics: • Low intensity of precipitation reaching the soil • High water infiltration capacity in the soil • High water storage capacity within the land and subsoil

Table 4. *Water Quality Regulation Ecosystem Service. Source: Prepared by the authors..*

3. WATER QUALITY REGULATION	
What is it?	The ability of a catchment to improve the physical, chemical, and biological characteristics of water as it flows through the catchment, as well as its ability to reduce the concentration of pollutants. Water quality regulation can be understood as the "filter" of the catchment.
Who is it important for?	Water quality is significant for users who require specific or controlled physical, chemical, or biological characteristics. For example, drinking water utilities use treatment plants – the better the water quality, the lower the cost of treatment. Similarly, hydroelectric power plants require water with good physical characteristics to avoid clogging of reservoirs and deterioration of turbines during the power generation process. Other users, such as local municipalities, require good physical, chemical, and biological characteristics to be able to supply drinking water to their inhabitants.
What elements or processes influence it?	Sound water quality depends on the following natural characteristics: • A tight plant cover that protects the soil • A good soil structure and a low slope to prevent erosion • A low input of physical, chemical, or biological pollutants from human activities and, sometimes, natural sources

Table 5. *Extreme Events Regulation Ecosystem Service. Source: Prepared by the authors.*

4. EXTREME EVENTS REGULATION	
What is it?	The ability of a catchment to regulate extreme events, for example, intense precipitation, to control the magnitude, frequency, and duration of extreme events, such as floods or landslides.
Who is it important for?	Peak flow control is essential to reduce the effects of floods on people and activities located in flooding areas. Landslides are among the most destructive disasters in the Andes.
What elements or processes influence it?	Sound disaster control depends on the following main characteristics: • A low occurrence of extreme precipitation events • A good hydrologic regulation to mitigate peak flows • A good soil structure and a low slope that offers mechanical resistance to water and debris

Table 6. *Soil Conservation Ecosystem Service. Source: Prepared by the authors.*

5. SOIL CONSERVATION	
What is it?	The ability of a catchment to contain laminar erosion (diffuse erosion) and improve soil fertility, soil moisture, and plant production. Soil properties affect the provision of other ecosystem services, such as plant production, forage, crops, etc.
Who is it important for?	Soils in good condition benefit multiple sectors directly, such as the agricultural sector, which uses the soils for crops or pasture. Others benefit indirectly, for example, reservoirs that are prone to clogging benefit from reduced erosion.
What elements or processes influence it?	Many ecosystem properties and functions influence the condition of soils, their fertility, and plant production including: • A tight vegetation cover, which can reduce erosion and nutrient loss • A high soil moisture, to protect against climatic variations • A good soil structure, such as texture or high organic matter content

A set of **co-benefits** can result from ecosystem protection and restoration actions, such as biodiversity maintenance, air purification, carbon sequestration in land and biomass, scenic beauty, maintenance of cultural values, and maintenance and support of biochemical processes.

There are also potential ***trade-offs*** between ecosystem services. That is, sometimes increasing one service can result in the reduction of another. A common example is afforestation with exotic species. Forest plantations can help improve water regulation and carbon storage in the biomass; however, trees often consume more water than native herbaceous vegetation, which can reduce the total water yield. The introduction of exotic species can also result in reduced biodiversity. While afforestation on degraded soils can improve water infiltration into the land by root action, a trade-off exists between how much more water can infiltrate and be stored in the land because of tree root effects and how much water is consumed through evapotranspiration and internal tree storage **(Figure 7)**.

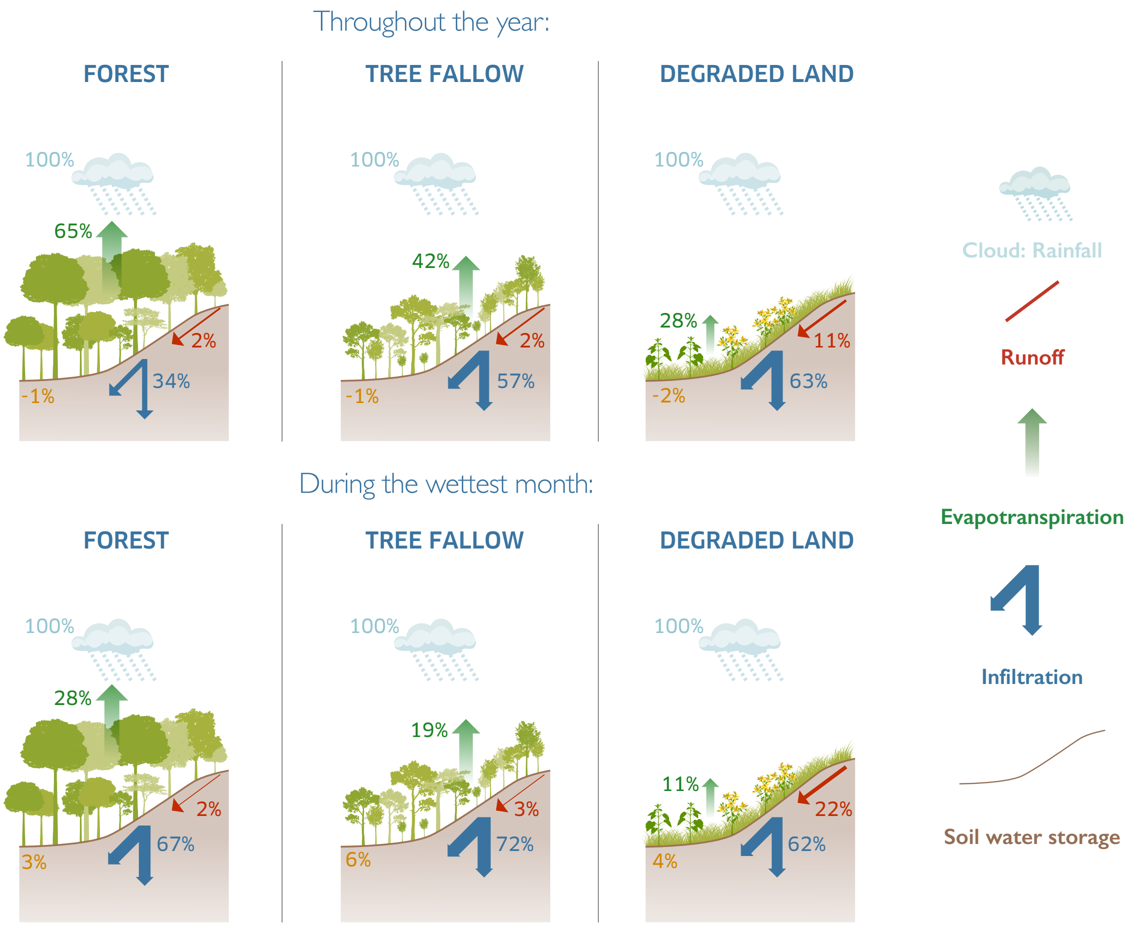


Figure 7. *Water balance as a function of precipitation. Adapted from van Meerveld et al. (2021).*

Co-benefits and trade-offs between ecosystem services need to be considered when making intervention decisions in the catchment's ecosystems. Only hydrologic models that can model priority ecosystem services and provide relevant quantitative information to inform decision-making should be considered for the question in hand. In general, one model will not be able to simulate all ecosystem services, and depending on the project, it may be necessary to use more than one. While multiple impacts on the territory should be considered, priority should be accorded to services that are of greatest interest and those that will be most impacted by interventions on the NI under consideration.

> ***"Co-benefits and trade-offs between ecosystem services need to be considered when making intervention decisions in ecosystems."***


Photo: Víctor Idrogo


Photo: Ana Castañeda

STEP 3. NATURAL INFRASTRUCTURE INTERVENTIONS TO BE MODELED

The aim of this step is to define the types of ***natural infrastructure interventions*** to be modeled and to compare this with available hydrologic models to determine whether and how they are capable of representing the interventions. The expected outcome of this step is a set of defined interventions and conceptual proposals to compare the available models and evaluate their usefulness to the decision context. The following guiding questions can be used:

- **Q8.** What NI is important for our purpose?
- **Q9.** Which NI interventions are to be modeled?
- **Q10.** How will the functioning of the modeled interventions be conceptualized?
- **Q11.** How can the functioning of NI interventions be represented in hydrologic models?

To model interventions on NI, hydrologic models use variables such as land use and land cover, water storage properties in the soil, surface interventions in the terrain (barriers, depressions, etc.), or groundwater recharge characteristics. The conceptualization of the performance of the interventions should be based on the variables and representations used by the hydrologic models to ensure their usefulness. This allows users to evaluate differences in the hydrologic response resulting from a change in these variables, as well as the effects of a variation in the complexity, accuracy, and sensitivity of the models. **Table 7** shows a set of NI interventions, the effects of which can be conceptualized to model them using simple processes.

Table 7. *List of Natural Infrastructure Interventions and Possible Evaluation-Modeling Approach. Source: Prepared by the authors.*

Intervention on natural infrastructure	Possible conceptual approach	Land use and land cover	Soil water storage characteristics	Structural interventions on the surface	Subsurface characteristics and flows
Creation of protection areas in the upper basin.	Reduce runoff velocity with different covers. Use land cover and land use map for preservation.	✔	✔	✔	
Installation of sustainable productive use systems, such as agroproductive, agroforestry, or silvopastoral systems.	Compare scenarios before and after a sustainable use project through land and vegetation characterization variables in intervened areas.	✔	✔	✔	
Purchase of land in areas of water interest for protection or restoration.	Support reduction of liquid and solid flows to protect areas.	✔	✔	✔	
Installation of forest fire prevention and control systems.	Quantify the impacts of burning on water services and compare to the impacts of a "natural" scenario.	✔	✔	✔	
Establishment or strengthening of a conservation area's surveillance system.	Compare with degraded conservation area scenario and establish its management.	✔	✔	✔	
Protection of springs and their access area.	Compare to a degraded or eliminated source. Determine area of contribution and area of influence of the water source.	✔	✔	✔	
Closure of natural pastures to prevent the entry of livestock.	Use land cover and land use map for conservation. Modify land water storage property variables (e.g., increase storage capacity to represent less soil compaction and less runoff).	✔	✔	✔	
Implementation of rotational grazing systems in natural pastures to eliminate overgrazing.	Modify land water storage property variables (e.g., create cyclical changes in storage capacity to represent presence or absence of livestock).	✔	✔	✔	
Sowing of natural pastures to (re) generate coverage.	Comparison of vegetated and bare ground cover.	✔	✔	✔	

Intervention in natural infrastructure	Possible conceptual approach	Land use and land cover	Soil water storage characteristics	Structural interventions on the surface	Subsurface characteristics and flows
Genetic improvement of livestock in areas outside protected ecosystems.	Change vegetation cover in ungrazed areas.	✔	✔	✔	
Switching from dairy cattle to South American camelid breeding.	Change vegetation cover with grazing and land properties to evaluate erosion reduction to represent the lesser impact of camelids vs. dairy cattle.	✔	✔	✔	
Planting of fodder oats or forage grasses in natural pasture areas.	Use land cover and land use map with cover changes in productive areas.	✔	✔	✔	
Wetland restoration.	Modify land water transport and storage property variables. Modify model processes to simulate changes in hydrology.	✔	✔	✔	
Elimination of the practice of extracting the organic layer of peatlands and bofedales.	Modify land carbon and organic matter content variables and their influence on hydrology.	✔	✔	✔	
Restoration of ancestral infiltration infrastructure.	Evaluate the effectiveness of infrastructure in storage. Use flow delay and residence time functions.	✔	✔	✔	✔
Construction of infiltration ditches.	Evaluate storage capacity and flow rate reduction; evaluate erosion control and runoff reduction capacity; evaluate water infiltration rates.	✔	✔	✔	✔
Construction of rustic storage or infiltration infrastructure.	Evaluate the effectiveness of storage infrastructure. Use flow retardation and residence time functions.	✔	✔	✔	✔
Afforestation with exotic species.	Determine changes in evapotranspiration, tree water consumption, and changes in water yield.	✔	✔	✔	

Intervention in natural infrastructure	Possible conceptual approach	Land use and land cover	Soil water storage characteristics	Structural interventions on the surface	Subsurface characteristics and flows
Afforestation of degraded or highly vulnerable areas.	Support the use of plant cover to recover degraded areas; use degradation maps; modify water infiltration variables in soils.	✓	✓	✓	
Recovery of native primary and secondary forests.	Evaluate current land cover and propose new land uses.	✓	✓	✓	✓
Recovery of terraces and pre-Hispanic terraces.	Propose recovery alternatives based on cover and infrastructure; evaluate the capacity for erosion control and runoff reduction; evaluate water infiltration rates.	✓	✓	✓	✓
Construction of dams in steep slope areas.	Evaluate infrastructures and covers for flow and sediment reduction and storage.	✓	✓	✓	✓
Construction of stone collars.	Evaluate infrastructures for flow reduction and storage.				✓
Construction of slow-forming terraces.	Evaluate infrastructures and covers in flow reduction and storage.	✓	✓	✓	✓
Construction of gabions to stabilize gullies.	Evaluate infrastructures for flow and sediment reduction and storage.				✓
Construction of drainage works for dirt roads.	Encourage speed reduction and road maintenance; evaluate suspended solids production and increase in sediment load and concentration.	✓	✓	✓	✓

Once the NI interventions have been identified, a modeling approach for evaluating their effectiveness has been conceptualized, and modeling proposals have been made, the team can select hydrologic models that may be suitable for producing the expected simulations. These models must be able to satisfactorily address the policy question posed, quantify the priority hydrologic ecosystem services, and simulate the proposed NI interventions. Next, these models need to be evaluated on how well they can support the specific institutional context and align with the available resources, data, and capacities.

STEP 4. DECISION-MAKING CONTEXT, RESOURCES, AND CAPABILITIES

The first aim of this step is to determine decision-maker ***interests,*** initial ***resources*** and ***capabilities,*** and the **time** available to contribute to the decision-making process. The expected outcome is to gain an understanding of the decision-making context to generate relevant information, consider the specific resources and capabilities of the decision-making institution, and establish the time frame for making those management decisions. The following guiding questions can be used:

- **Q12.** What is the intended use of the hydrologic model results?
- **Q13.** What is the timeframe to inform the decision in question?

The intended uses of modeling results depend directly on the interests of decision-makers, project formulators, and policy makers involved in a NI intervention. Such uses may include developing future scenarios, conducting policy evaluations, valuing ecosystem services for a region, determining the best actions to maintain or enhance priority ecosystem services, calculating trade-offs between different ecosystem services, evaluating the benefits and co-benefits of interventions on NI, and so on. This context must be communicated to and understood by those who will carry out the modeling exercise, regardless of their professional background, to ensure that hydrologic modeling serves an integral interest beyond hydrologic theory.

Another significant factor is urgency –whether there is a decision-making deadline and how soon modeling results are needed. The time required to obtain model results depends on model complexity, data availability, and the technical capabilities of the team to run the model. If a quick decision is required, time becomes an important constraint that influences model selection. Some common time windows include, in less than three months, within three to six months, six months to one year, or after one year.

There is an interrelationship and trade-off between data availability, local resources and capabilities, and time available. More complex models generally provide more accurate predictions and have a wider range of uses, but they can also be more difficult to use and understand, require more and better input data and resource availability, as well as technical modeling skills and expertise. A complex model may not be the best option in a context where a quick decision must be made with scarce data and resources. Grayson & Bloschl (2000) illustrate how the ***availability of information*** dictates the choices made regarding ***model complexity*** and its ***predictive capacity (Figure 8).***

"A complex model may not be the best option in a context where a quick decision must be made with scarce data and resources."

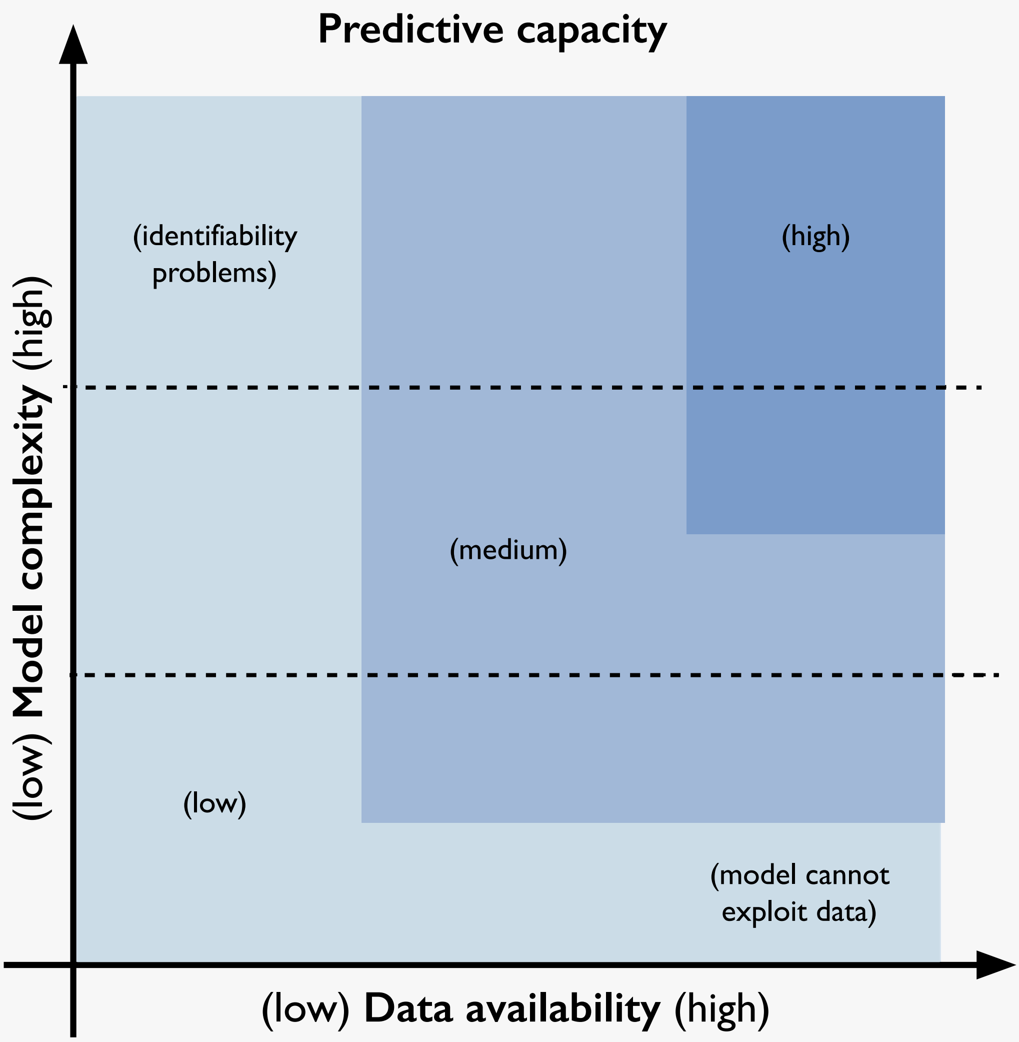


Figure 8. *Relationship between Data Availability, Model Complexity, and Predictive Capacity. Adapted from: Grayson & Bloschl (2000).*

The second objective of this step is to evaluate the availability of data and resources, as well as the technical capacity of the team, to determine the feasibility of implementing the relevant hydrologic models and tools. This will help determine how complex a model should be. The expected outcome is an inventory of available data for the defined modeling area that addresses the policy question of interest, and an evaluation of internal resources and technical capacity. It is necessary to characterize whether in-house capacity is high, medium, or low, for the purposes of operating selected hydrologic models. If in-house technical capacity is low or nonexistent, it needs to be determined whether there are resources to hire an expert team to build and run the model within the decision-making deadline. Some guiding questions include:

- **Q14.** What data are available to address the policy question?
- **Q15.** Is the necessary data quality available for the preferred hydrologic models?
- **Q16.** Is there in-house technical capacity and prior experience to operate potentially relevant hydrologic models?
- **Q17.** If necessary, are there resources to finance external technical experts to operate a selected model?

An efficient approach to evaluating data availability is to examine the input data requirements of different models and compare them with existing or readily available data. It is not unusual to have gaps in data availability. The simplest version of a hydrologic model requires a precipitation time series as the main input of water into a basin to simulate its conversion into a flow time series. In the absence of precipitation data for the basin (primary data), extrapolations from nearby precipitation and weather stations (secondary data) would have to be used. If there are no representative data available during the time period of interest, remote sensing can be used as a source of secondary data. At present, satellite products are available that can help address the scarcity of data in the territory (Kummerow et al., 1998; Tapley et al., 2011; Entekhabi et al., 2010; Hou et al., 2014; Tang et al., 2020).

In other scenarios, it is possible to have a wider range of relevant local data to use as inputs, such as a high-quality local dataset or local values for ecosystem services. Some more complex models require a digital elevation model (DEM) and a land cover or land use map, particularly if spatial analysis of NI interventions are needed. A flow time series is also often preferred or required to calibrate the parameters of hydrologic models. It tends to be less important for more complex models that employ physics-based equations. Some models automatically use information from global databases, while others require manual input of measured data. Secondary data available in global datasets can be used for several necessary variables (e.g., WorldClim: Fick et al., 2017), but local data have been shown to generate more accurate modeled estimates (e.g., Ochoa-Tocachi et al., 2016a & 2018a; Redhead et al., 2016). Sometimes, primary data may need to be generated through field experiments or hydrologic monitoring systems to support modeling or improve the model portfolio under consideration.

"Secondary data available in global datasets can be used for a number of necessary variables, but local data have been shown to generate more accurate modeled estimates."

It is also necessary to evaluate whether in-house technical capacity exists to operate hydrologic models. Local technical expertise may be small or restricted by availability of personnel or time. The simplest models can be operated by technical staff with limited hydrologic expertise, but there is a risk of higher uncertainties in results. Some models can be operated with online tools and require moderate technical skills. If there is no in-house technical capacity to collect data and operate models, the team must determine whether there are resources to hire experts to complete the analysis. If not, it is important to clarify the budget available and to engage with organizations and individuals who can provide that expertise and technical support.

More complex models require a high level of technical capacity. If in-house modeling expertise is not available, but there is some local data and sufficient resources to hire an external expert team, more complex models might be more feasible. If in-house technical capacity is moderate and a wide range of data is available, models that are moderately complex and provide user support may be an option. In these cases, capacity building of in-house technical staff should be considered as an efficient way to improve modeling results. In rare cases, the design, construction, codification, and operation of custom-made models are the best way to support a project, which requires high-level in-house knowledge that understands the decision-making context, interests, and available resources (e.g., Ochoa-Tocachi et al., 2019a).

Finally, it is advisable for decision-makers to consider more long-term efforts, such as regular model validation and replicating modeling runs to create more robust results or improve upon the model as more and/or better-quality data (and, potentially, better models) become available. With this comprehensive planning perspective, it is necessary to evaluate whether it would be more cost-effective to have a strong in-house team with continuous training, or whether it is sufficient to rely on an external expert team that is available on an ad hoc basis for short periods.

"Capacity building of in-house technical staff should be considered as an efficient way to improve modeling results."


Photo: Ana Castañeda

STEP 5. EVALUATION OF AVAILABLE HYDROLOGIC MODELS

Volume 1 of the modeling practice concludes with the **selection** of the hydrologic model (or modeling toolbox) to help address the policy question of interest (Step 1) by modeling the priority hydrologic ecosystem services (Step 2), generated or affected by the proposed natural infrastructure interventions (Step 3), in the identified decision-making context, using available data, resources, and technical capabilities (Step 4). The objective of Step 5 is to compare the results of steps 1 to 4 and select the most appropriate hydrologic model.

To guide decision makers in this selection, key concepts about hydrologic modeling theory and a short list comparing popular models are presented below. The expected results of this step are to provide a better understanding of the technical specificities of each model and to refine the number of modeling tools that can be used in the decision-making context. The goal is for decision-makers to be able to select one or two of the most appropriate hydrologic models to address their policy questions after performing steps 1 through 4 of Volume 1 of this guide, which will help them balance technical capacity, data availability, and financial resources to deliver results in time to inform their decisions. The guiding questions to be answered in this step are:

- **Q18.** Which hydrologic model responds positively to the highest number of criteria and how?
- **Q19.** What are the limitations and disadvantages of the selected hydrologic model?
- **Q20.** What other models are feasible alternatives?

Climate variability, the diversity of ecosystems, and the variety of questions to be answered at varying spatial and temporal scales encourage flexibility and multiple hydrologic modeling approaches to answer specific problems. There are many options, and even hydrologic professionals may struggle to choose between one model or another. Sometimes, this choice will be based on qualitative criteria, such as using a popular model because the institution has previous success with it. It is common to observe the cognitive bias of the ***"golden hammer"***, in modeling, which refers to excessive dependence on a tool, technology, or paradigm that one has familiarity with or receives exaggerated praise. This bias can be summed up by the phrase: "When the only tool you have is a hammer, every problem looks like a nail". This makes it particularly important to make an objective choice based on the objectives of the study, properly identify the critical processes to be represented in the modeling, and factor in the available information and technical capabilities at the time of modeling.

"When the only tool you have is a hammer, every problem looks like a nail."

Types of Hydrologic Models

Choosing an appropriate model requires researching what is available and whether it will meet project needs. Several classifications of hydrologic models are available in the literature (e.g., Singh, 1995; Refsgaard, 1996; Solomatine & Wagener, 2011). In this guide, hydrologic models are divided according to: a) how hydrologic processes are classified; b) the type of spatial disaggregation; and c) how randomness is introduced in the modeling.

Based on the level of description of hydrologic processes, models are classified as: i) ***empirical***; ii) ***conceptual***; and iii) ***physically based***. Empirical models, also called **"data-driven"** or **"black-box"** models, are developed without explicit consideration of physical processes (Resfgaard, 1996). Common examples are models using statistical or machine learning algorithms, such as artificial neural networks (ANN) or support vector machines (SVM). Conceptual models characterize system performance through simplified representations of hydrologic processes using simple parametric relationships. These models use reservoir parameterizations as the main element to represent the storage capacity of the hydrologic system components and simplified flows to fill or empty these reservoirs (Solomatine & Wagener, 2011). Finally, physics-based models, also called process-based or **"white box"** models, make use of more detailed and rigorous representations. They are generally made up of equations based on laws of conservation of mass, momentum, and energy. Physics-based models are the most complex and typically require significant computational resources, as well as a highly skilled and experienced operator.

Similarly, models can be classified depending on their spatial disaggregation in: i) ***aggregated (or lumped)***; ii) ***semi-distributed***; and iii) ***distributed*** models (**Figure 7**). This influences information requirements at different scales, ranging from summary information (such as basin area) to very detailed information (such as spatial maps of elevation, soils or vegetation cover), and model outputs will also depend on these scales. Aggregate models treat each study catchment as if it were a unit, and the variables and parameters used represent average values for the entire catchment. Semidistributed models discretize a catchment into subunits, which are called hydrologic response units (HRUs) in some models, based on the concept that homogeneous areas (e.g., with similar slopes, similar land cover, and similar soils) behave in hydrologicly similar ways. Finally, distributed models discretize the entire basin into regularly spaced subunits, e.g., using the same raster cell size as the digital elevation model map. By using distributed model spatial variations of hydrologic variables can be taken into account and distributed results can be obtained. In general, these models are the most complex and require significant computational resources and a high technical capacity for their operation.

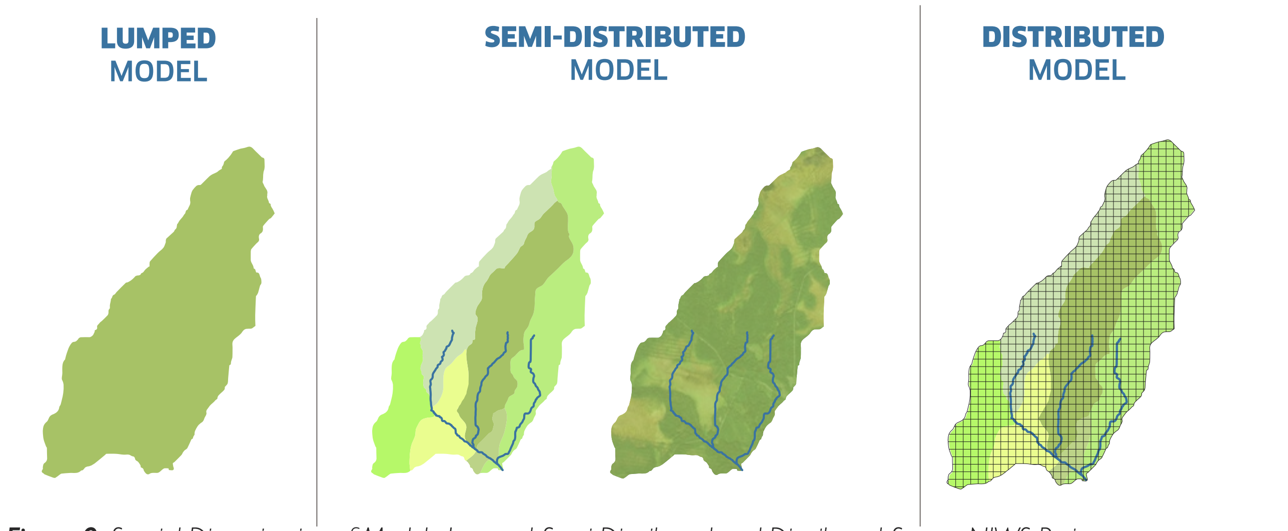


Figure 9. *Spatial Discretization of Models: Lumped, Semi-Distributed, and Distributed. Source: NIWS Project*

Models are classified according to the randomness they can introduce, as: i) **deterministic**; and ii) **stochastic**. In a deterministic model, the same inputs will invariably produce the same outputs, and not represent any randomness or uncertainty principles, meaning that if a simulation is repeated using a given set of model parameter values and input variables, the same results will always be obtained; likewise, if the input data are varied or the parameters are slightly modified, the results will be different. This principle is used for the calibration of model parameters. However, the ability to generate stochastic (i.e., random) behavior in different components of the model is sometimes required. Stochastic models allow for randomness or uncertainty in the outputs, which may be considered in the input variables, boundary conditions, or model parameters (Beven, 2011). This is useful in scenarios and processes that are intrinsically random or uncertain, such as precipitation or weather predictions (Beven, 2011).

In practice, model classification is somewhat less strict. Some physically based models incorporate empirical models; for example, many groundwater modeling packages, such as Modflow or Feflow, are described as physically based models. Yet these models use Darcy's Law, which was derived empirically through laboratory experiments (Darcy, 1856). Likewise, a model can have extensions or configurations that give it the capacity to be stochastic and to be operated deterministically.

It should be noted that no one type of model is better than another – they simply have different applications. At first glance, physically based models may appear to reproduce reality more faithfully because they approximate the laws accepted by the scientific community. However, their high parameterization and high data demand lead to modeling and calibration problems, such as nonlinearity, equifinality, and overparameterization (Beven, 2011). On the other hand, the spatial features of aggregated or semi-distributed models make it impossible to evaluate scenarios that require high spatial resolution (disaggregation of locations. i.e., high detail). For example, evaluating the impact of NI interventions for flood control may require a distributed model if one wants to estimate how NI might reduce floodable areas in a micro-catchment for a given design storm. The original siting of a portfolio of NI interventions may require a disaggregated look at the basin to evaluate where they should be placed to efficiently maximize benefits. However, it is not always possible, or necessary, to have a spatial model, either because of limitations in available data, computational resources, or technical capabilities, or because the expected results do not require considering their spatial distribution (e.g., estimating the total water yield of a basin or its water balance).

"No one type of model is better than another – they simply have different applications."

Popular Hydrologic Models

Table 8 lists hydrologic models and their main features. The list includes different approaches to water cycle and ecosystem analysis to represent the varying possible objectives of NI interventions. **Table 9** shows the characteristics of the hydrologic models considered. **Table 10** combines an overview of hydrologic modeling tools into functions that address the intended use of results, output format and modeled natural infrastructure.

Table 8. *Characteristics of Relevant and Frequently Used Hydrologic Models. Source: Prepared by the authors..*

Hydrologic Model	Description	Required Information	Natural Infrastructure Component	Model Outputs
SWAT **(Arnold et al., 2012)**	Soil and Water Assessment Tool. A semi-distributed model for land and water evaluation. It was developed to predict the impact of management practices on the generation of flows, sediments, and agricultural chemicals. It is a continuous model that requires information on climate, land properties, topography, vegetation, and land management.	Topography distributed in digital elevation model (DEM); land use and land cover map (LULC); land biophysical table; land type map (texture and structure); daily climatological information on precipitation; maximum and minimum temperature.	Soil use and vegetation cover, surface intervention, and subsurface storage capacity characteristics.	Daily, monthly, and annual time series of runoff sheets; infiltration (storage and subsurface supply); evaporation; and amount of sediment transported.
KINEROS **(Woolhiser et al., 1990)**	KINematic runoff and EROSion model. A semi-distributed, event-driven model based on the processes of interception, infiltration, runoff, and erosion. The basin is represented by a cascade of planes and channels. The equations are solved using finite difference techniques.	Topography distributed in digital elevation model (DEM); land use and land cover map (LULC); land biophysical table; land type map (texture and structure); channel granulometry; and precipitation hyetogram.	Soil use and vegetation cover, surface intervention, and subsurface storage capacity.	Time series of flow, infiltration sheet, concentration, erosion, deposition, and sediment production in the basin.
TOPMODEL **(Beven & Kirkby, 1979; Buytaert & Beven, 2010)**	TOPographic index MODEL. A semi-distributed hydrologic model based on the concept of the topographic index. It is an indicator of the susceptibility of certain areas of the basin to become completely saturated and is therefore based on the mechanism of runoff due to excess saturation.	Topography distributed in digital elevation model (DEM); precipitation and evapotranspiration information at different time scales; and physical parameters of the subsoil.	Intervention on the surface, and characteristics of the subsurface storage capacity.	Time series of runoff sheet, time series at event, day, month and year scales.

Hydrologic Model	Description	Required Information	Natural Infrastructure Component	Model Outputs
HEC-HMS & HEC–GeoHMS **(Scharffenberg, 2016; Fleming & Doan, 2000)**	Hydrologic Engineering Center - Hydrologic Modeling System. Together with the HEC-GeoHMS preprocessing interface, it is designed to simulate complete hydrologic processes of basins, including many traditional analysis procedures for events (abstraction, unit hydrograph, hydrologic transit), as well as continuous simulation (evapotranspiration, snowmelt, moisture content).	Topography distributed on digital elevation model (DEM); land use and land cover map (LULC); land biophysical table; land type map (texture and structure); precipitation hyetogram for event scale. For continuous scale, subsurface physical parameters are required.	Soil use and vegetatiton cover, surface intervention, and subsurface storage capacity characteristics.	Time series of flow rate for a precipitation event or continuous scale of precipitation and infiltration losses.
MODFLOW **(Harbaugh, 2005)**	Finite difference groundwater model. The structure consists of a main program and a series of highly independent subroutines. The subroutines are grouped into packages, each dealing with a specific feature of the hydrologic system to be simulated, such as river flow.	Depending on the purpose of the study, it may require topographic information, sources and sinks (rivers, drains, lakes, wells, etc.) and hydrogeological properties. It performs simulations in static and transient regime.	Land use and vegetation cover, surface and subsurface intervention, and subsurface storage capacity characteristics.	Water balance results between zones (rivers, wells, lakes, areas of interest).
WEAP **(Sieber & Purkey, 2015)**	Water Evaluation and Planning. WEAP is more of a modeling software platform than a model itself, mainly for water planning and distribution. It also performs hydrologic modeling at continuous scale.	Area, subsurface storage characteristics information, monthly precipitation, and monthly average temperature.	Subsoil storage capacity characteristics.	Water sheet and monthly flow.
RS MINERVE **(García Hernández et al. 2014, 2020)**	Routing System Modélisation des Intempéries de Nature Extrême dans le Rhône Valaisan et leurs Effets. Modeling software originally developed for flood forecasting and water resource management in mountain regions. Today it offers a wide range of applications, including a water balance with management objects, such as hydraulic infrastructure. RS MINERVE covers different hydrologic models: GSM, SOCONT, SAC-SMA, GR4J and HBV.	Topology of rivers, confluence points, and sub-basin areas with altitude and topography information distributed in elevation bands, daily or monthly climatology of precipitation and temperature, optionally ETP. RS MINERVE builds a model from three inputs in vector format: rivers, junctions, and subbasins, plus a climatological database of P, T, and ETP (the latter can be automatic, from a global database or a manual dataset).	Characteristics of subsurface storage capacity.	Flow at different time scales, including all components of the water balance (e.g., water consumption, water infrastructure).

Hydrologic Model	Description	Required Information	Natural Infrastructure Component	Model Outputs
HBV **(Bergström, 1992, 1995)**	Hydrologiska Byråns Vattenbalansavdelning. An aggregate conceptual model that considers water and energy balance. It consists of four modules: snowmelt and snow accumulation, land moisture and effective precipitation, evapotranspiration, and runoff estimation).	Area, river profile, subsurface storage characteristics, precipitation, and temperature information on a continuous scale.	Subsoil storage capacity characteristics.	Flow rate at different time scales.
LUTZ SCHOLZ **(Scholz, 1980)**	An aggregate conceptual model for monthly-scale flow forecasting using a model that combines the water balance for the average year with a Markovian process for flow generation.	Area, precipitation, runoff coefficient, ETP, effective precipitation, monthly water balance, periods of the hydrologic cycle, basin retention, observed flow.	Subsoil storage capacity characteristics.	Time series of monthly mean flow.
InVEST **(Sharp et al., 2018)**	InVEST is a toolbox for exploring how changes in ecosystems can lead to changes in ecosystem services that benefit people. The toolbox has three modules for hydrologic ecosystem services (annual and seasonal water yield, sediment delivery ratiom, and nutrient delivery ratio).	Topography distributed in digital elevation model (DEM), land cover and land use map (LULC), biophysical table, land type map, monthly precipitation map, ETP, and precipitation events by month. For the sediment delivery ratio (SDR) module, a DEM, precipitation erosivity index, land erodibility, LULC map, biophysical table, and basins are required.	Soil use and vegetation cover and subsoil storage capacity characteristics.	Map of curve number, instantaneous flow, base flow. For the SDR module, sediment yield maps and sediment delivery ratio as main outputs.

Hydrologic Model	Description	Required Information	Natural Infrastructure Component	Model Outputs
FONAG 2.1 by ATUK **(Ochoa–Tocachi et al., 2019b, 2020a, 2020b)**	Distributed model developed for water funds, based on the concept of hydrologic regulation hydrozones; distributed and cumulative water balance for water yield; and pollutant and sediment transport, based on source and transport coefficients. The water balance implemented by the model considers water production by natural infrastructure, as well as uses and flow returns from human activities.	Topography distributed in digital elevation model (DEM); land use and land cover map (LULC); hydrozone correspondence table; table of abstractions and anthropic flow returns; table of evapotranspiration coefficients, hydrologic regulation and compound hydrozone; climatic raster maps (precipitation, temperature, evapotranspiration); points of interest for outputs.	Land use and vegetation cover, surface intervention, and characteristics of natural infrastructure interventions. Interventions can be categorized as independent hydrozones with their own characteristics of water yield, hydrologic regulation capacity, or compound transport.	Raster maps distributed on a continuous monthly scale and monthly interannual averages of flow and concentration of compounds. Continuous monthly time series of precipitation, flow, concentration of compounds, at the points of interest. Distributed water stress map. Calibration and sensitivity analysis results.
iMHEA REGIONALIZATION MODEL **(Ochoa–Tocachi et al., 2016b)**	Paired catchment regionalization of the Regional Initiative for Hydrologic Monitoring of Andean Ecosystems (iMHEA). It is an aggregated empirical model that correlates the biophysical properties of the catchments and their climatic characteristics to estimate values of hydrologic indicators and their uncertainty ranges. It uses multivariate linear regressions based on data from iMHEA catchments in the tropical Andes.	Aggregate characteristics of the catchment of interest in the categories of shape, drainage, elevation, topography and slope, soils and geology, meteorology, precipitation intensity, land cover, and land use.	Interventions should be translated into the values of the biophysical characteristics of the catchment: shape, drainage, topography and slope, soils and geology, land cover, and land use.	Hydrologic indicators estimated in average value and uncertainty range at %95 confidence. Indicators in the categories of water yield, hydrologic regulation, water balance, flow magnitude, frequency, duration, temporality, and rate of change.

Hydrologic Model	Description	Required Information	Natural Infrastructure Component	Model Outputs
CUBHIC 2.0 **(Foster et al., 2020; Ochoa–Tocachi et al., 2022)**	The CUBHIC (Quantification of Hydrologic Benefits of Catchment Interventions) methodology is a set of hydrologic models that aim to perform a rapid evaluation of the water quantity and quality benefits of interventions in natural infrastructure.	Site-specific data, such as soil and vegetation characteristics; PISCO or field-measured climatic data of precipitation and temperatures (maxima, minima and averages); curve numbers based on observable site characteristics.	CUBHIC methodologies have been developed for 6 types of interventions in natural infrastructure: 1. High Andean grassland conservation and restoration; 2. Infiltration trenches; 3. Forest protection and restoration; 4. Restoration and protection of wetlands; 5. Qochas (permeable micro-reservoirs) reservoirs); and 6. Construction and restoration of amunas (ancient infiltration canals).	Each methodology includes a downloadable spreadsheet to calculate NI intervention benefits (Microsoft Excel), as well as a list of information needed to apply it. The results are generated in the same Excel sheet.

Table 9. *Summary of Hydrologic Model Characteristics Under Consideration. Source: Prepared by the authors.*

Hydrologic Model	Model outputs			Spatial scale			Temporal scale		Access	
	Water quantity	Water quality	Thematic map	Lumped	Semi-distributed	Distributed	Continuous	Event	Open	Commercial
SWAT	✓	✓	✓		✓		✓		✓	
KINEROS	✓	✓	✓		✓			✓	✓	
TOPMODEL	✓		✓		✓		✓		✓	
HEC–HMS	✓	✓	✓	✓	✓	✓	✓	✓	✓	
MODFLOW	✓	✓	✓			✓	✓		✓	✓
WEAP	✓			✓			✓		✓	✓
RS MINERVE	✓		✓		✓		✓		✓	
HBV	✓			✓			✓		✓	
LUTZ SCHOLZ	✓			✓			✓		✓	
InVEST	✓	✓	✓			✓	✓		✓	
FONAG 2.1	✓	✓	✓			✓	✓		✓	
iMHEA	✓			✓				✓	✓	
CUBHIC	✓	✓		✓			✓		✓	

Table 10. *Identification of Hydrologic Models by Capacity to Model Natural Infrastructure Interventions. Source: Prepared by the authors.*

Hydrologic Model	Compare scenarios	Natural Infrastructure Intervention Modeling			
		Land use and land cover	Soil water storage characteristics	Structural interventions on the surface	Subsurface characteristics and flows
SWAT	Yes	Yes	Yes	Yes	No
KINEROS	Yes	Yes	Yes	Yes	No
TOPMODEL	Yes	*	Yes	No	No
HEC–HMS	Yes	Yes	Yes	Yes	No
MODFLOW	Yes	Yes	Yes	Yes	Yes
WEAP	Yes	*	Yes	Yes	No
RS MINERVE	Yes	*	Yes	Yes	No
HBV	Yes	No	Yes	No	No
LUTZ SCHOLZ	Yes	No	Yes	No	No
InVEST	Yes	Yes	Yes	Yes	No
FONAG 2.1	Yes	Yes	Yes	Yes	No
iMHEA	Yes	Yes	*	No	*
CUBHIC 2.0	Yes	Yes	Yes	*	Yes

*Not by default. Requires preprocessing or an adapted modeling strategy.

Selecting the Most Appropriate Model

Guiding questions from **Table 1**, answered by following the five previous steps, can be used to select the most appropriate hydrologic model (or modeling toolbox) for the given policy question(s) and decision context. A case study from Peru is outlined below to demonstrate the process.

Illustrative Application

The Tambo-Ilo-Moquegua Basin, in southern Peru (Moquegua, Arequipa, and Puno), is one of the basins prioritized in the Natural Infrastructure for Water Security project (**Figure 10**). It comprises two hydrographic units of the Pacific slope: the Tambo Basin and the Ilo-Moquegua Basin, covering a total area of 16,492 square kilometers (km^2), with 126,715 inhabitants. The Tambo-Ilo-Moquegua basin has a marked altitudinal gradient from the high Andean zone (5,665 meters above mean sea level (masl)) to its mouth at the Pacific Ocean (29 masl). The basin encompasses diverse ecosystems, mainly: dry puna grasslands (34%), Andean scrubland (26%), coastal desert (22%) and periglacial zone (15%) (MINAM, 2019b).

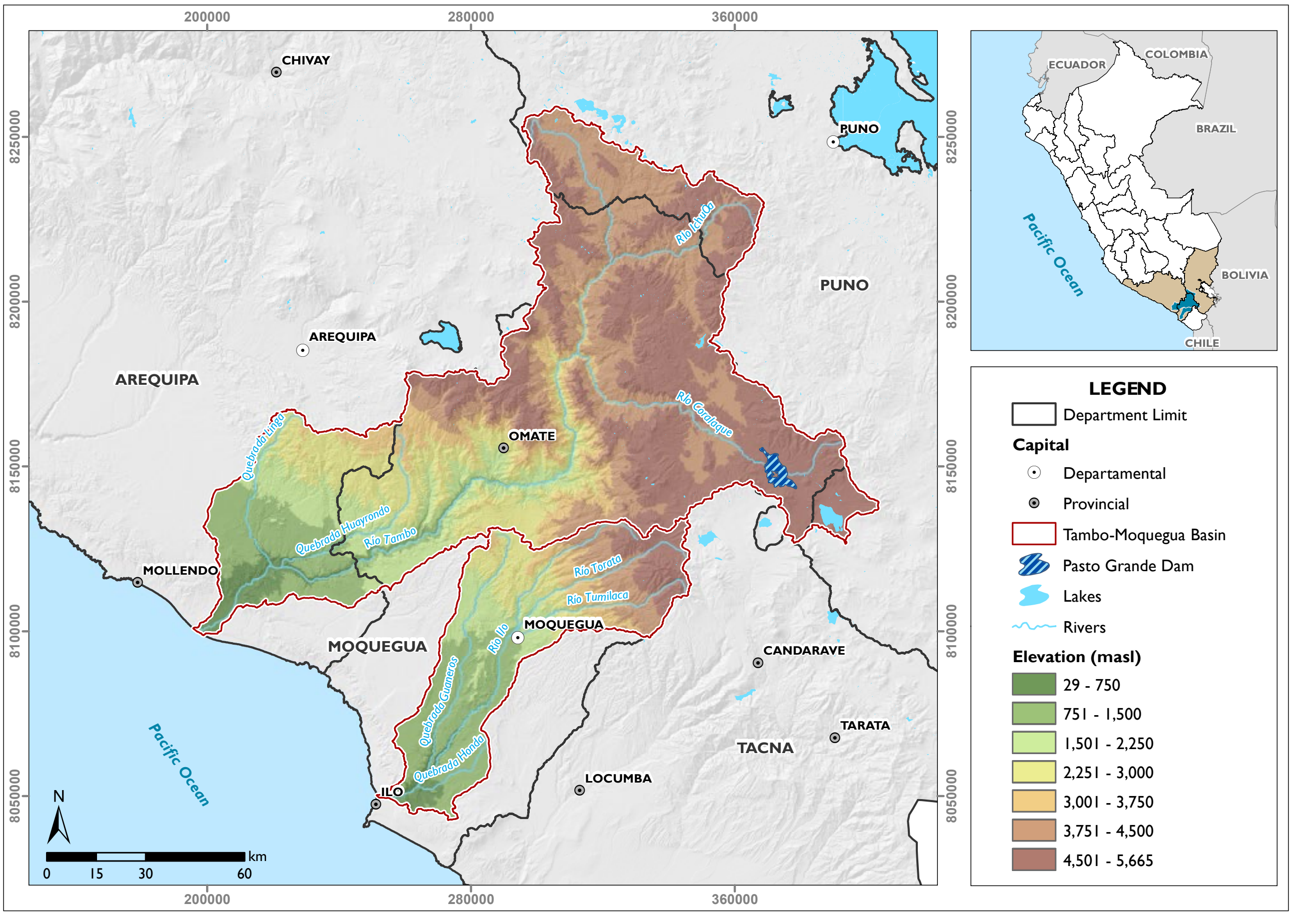


Figure 10. *Characterization of the Tambo-Ilo-Moquegua Basin. Source: Prepared by the authors.*

There is a marked difference in water availability between the two sub-basins. Water supply problems in the Ilo-Moquegua Basin led to the construction of the Pasto Grande Dam (Tambo basin), which supplies agricultural and residential users in the cities of Moquegua and Ilo. However, the dam is currently at 68% capacity, the dry season has worsened, and the bofedales have been degraded due to intensive alpaca grazing, which reduces their hydrologic regulation capacity. Heavy metals and algae have also been detected, and have a serious impact on water quality (SUNASS, 2019).

The main sources of water supply in the Ilo-Moquegua Basin are located farther upsteam. Subsurface and surface catchments in the Tumilaca River sub-basin supply 44% of the population of Moquegua (SUNASS, 2019). However, land use is impacting water security. Mining projects (Southern and Quellaveco) in the upper part of the basin have been impacting vital ecosystems, such as bofedales, and influencing river regimes and environmental services. This is compounded by the loss of vegetation cover due to agricultural and urban expansion (SUNASS, 2019).

Given these challenges, the Tambo-Ilo-Moquegua Basin is a complex case for integrated water resource management, where multiple interests and stakeholders converge (population, mining sector, agriculture, etc.). However, it is also a favorable scenario to coordinate and align efforts to recover and preserve NI, with a goal of laying the foundation for sustainable and water-safe development. **Table 11** shows the application of the five-step model selection process to this scenario.

Table 11. *Application of Steps 1-5 for the Selection of Hydrologic Models for Natural Infrastructure to the Tambo-Moquegua Basin. Source: Prepared by the authors.*

STEPS	GUIDING QUESTIONS	
STEP 1. Determine the policy questions and define their scope	**Q1. What is the policy question we want to address using the hydrologic model?**	How do we improve the water supply for the city of Moquegua, mainly in the Tumilaca, Alto Ilo-Moquegua sub-basins, as well as in the Pasto Grande Reservoir?
	Q2. How can we define the scope of this question so that we get relevant outputs from the model?	Can NI interventions support high flows in rainy seasons to: i) Increase available flows during the dry season; and. ii) Reduce sediment load in rivers and dams
	Q3. What formats (e.g., flow time series, maps of biophysical variables) are required of the model outputs?	Projected scenarios with and without intervention: time series of flow and sediment load at high temporal resolutions (at a minimum, daily); spatially disaggregated maps of erosion and sediment sources.
	Q4. What levels of accuracy and precision are needed to make a decision on this question?	We expect to obtain results of the impacts of the interventions in at least 5% of the baseline. The average flow value (baseline) at the Santa Rosa station is 30 m^3/s; therefore, the precision defined here as acceptable (5%) would be 1.5 m^3/s.

STEPS		GUIDING QUESTIONS
STEP 2. Determine priority ecosystem services	**Q5. What ecosystem services are important for our decision making and at what geographic scale?**	i) Hydrologic regulation (flood attenuation and dry season flow). ii) Annual erosion control. Scale: Tumilaca and Alto Ilo-Moquegua sub-basins; and Pasto Grande Reservoir.
	Q6. What hydrologic processes and functions control priority hydrologic ecosystem services?	Soil water storage in grasslands and bofedales is one of the major drivers of the water regulation service. Bare soil is more vulnerable to erosion, which leads to more sediment in the watercourses.
	Q7. What hydrologic indicators can be used to obtain representative responses for these services?	i) Accumulated volume of water during the dry season. ii) Accumulated sediment load on an annual scale.

STEPS		GUIDING QUESTIONS
STEP 3. Conceptualize the interventions in the natural infrastructure to be modeled	**Q8. What natural infrastructure is important for our purpose?**	The priority NI are grasslands and bofedales.
	Q9. What are the interventions on the natural infrastructure in the field to be modeled?	The interventions considered are: i) Conservation and protection of grasslands and bofedales; and ii) Recovery of bare soil areas using vegetation cover..
	Q10. How is the performance of the interventions sought to be modeled conceptualized?	Conservation: maintain the soil's water storage capacity, as well as the soil's water infiltration capacity. Recovery of vegetation cover: reduce erosion and the effect of precipitation intensity on the soil.
	Q11. How can the performance of interventions on natural infrastructure be represented in hydrologic models?	Changes in the land cover and land use map, or ecosystem types. Changes in soil characteristics (water storage, hydraulic conductivity, soil depth, infiltration). Ability to consider different types of erosion.

STEPS	GUIDING QUESTIONS	
STEP 4. Consider the decision-making context, data and capabilities.	**Q12.** What is the intended use of the hydrologic model results?	Exploratory exercise to decide if and where it is necessary to allocate a budget for investment projects for NI interventions at the pre-feasibility level.
	Q13. What is the timeframe to inform the decision in question?	Modeling is required within a maximum of two weeks.
	Q14. What data are available to address the policy question in question?	Available data: ALOS-QALSAR DEM, at a 12.5 meter (m) resolution; MINAM ecosystem map (2019b); commissioned land cover and land use map, at a 30 m resolution; global soil maps (FAO); daily PISCO climate data (precipitation and temperature); daily SENAMHI in situ climate data; daily flow data for calibration (ANA).
	Q15. Is the required data quality available for the preferred hydrologic models?	The quality of data available is suitable for use in a model such as SWAT, KINEROS, or InVEST.
	Q16. Is there in-house technical capacity and prior experience to operate potentially relevant hydrologic models?	There is technical capacity within the NIWS project and previous experience in the identified models.
	Q17. If necessary, are there resources to finance external technical capacity to operate a selected model?	Not required.

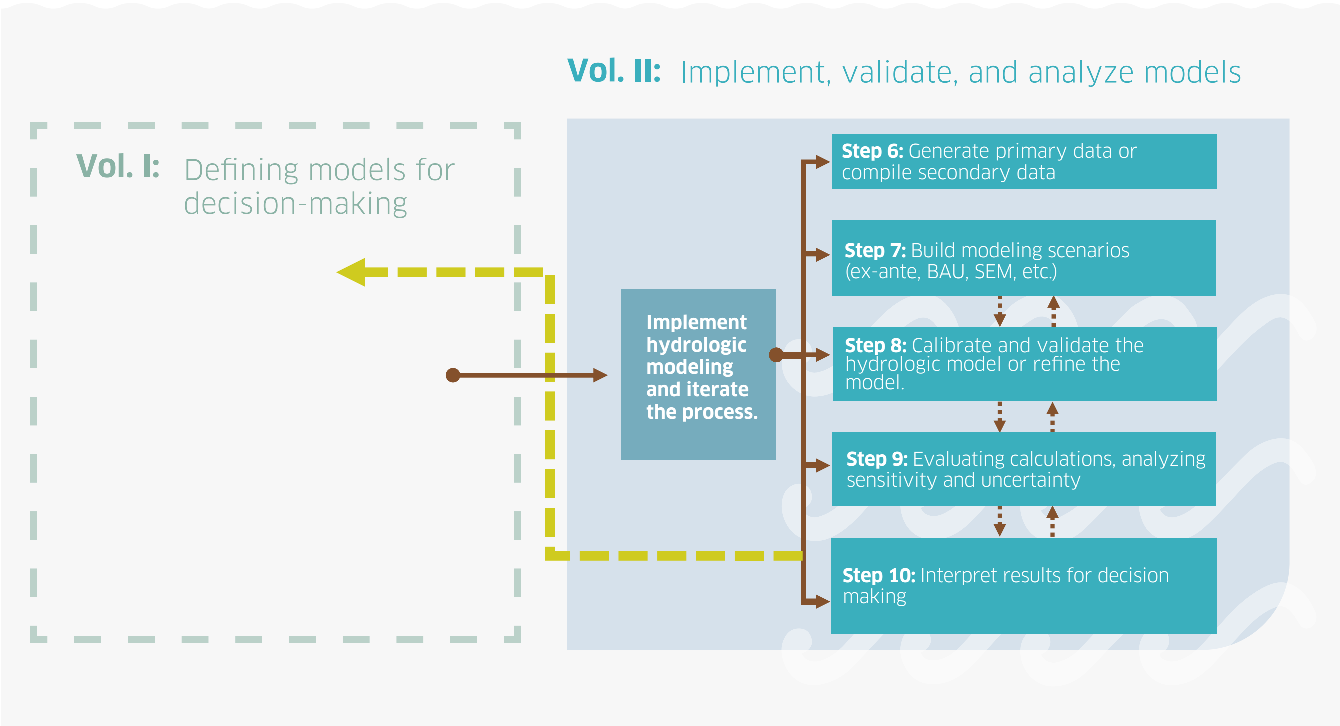
STEPS	GUIDING QUESTIONS	
STEP 5. Select the most appropriate hydrologic model	**Q18. Which hydrologic model responds positively to the highest number of criteria and how?**	The SWAT hydrologic model has been selected. This model can simulate daily flows and sediment loads in the same exercise. It is a semi-distributed model that can generate outputs in the form of maps to spatially analyze sediment sources and NI interventions. It is capable of assimilating data on vegetation cover and soil characteristics. It does not require further calibration to obtain results at the required level. The NI interventions under consideration related to land cover can be modeled with SWAT. Prior team experience in the use of the model means that results can be obtained within the given time constraints.
	Q19. What are the limitations and disadvantages of the selected hydrologic model?	SWAT uses the MUSLE method to model erosion. It is not possible to simulate other types of erosion, particularly mass movements and gullies. The hydrology of bofedales is not properly represented by the curve number method.
	Q20. What are the limitations and disadvantages of the selected hydrologic model?	KINEROS is a feasible alternative, with the disadvantage that modeling would be done on a precipitation event scale, rather than a continuous time scale. InVEST could be used to generate annual and seasonal average values, but not continuous time series.


Photo: Michell Leon

VOLUME II

Photo: Michell Leon

IMPLEMENT, VALIDATE, AND ANALYZE HYDROLOGIC MODELS



Volume II is intended for hydrologic modeling experts who need to generate model outputs to help inform decisions about NI interventions, experts in project formulation and evaluation, and decision-makers who are not familiar with hydrology or modeling but who can participate in the moddeling process to maximize its relevance for decision making. Rather than focusing on fundamental concepts of hydrologic processes, we will provide guidance on the importance of model inputs; defining and developing simulation scenarios; calibrating and validating models to make more realistic estimates; analyzing the sensitivity and uncertainty associated with these estimates; and interpreting results in a way that contributes to informed decision-making.

Several modeling ***softwares*** include calibration and uncertainty calculation tools. However, understanding these procedures and the implications for scenario simulation is useful for improving the reliability of predictions.


Photo: Ana Castañeda


Photo: Jorwi Deiby Barboza Castro

STEP 6. DATA & INFORMATION

This step seeks to clarify the role of observations and measurements in modeling practice. Modeling does not replace primary data generation or on-site monitoring – rather, it is a tool for extracting knowledge from ***"proxy"*** observations and making estimates where direct observation is not possible (e.g., predictions or ex ante evaluations). As such, the outputs of a model will, in principle, always be inferior to direct observations, and its quality will be directly related to the quantity and quality of the input data. Model input data can come from primary or secondary sources. ***Primary data*** are those that we obtain directly from reality, by collecting or producing them with our own instruments. ***Secondary data*** are those already produced by others and can be reused in various applications.

"The outputs of a model will always, in principle, be inferior to direct observations, and its quality will be directly related to the quantity and quality of the input data."

Salience, Credibility, and Legitimacy

Information has an intrinsic value that can be explained by the concepts of relevance, credibility, and legitimacy (Cash et al., 2003). Data are simple facts and numbers. We can collect data about a basin and the water flowing through it without extracting meaning from the data. Data become information when they are analyzed, interpreted, and put into context (Brauman et al., 2021a). However, not all information is equally meaningful or useful.

Brauman et al. (2021a, 2021b) analyze the usefulness of information as a function of the decision-making context. One way to analyze usefulness is to consider whether it is salient, credible, and legitimate (Cash et al., 2003). ***Salience*** is concerned with the relevance and type of information provided. Information is relevant when it is important to the question of interest. When evaluating the benefit of conserving, restoring, or protecting NI, salience describes how well the decision-making context is considered by the selection and configuration of the hydrologic model. ***Credibility*** is concerned with whether the information is transparent and robust. Information is credible when it is perceived as correct and reliable. Credibility depends on several factors, ranging from how well decision-makers understand the principles and tools used in hydrologic modeling to the level of precision and accuracy of the results. ***Legitimacy*** is concerned with how the information has been produced and communicated. Information is legitimate when productive and honest interactions with decision-makers result in relationships, trust, and processes that meet standards of political and procedural equity.

Salience is the importance of the information to a decision.

- Does the information provided address the issues of interest for decision-making?
- Were the correct hydrologic variables modeled?
- Do the findings inform the objectives?

Credibility is the perception that the information meets standards of scientific and technical robustness.

- What is the quality of the information produced?
- Is it authoritative, trustworthy, and reliable?
- Do decision-makers understand how the information was generated and trust that it is accurate and precise?

Legitimacy is the perception that the information was generated in a fair and unbiased manner, considering appropriate values, concerns, and perspectives of diverse individuals and institutions.

- Does the process of information generation and communication meet standards of political and procedural equity?
- Are decision-makers confident in the technical staff?
- Is the evaluation and communication of information trusted to be transparent?

Source: Brauman et al. (2021a); definitions adapted from Cash et al. (2003)

In the context of NI, salience requires that a hydrologic model incorporate the objectives and interventions relevant to the decision-making context in question (Brauman et al., 2021). A precise and accurate, and therefore credible, model that predicts how changes in precipitation affect downstream flows but does not incorporate the effects of changes in NI would not be not salient. A salient model that informs the decision-making context can be considered credible, even if it is less accurate or precise, if decision-makers understand the simplifications of the modeling process. Whether the outputs of a hydrologic model are considered salient and legitimate reflects whether decision-makers were consulted during the modeling development process and therefore trust the organization producing the information. (Hamel et al., 2020; Bremer et al., 2020).

Importance of Observations and Measurements

All hydrologic modeling depends, to different degrees, on field observations, or primary data. As direct observations, primary data tend to offer a better understanding of reality and produce a model with better results. The usefulness of observations in the modeling process can be evaluated during: i) the definition of the decision context; ii) model choice; iii) configuration; and iv) calibration. All members of a project should be reminded that models are approximations of reality. The information available will determine how well the modeled area is understood, whether information needs to be collected to complete modeling, and any other limitations or considerations for the process. In many cases, there will be qualitative information that must be transformed into quantitative information; for example, a sandy type soil can be represented using its saturated hydraulic conductivity (between 10^{-5} and 10^{-3} m^3/s) so it can be incorporated into the model (Boeker & van Grondelle, 1995). At other times, the information will have limitations; for example, V-notch triangular weirs have high accuracy for low flows, but their uncertainty increases for high flows (Herschy, 2009). Another common situation is the lack of flow information to calibrate or evaluate a hydrologic model. This is problematic because it prevents model evaluation in the absence of data. Making predictions in basins that do not have flow observations is a common challenge for hydrologic sciences and is known as ***"predictions in ungauged basins"*** (Blöschl et al., 2013; Ochoa-Tocachi et al., 2016b).

The available information allows a modeler to build an understanding of the hydrologic system, such as choosing the hydrologic processes and the equations to represent them. For example, the type of vegetation cover in an area can indicate how important it is to include interception in a hydrologic model. The importance of interception in a basin dominated by grasslands may be negligible, but it may be a very influential process in a forest. The presence of bofedales could suggest the need to include a ***"reservoir"*** component in the hydrologic model that mimics their capacity to buffer surface runoff. Bioclimatic information, such as precipitation intensity, the biome to which a basin belongs, and topographic information can inform the order of magnitude of flows or the way runoff is generated (Ochoa-Tocachi et al., 2018a), which indicates how to mathematically represent them in a model.

The predictive capacity of a model depends on the balance between its complexity and the amount of information available. Existing models can be as simple as equations converting precipitation to runoff using a dimensionless factor, such as the runoff coefficient, or complex, like a system of higher order differential equations. These features have implications for computational cost, model parameters, and predictive capability. As mentioned before (**Figure 8**), greater complexity in a hydrologic model does not necessarily translate to better results. Implementing complex models with little information or with poor quality data produces poor results due to model overparameterization problems and the need to make several assumptions about parameter spatial variability (cf. Muñoz et al., under review).

> ***"Implementing complex models with poor quality data produces poor results due to model overparameterization problems and the need to make several assumptions about parameter spatial variability."***

The concepts of ***accuracy*** and ***precision*** are used to define the quality of observations and measurements. Sometimes these terms are used interchangeably, but they describe different elements of information quality. Accuracy refers to the time lag of the data, while precision refers to the dispersion of the data **(Figure 11).** For example, for an existing real datum, represented by the red circle in **Figure 11,** several measurements are generated with different degrees of accuracy and precision represented by the gray dots. If the average of the measurements is within the red circle, we can say that the measurements have a high level of accuracy **(Figures 11a, 11d).** If the average of the measurements is outside the red circle, the accuracy of the measurements is low **(Figures 11b, 11c).** If the variability of the observations is high (i.e., there is high dispersion in the data), we can say that the measurements have low precision **(Figures 11a, 11c).** In turn, if the dispersion is low (i.e., the data are concentrated in one area), we can say that the measurements have high precision **(Figures 11b, 11d).**

Accuracy and precision do not necessarily go hand in hand. Ideally, both accuracy and precision are high (i.e., the average of the measurements is close to the true value and the dispersion of the data is low). From a statistical point of view, the worst case is low accuracy and low precision **(Figure 11c).** However, the most unfavorable case in field data measurement is one in which accuracy is poor (i.e., measurements are far from the true value) but precision is high **(Figure 11b).** This case is dangerous because the high precision generates a false sense of certainty, when, in fact, the measured values may be very far from the true value and lead to erroneous conclusions. Even if the precision of the data is low (i.e., there is high dispersion), one should ideally ensure that the accuracy is high so that at least the average of the dispersed data is a good approximation of the true value.

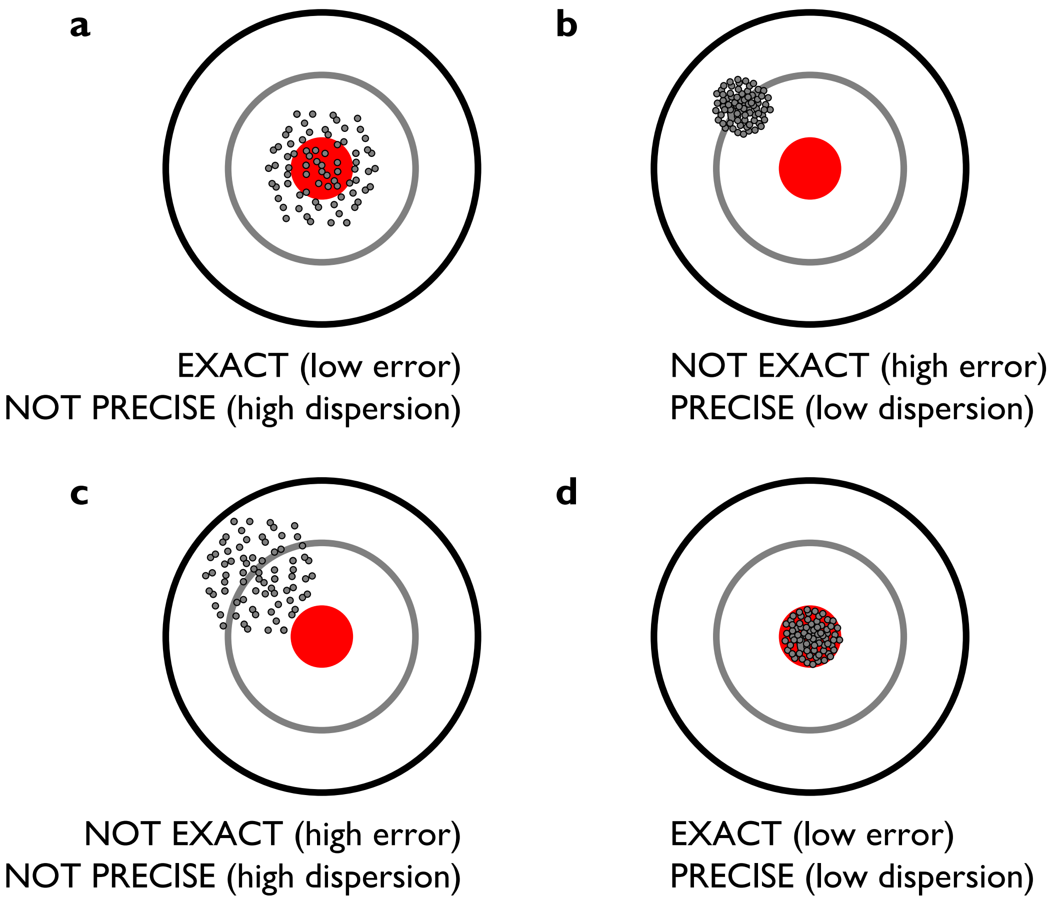


Figure 11. *Data Accuracy versus Data Precision. Source: NIWS Project*

"The most unfavorable case in field data measurement is one in which accuracy is poor but precision is high, because it generates a false sense of certainty."

The nature of the data needed for the models depends on the modeling objectives. Different criteria can be considered in data generation to meet those aims, depending on spatial and temporal scales:

- **Spatial scale**
  - The monitoring area should be representative of the processes and conditions of the surrounding environment.
  - **At the basin scale:** precipitation, flow, biophysical properties.
  - **At plot scale:** weather station, hydrologic processes, vegetation processes, specific properties.

- **Time scale**
  - **Immediate:** speed of response of the catchment, hydrologic regulation.
  - **Hourly:** intensity of precipitation.
  - **Daily:** medium and long-term monitoring, water yield.
  - **Monthly:** seasonality, long-term modeling, water resources management.
  - **Annual:** climate variability, interannual cycles (e.g., El Niño), climate change.

Observations, especially those of precipitation and flow, are used to calibrate a hydrologic model. Once existing conditions are understood, assumptions have been made, and the model configured, it is necessary to estimate the best parameters that can approximate observed measurements using optimization algorithms. If these observations did not exist, the parameter estimation process would not be able to be performed quantitatively, which would decrease the reliability of the model results. Flow or precipitation measurement errors can also be amplified during calibration. For example, if a hydrologic model is intended to simulate flood events (i.e., peak flows), but is calibrated with data from a hydrometric station designed for gauging low flows, this uncertainty in the measurement of high flows will propagate and contribute to overall model uncertainty.

Both existing conditions and the choice of model complexity are translated into **model structure.** which is the set of explicit or implicit assumptions and choices made in a model. These may include the description of hydrologic processes, how they are coupled, numerical discretization, the representation of spatial variability, and so. (Butts et al., 2004). However, evaluating data usefulness extends beyond model structure. Various sources can be used to bring the information available in a study area closer to a specific model configuration. For example, delimiting a basin requires topographic information, which has been derived from remote sensing and is available through satellite products, such as ASTER or ALOS-PALSAR. To estimate the water infiltration capacity of soils, hydrophysical tests can be performed in the field to obtain more accurate values. Greater precision on model parameter information translates into greater certainty in results. Another alternative is to combine layers of geographic information of land use and land cover obtained indirectly. This type of information is known as secondary data, which are valuable assets in the face of primary data scarcity.

Availability of Secondary Data

Where primary data cannot be generated, it is important to consider the compilation of available secondary data. Several models need a set of secondary data to be operational, including distributed topography information in the form of a ***digital elevation model*** (DEM), land use and land cover maps usually available at a regional scale, maps of soil types and their water transport and retention characteristics, information on abstractions and flow returns, usually from human consumption authorization databases, etc. DEMs, for example, are constructed by radar interferometry, stereoscopy from pairs of aerial images or satellite shots, digitization of the contour lines of a map, direct input of the coordinates (x, y, z) of terrain points measured by GPS, triangulation (by topography), or by means of a laser rangefinder, airborne laser system (LIDAR), etc. In most cases, DEMs are generated by institutions, such as the United States Geological Survey (USGS) and other national and international geographic authorities, and are compiled by a modeling team to be used as secondary data.

"Several models need a set of secondary data to be operational."

At times it is not possible to obtain the information required. For example, while a DEM seeks to represent the spatial distribution of terrain elevation without all natural and built elements, there are variants. A digital surface model (DSM) is a faithful estimation of the earth's surface, including all the objects it contains, incorporating natural and man-made or constructed features. A digital terrain model (DTM) typically augments a DEM, including vector elements of the natural terrain, such as rivers or foothills, but without any objects such as vegetation or buildings (Li et al., 2005). A DTM can be interpolated to generate a DEM, but not vice versa. Each type of digital model can be used for different applications, but sometimes they are used interchangeably, depending on availability and access. In turn, these secondary data may be available at different scales, resolutions, and quality. For example, it is common to have DEM rasters with a resolution of 90 m openly available worldwide (USGS EarthExplorer), while DEM rasters of finer resolutions must be produced with topographic surveys, LIDAR systems, or purchased from third parties **(Table 12)**. Sources of secondary data for hydrologic modeling are shown in **Table 13**.

Table 12. *Different Digital Elevation Model (DEM) Resolutions and Applications. Source: Virtual Terrain Project (2021).*

APPLICATION	ARC	METERS (APPROXIMATE)
Global world map	1 degree	110 kilometers (km)
Size of a traditional map quadrant	7.5 arcminutes (arcmin)	14 km
Global elevation map	30 arcseconds (arcsec)	1 km
SRTM available worldwide	3 arcsec	90 meters (m)
Topography DEM, USA SRTM available	1 arcsec	30 m
Topography DEM	1/3 arcsec	10 m
LIDAR DEM (new products)	1/9 arcsec	3.4 m

Table 13 . *Secondary Data Sources for Hydrologic Modeling. Source: Prepared by the authors.*

TYPES OF DATA	SOURCE OF DATA
Digital elevation models, land uses, other global data	**USGS EarthExplorer, SRTM,** at https://earthexplorer.usgs.gov **ASTER,** at https://asterweb.jpl.nasa.gov/gdem.asp **ALOS–PALSAR,** at https://earth.esa.int/eogateway/missions/alos **NASADEM,** at https://earthdata.nasa.gov/esds/competitive–programs/measures/nasadem **MERIT DEM,** at http://hydro.iis.u-tokyo.ac.jp/~yamadai/MERIT_DEM/
Climate data from in situ stations	**SENAMHI,** at https://www.senamhi.gob.pe **NOAA NCDC,** at https://www.ncdc.noaa.gov/cdo-web/ **SO-HYBAM,** at https://hybam.obs-mip.fr

TYPES OF DATA	SOURCE OF DATA
Hydrometeorological data from Andean ecosystems	**iMHEA,** at http://imhea.org https://figshare.com/collections/High-resolution_hydrometeorological_data_from_a_network_of_headwater_catchments_in_the_tropical_Andes/3943774 **SEDC FONAG,** at http://sedc.fonag.org.ec
Satellite precipitation products	**TRMM,** at https://climatedataguide.ucar.edu/climate-data/trmm-tropical-rainfall-measuring-mission **GPM,** at https://gpm.nasa.gov/data/directory **Precipitation Processing System,** en https://arthurhou.pps.eosdis.nasa.gov
Reanalysis climate data, temperature, precipitation, meteorological variables	**PISCO,** at https://www.senamhi.gob.pe **WorlClim,** at https://www.worldclim.org/data/index.html **ERA5,** at https://cds.climate.copernicus.eu/cdsapp#!/dataset/reanalysis-era5-single-levels **CHIRPS,** at https://www.chc.ucsb.edu/data/chirps https://data.chc.ucsb.edu/products/CHIRPS-2.0/global_daily/netcdf/p05/
Land cover and land use maps	**GlobaLand 30,** at http://www.globallandcover.com/defaults_en.html?src=/Scripts/map/defaults/En/download_en.html **Sentinel,** at https://sentinel.esa.int/web/sentinel/thematic-areas/land-monitoring/land-cover-use-and-change-detection-mapping **MODIS,** at https://modis.gsfc.nasa.gov/data/dataprod/mod12.php **MODIS Terra,** at https://terra.nasa.gov/about/terra-instruments/modis **MODIS Aqua,** at https://oceancolor.gsfc.nasa.gov/data/aqua/ **MINAM,** at https://sinia.minam.gob.pe/mapas/mapa-nacional-ecosistemas-peru

TYPES OF DATA	SOURCE OF DATA
Soil type and property data	**SMAP soil moisture,** at https://smap.jpl.nasa.gov **GRACE gravity,** at https://podaac.jpl.nasa.gov/GRACE **FAO Soils Portal,** at http://www.fao.org/soils-portal/soil-survey/soil-maps-and-databases/en/ http://www.fao.org/soils-portal/soil-survey/soil-maps-and-databases/harmonized-world-soil-database-v12/zh/ **ISRIC,** at https://data.isric.org
Global geology data	**OneGeology,** at http://www.onegeology.org **USNA,** at https://www.usna.edu/Users/oceano/pguth/md_help/html/global_geology.htm
Global compendium of hydro-environmental variables	**HydroATLAS,** at https://www.hydrosheds.org/page/hydroatlas **HydroSHEDS,** at https://www.hydrosheds.org

In short, modeling is a very useful tool for making estimates when direct observation is not possible. The reliability and certainty of these results will depend on the quantity and quality of the available observations. Observations are the starting point for understanding the reality of the system to be modeled, choosing the complexity of the model, setting up operations, then calibrating and validating it. When it is not possible to generate or collect primary data, alternative information may be sought from external sources in the form of secondary data. Secondary data are common and necessary in various types of models. In principle, model estimates will always be more uncertain than direct observations, and the quality of the model outputs will depend directly on the quality and quantity of the input data.


Photo: Natural Infrastructure for Water Security

Photo: Will Espinoza

STEP 7. MODELING SCENARIOS

This step describes how the hydrologic processes will be represented in a model as "scenarios". In computer science, "garbage in, garbage out" is the concept that flawed, or nonsense input data produces nonsense output. This concept is related to information or product quality that enter a system: if the quality of what enters is poor, the result and conclusions are therefore generally also poor.

A simulation will be a direct result of the assumptions and the parameterization of evaluation "scenarios", as much as the quantity and quality of the input data. For example, when replacing a degraded soil infiltration capacity A with a recovered soil infiltration capacity B, the model will show a change equivalent to the difference between B–A. The development or selection of a scenario to evaluate is as key to soundly answering policy questions as the selection of the model. To define scenarios, it is necessary to answer the following questions: What is the purpose of modeling? What needs to be represented?

> ***"A simulation will be a direct result of the assumptions and the parameterization of evaluation scenarios, as much as the quantity and quality of the input data."***

Several concepts are used to define modeling scenarios, the purpose of which are to build ***modeling experiments*** to evaluate results. For example, if we want to estimate the benefits of a NI intervention, we must remember that this infrastructure already exists and that we cannot attribute all the ecosystem services and the associated benefits to the interventions implemented. It is necessary to define differences that quantify the additional value of the benefits resulting from interventions. One way to conceptualize these differential effects is to compare possible future scenarios. For example, if the focus is on the value of conservation or protection of NI, conservation scenarios versus degradation scenarios can be compared. If the focus is placed on the value of recovery or restoration of NI, restored state scenarios versus current state scenarios can be compared. Finally, if the focus is placed on the implementation of a project (e.g., intervened vs. non-intervened NI, gray vs. natural infrastructure, or gray and natural infrastructure combined), scenarios with project versus no project can be compared.

The time variable and assumptions about the future state are determining factors in creating scenarios. The study area is subject to uncertainty, uncontrolled conditions, in addition to the effect of deliberate human interventions. In other words, as time progresses, some areas or elements of the NI that are currently degrading will continue to decline unless prevention, mitigation, or remediation actions are carried out. Similarly, nature has its own regenerative capacity, and some areas or elements of the NI will be subject to this force. In addition to these natural processes, human beings interact daily with the landscape, causing land-use change that can affect or benefit NI intervention projects. Finally, there are environmental changes, such as those caused by climate change, which can have a considerable effect on current and future conditions and which may or may not be considered during model and simulation development. Various scenarios can be developed to evaluate NI through modeling that take these factors into account.

Current Scenario or Baseline (BASE)

The current state or existing condition of the study area is the foundation of all NI interventions and modeling scenarios. This serves as the ***"baseline"*** for evaluation, for both positive and negative impacts expected from interventions, or the absence of those interventions **(Figure 12).** For this first activity, the following information must be collected:

- Meteorological and hydrologic information of the areas of interest during the most recent available years.
- Biophysical data on the ecosystem: topography, elevation, soils, geology, vegetation, land use and cover, among others.
- Time series or average daily, monthly, or annual flow data at the points of interest during time period(s) of interest.
- Flows and points that have been intentionally derived, diverted, or captured from the natural drainage network for various water uses.
- Flows and return points the drainage network after being used (if known).
- Flows and transfer points from the basin outwards and from outside into the basin.
- Time series or data of daily, monthly, and annual average concentrations of water quality parameters relevant to the evaluation. The quality indicators must respond to the needs and availability of information.
- Existence of hydraulic infrastructure or other elements in the study area that modify the priority ecosystem services (storage reservoirs, flood control works, erosion control, sand trap works, etc.).

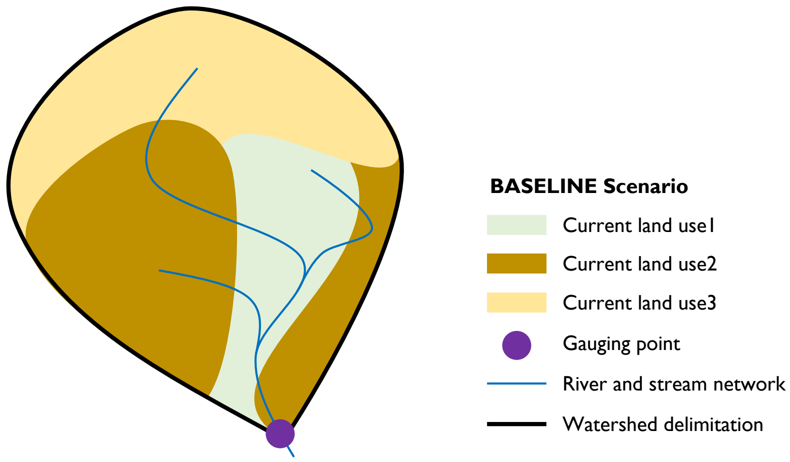


Figure 12. *Conceptual representation of the baseline scenario in an analysis basin. Source: Prepared by the authors.*

Baseline information can be used to determine the initial values of the hydrologic ecosystem service of interest (SEH_A), which generates time baseline T_I, conceptually represented by point A in **Figure 13.**

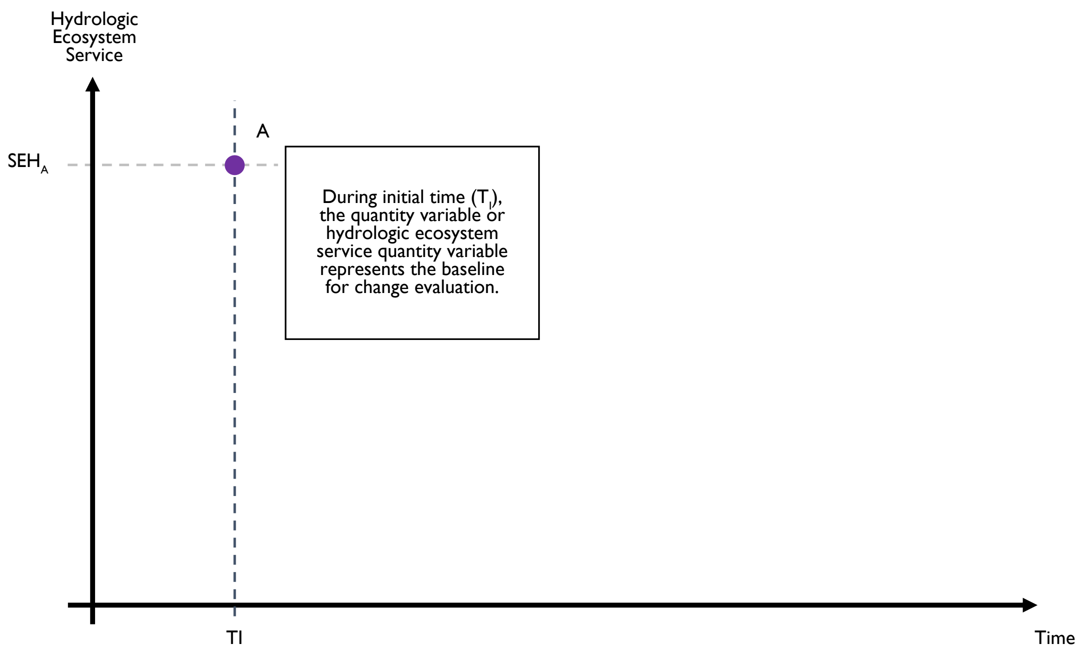


Figure 13. *Initial status of the hydrologic ecosystem service quantity or quality of interest at the point of evaluation. Source: Ochoa–Tocachi et al. (2018b).*

BAU (Business As Usual) or Trend Scenario

A scenario resulting from a ***"trend"*** of anthropomorphic land use and natural degradation mechanisms, without interventions for conservation, protection, or restoration of NI is known as ***"business as usual"*** (BAU) (e.g., current development pattern/expansion of urban, agricultural, and livestock areas). Business as usual can be used to compare how ecosystem services of interest might change from the baseline **(Figure 14).**

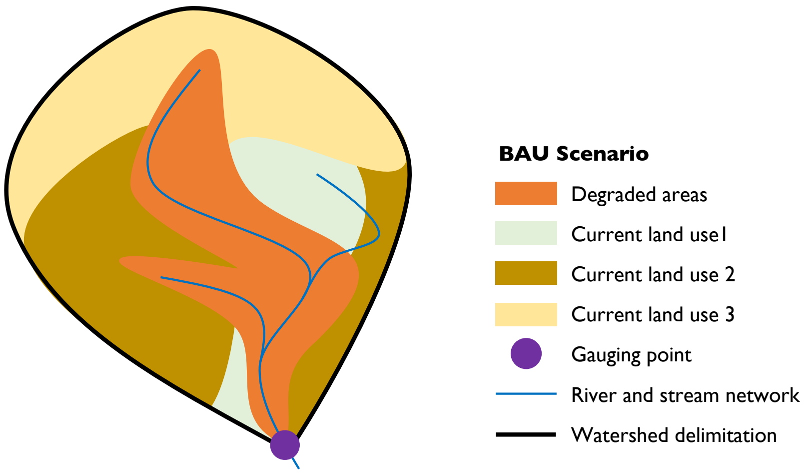


Figure 14. *Conceptual representation of the trend or business as usual scenario in an analysis basin. Source: Prepared by the authors.*

In Figure 15, the variables that determine the ecosystem services of interest are projected within a defined time period (e.g., 40 years), using biophysical and historical information previously collected through a geographic analysis tool. Ideally, there is a multi-temporal analysis of the changes in vegetation cover and land use in the study area, which allows historical change rates and the most critical locations to be identified when designing scenarios and interpreting model projections. Conceptually, **Figure 15** shows a change function with negative slope which theoretically represents that, without interventions on the natural infrastructure, the performance of hydrologic ecosystem services decreases over time ($SEH_B < SEH_A$), represented by point B on T_2.

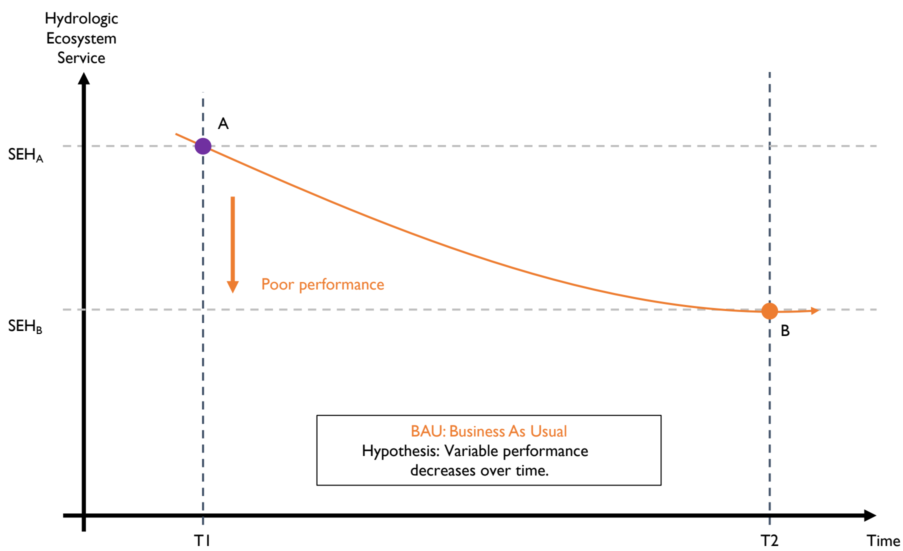


Figure 15. *Projected trend scenario of the hydrologic ecosystem service at the study point within the time period defined for the evaluation. Source: Ochoa–Tocachi et al. (2018b).*

PES (Pessimistic) or Maximum Potential Degradation Scenario

The ***"pessimistic"*** scenario (PES) represents the worst possible state of degradation of NI and, consequently, of the priority hydrologic ecosystem services. Usually, this scenario is simply constructed with the consideration that all the important areas of water are degraded by human factors and environmental changes. Examples of this type of scenario are the burning of large tracts of natural grasslands and forests, or full withdrawal and disappearance of glaciers **(Figure 16).** On many occasions, the PES scenario is used as the point of comparison for the benefits of NI projects to demonstrate the extreme consequences of taking no action at all. It is also an easy scenario to build and does not require many considerations. However, we must try not to overestimate the benefits of NI projects by comparison with a PES scenario, because it is possible that this scenario would require very specific conditions to happen in real life.

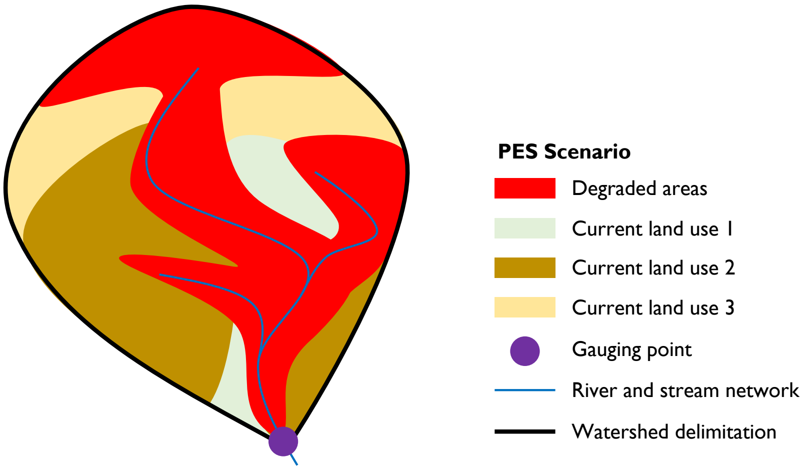


Figure 16. *Conceptual representation of the pessimistic scenario in an analysis basin. Source: Prepared by the authors.*

Conceptually, **Figure 17** represents that in a worst-case scenario, the performance of hydrologic ecosystem services falls in time to minimum values ($SEH_P < SEH_B < SEH_A$), represented by point P on T_2.

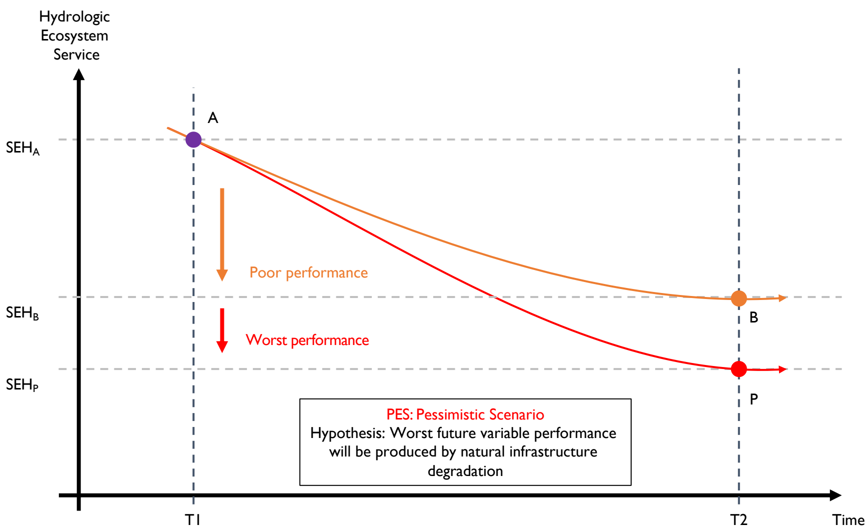


Figure 17. *Pessimistic scenario of the hydrologic ecosystem service performance in which the state of the natural infrastructure is totally degraded. Source: Ochoa–Tocachi et al. (2018b).*

NI or Natural Infrastructure Management Scenario

In this scenario, the performance of ecosystem services resulting from interventions is projected onto the NI over a given time, particularly at the points of interest. With the use of available spatial information and geographic analysis tools, such as HIRO (Rapid Opportunity Identification Tool: Román et al., 2020), priority areas for water retention, erosion control, and hydrologic ecosystem services can be identified. These areas can be focused on the design of NI interventions, such as wetland restoration, forest conservation, control of agricultural border expansion, and other changes in land use **(Figure 18).**

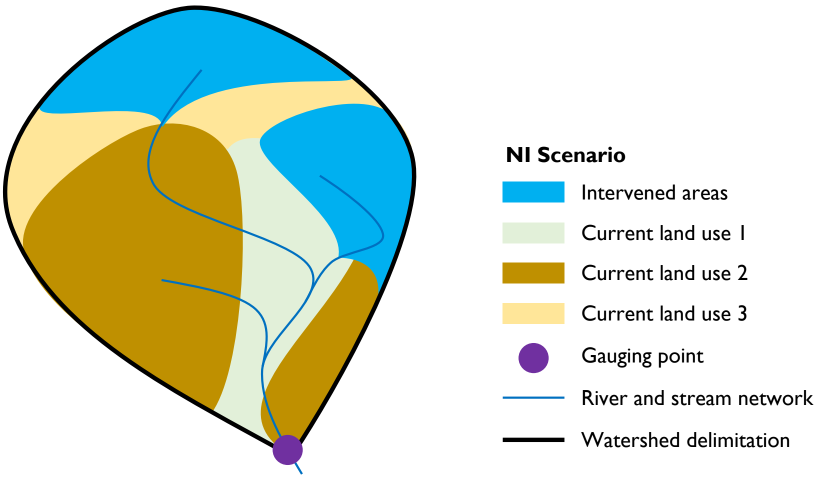


Figure 18. *Conceptual representation of the natural infrastructure management scenario in an analysis basin. Source: Prepared by the authors.*

The model must be able to determine whether an intervention produces a positive impact, no impact (i.e., null), or negative impact. The proposed hypothesis is that NI interventions generate positive changes in the quality and quantity of hydrologic ecosystem services in the study area (SEH_R, point R, **Figure 19**), resulting in impacts at least better than those of the projected situation in the BAU scenario, even if they are not as good as the baseline ($SEH_B < SEH_R < SEH_A$ at T_2)

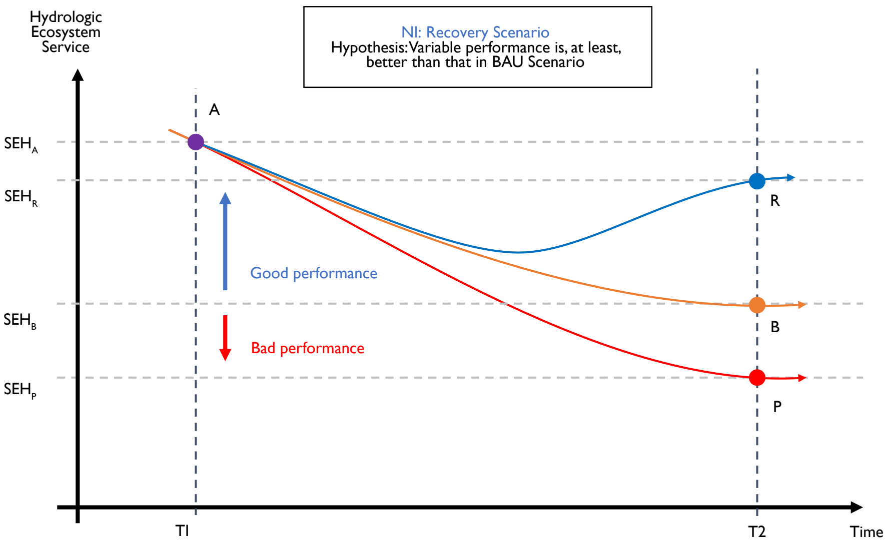


Figure 19. *Scenario of hydrologic ecosystem service performance recovery by interventions in the natural infrastructure. Source: Ochoa–Tocachi et al. (2018b).*

SEM *(sustainable ecosystem management)* or Optimistic Scenario

Natural infrastructure interventions applied to the maximum potential scale generate the ***"sustainable ecosystem management"*** (SEM) scenario. This scenario encompasses the development of activities for the protection, conservation, maintenance, and recovery of NI in water sources, including projects for the conservation of biodiversity, ecological restoration, and sustainable production, to increase the supply of hydrologic ecosystem services of interest, especially more water of better quality **(Figure 20).**

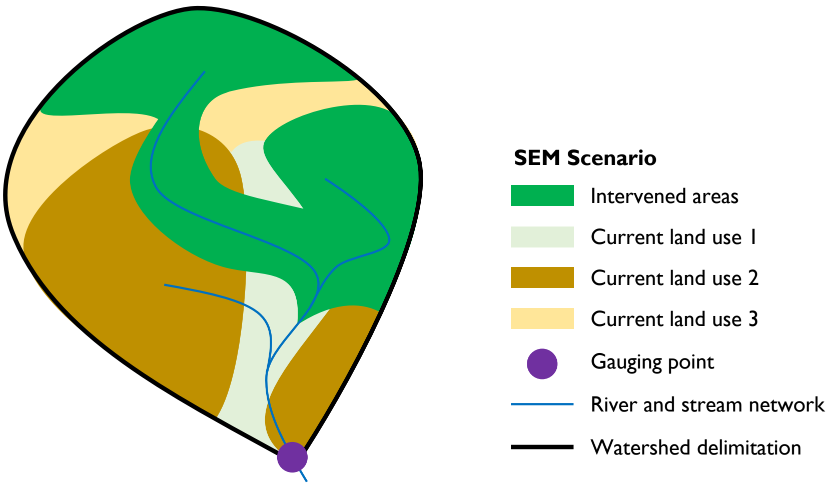


Figure 20. *Conceptual representation of the sustainable ecosystem management scenario in an analysis basin. Source: Prepared by the authors.*

Compared to the NI scenario, the interventions can have at least two other outcomes: i) the water quantity or quality decreasing trend in the ecosystem service is reversed, so that the initial characteristics are maintained ($SEH_R = SEH_A$); and ii) NI interventions allow the basin to exceed quantity or quality standards of initial time T_I ($SEH_S > SEH_A$, point S in **Figure 21**).

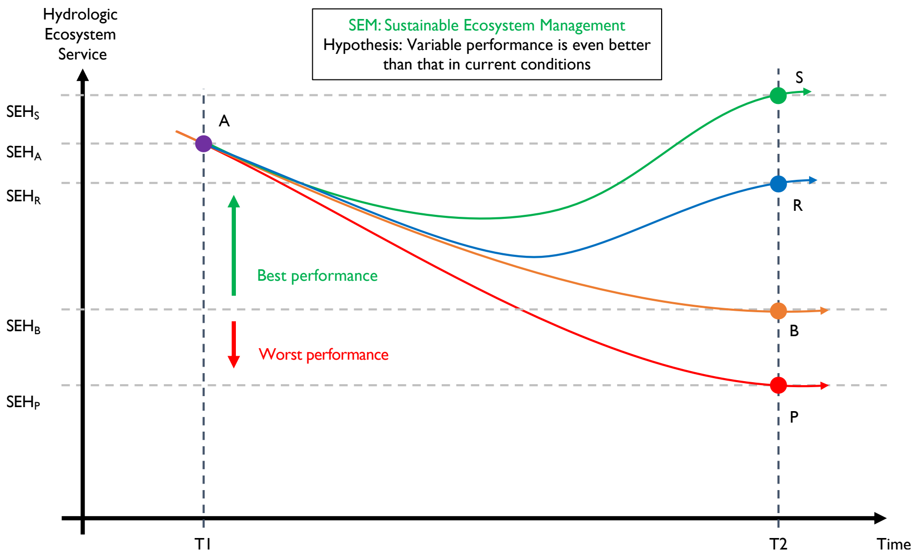


Figure 21. *Scenario of sustainable ecosystem management resulting from interventions on natural infrastructure. Source: Ochoa–Tocachi et al. (2018b*

When comparing different scenarios **(Figure 22)**, net benefits are calculated as the difference between indicators that quantify ecosystem services produced under simulated scenarios **(Figure 21)**. This approach raises the concept of *additionality*, which is defined as the **additional** benefits that would be generated as a result of an NI intervention project that would not have occurred in the absence of the project.

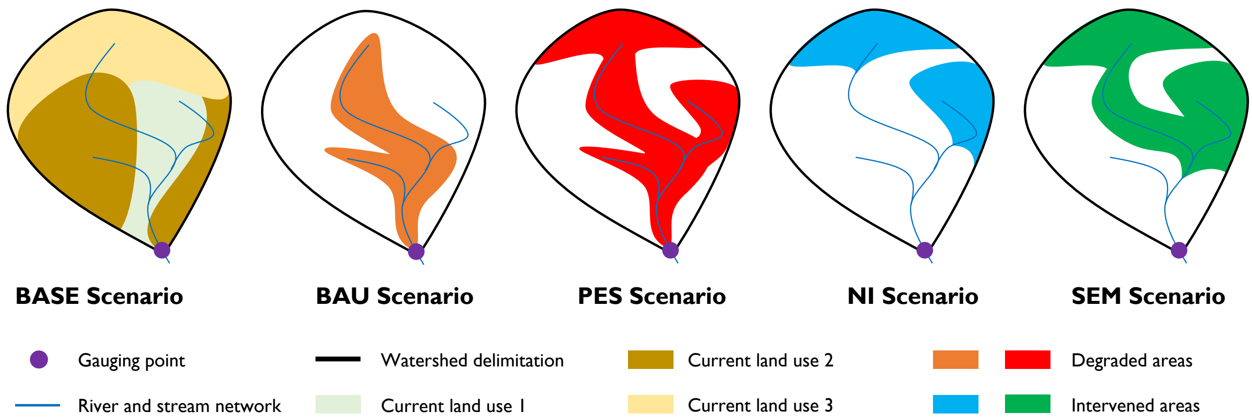


Figure 22. *Conceptual summary of the types of scenarios suggested to evaluate natural infrastructure benefits. Source: Prepared by the authors.*

Climate Change Scenarios

Finally, future projections involve not only changes in the biophysical characteristics of the basin or study area (soil, vegetation, infrastructure, etc.), but also environmental ones, such as climate change. Representative concentration pathways (RCP) are common ways to represent and analyze climate change. An RCP is a greenhouse gas concentration pathway adopted by the Intergovernmental Panel on Climate Change (IPCC). For the *Fifth Evaluation Report of the IPCC* (2014) four pathways were used for climate modeling, describing different possible climate futures that depend on the volume of greenhouse gases emitted in the coming years. RCPs 2.6, 4.5, 6.0 y 8.5 are labeled from a range of possible radiative forcing values in the year 2100 (2.6, 4.5, 6.0 and 8.5 W/m^2, respectively) **(Figure 23).**

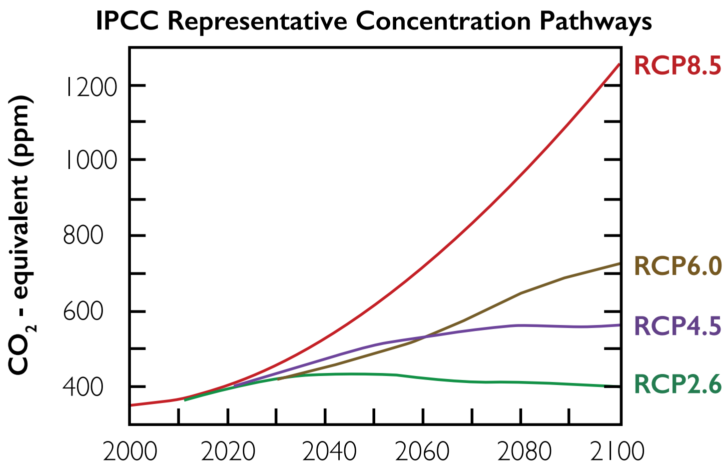


Figure 23. *CO_2 equivalent concentrations for various emission scenarios (parts per million) by the four RCPs used in the Fifth Evaluation Report of the IPCC for predictive models. Source: IPCC (2014).*

A simple way to consider climate change scenarios is modifying the meteorologic input data to a model (temperature, precipitation, evapotranspiration, etc.). Commonly, several hydrologic models perform a sensitivity analysis (Step 9a) of climatic data. For example, this can be done considering variations in precipitation of ±5%, ±10%, ±20%, etc., or considering variations in average temperature of 1°C, 2°C, 5°C. Changes in temperature result in changes in evapotranspiration, i.e., water flowing to the atmosphere, as well as glacial contribution to streamflow, because glacier retreat depends on temperature. Changes in other meteorologic variables, such as air humidity or wind velocity, can also be considered in climate change scenarios, although they are more uncertain and require in situ measurement data. Another way to represent climate change scenarios is using future temperature projections (high temperature hazard maps) and future precipitation (high intensity rainfall hazard maps). Although it is possible that changes in future total precipitation amounts are uncertain, other modeling scenarios can consider effects in variables such as rainfall intensity or climatic extreme events.

However it is important to note that as more assumptions are made regarding the future, the greater the uncertainty of projections. For this reason, it is not recommended to combine land use and land cover change analysis with climate change analysis, because the climate change uncertainty could mask the differences in these more complex scenarios and hide the benefits of NI interventions. This uncertainty could also under or overestimate the projected benefits and obscure the identification of evidence of the expected benefits of the interventions. If it is deemed appropriate to combine various types of assumptions about future changes, a careful uncertainty analysis should be applied, which may significantly increase the resources, technical expertise, and modeling time required.The construction of more complex scenarios does not necessarily mean that better results will be obtained. Simulation experiments with simple scenarios to reflect the expected impacts of NI interventions are recommended – representing and communicating in the clearest, most direct way.

"The construction of more complex scenarios does not necessarily mean that better results will be obtained."

STEP 8. MODEL CALIBRATION AND VALIDATION

Manual and Automatic Parameter Calibration

Calibration is a process of trial and error where parameter configuration of the model is tested by comparing results against observed data for the study area (i.e., "real world" data) **(Figure 24)**. This process has three main steps: 1) defining the input parameters; 2) modeling based on said parameters to generate results; and 3) performance analysis of the results (Pianosi et al., 2015, 2016).

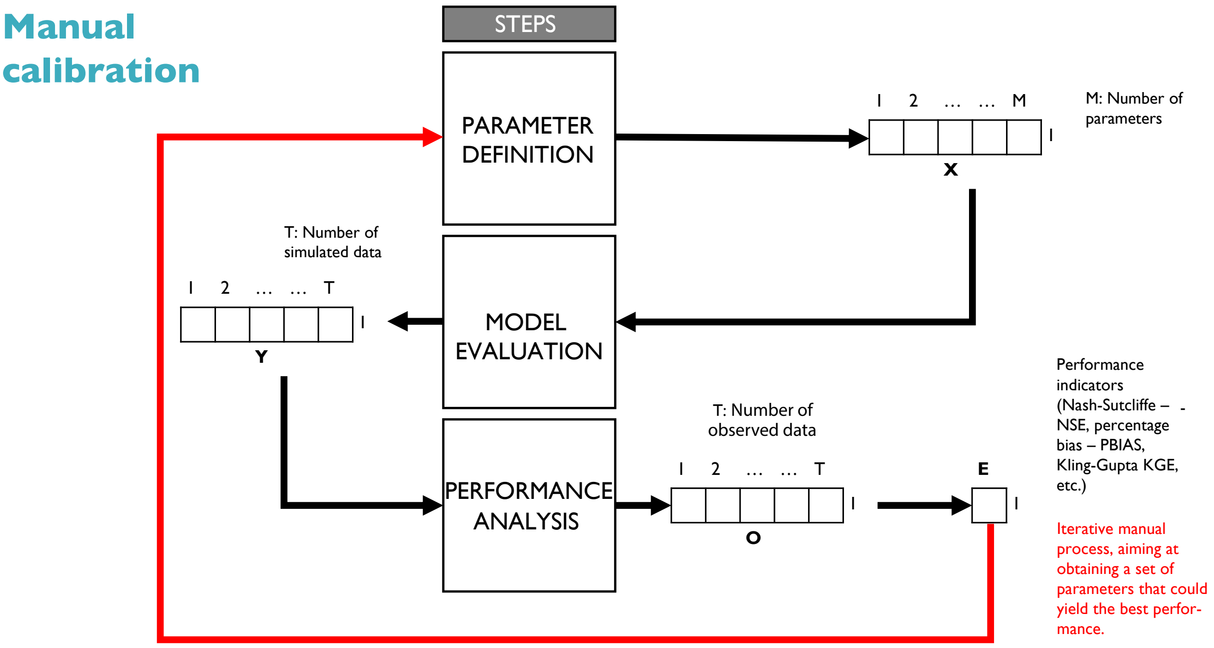


Figure 24. *Manual calibration process of a simulation model. Source: Ochoa–Tocachi et al. (2020a).*

In a ***manual calibration*** the number **M** of parameters are defined. These parameters are likely to be modified during the calibration exercise and their sensitivity can be analyzed as indicated in the following subsection. This list of parameters is entered into the model, which then generates a set of results with a length of **T** (for example, T can be the number of months simulated or the number of data generated). This set of simulated results is analyzed to measure its performance against a set of observed data. Ideally, there is one observed data point for each simulated data point for comparison, which is quantified using performance indicators, such as Nash-Sutcliffe efficiency or percentage bias. When there are several criteria for performance evaluation, the exercise becomes a ***multi-objective optimization analysis***. If performance is not satisfactory, the process is repeated, changing the definition of one or more of the parameters and evaluating their performance each time. This process can therefore be iterative, take considerable time and effort, and does not necessarily result in an optimal set of parameters. In such cases, automatic calibration can help improve this process.

> ***"Calibration is a trial and error process where parameter configuration of the model is tested by comparing results against observed data."***

In an ***automatic calibration***, the sampling process produces not just one combination of parameters, but hundreds or thousands of unique combinations. This sampling is generated following a formal statistical strategy, such as Monte Carlo or Latin Hypercube Sampling, and assumes a probability distribution for each parameter. The result is a matrix **X** of size N x M, where **N** is the number of individual combinations and M is the number of parameters **(Figure 25).**

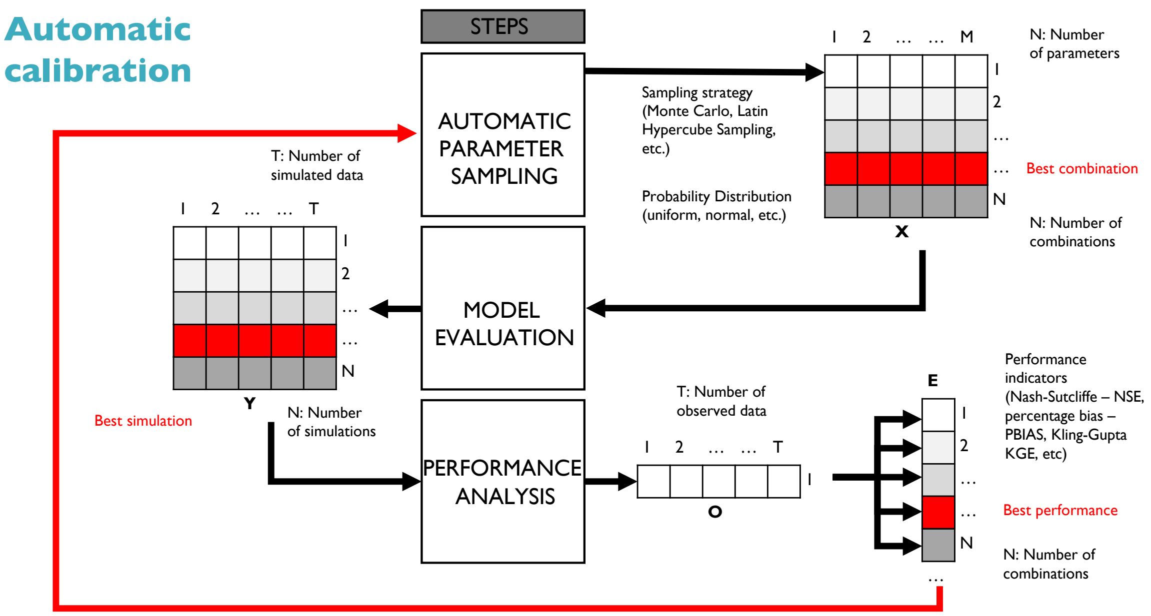


Figure 25. *Automatic calibration process of a simulation model. Source: Ochoa–Tocachi et al. (2020a).*

Matrix X of parameters is entered in the model, which generates **N** sets of results, one for each combination of parameters entered. The result of the modeling is a matrix of size N x T, where **T** is the number of data generated (e.g., the number of simulated months of flow). This Y matrix of simulated results is analyzed to measure the performance of each line against the set of observed data.

Instead of a single modeling result, N results are generated for each performance indicator. An **E** matrix of size N x L is obtained, where **L** is the number of indicators analyzed (in **Figure 25**, a single indicator). From this list of indicator results, the one showing the ***"best performance"*** is selected. This performance will be linked to the ***"best simulation"***, which is the result of the ***"best combination"*** of parameters. The process can be performed once or repeated for as many iterations as necessary to refine the parameter sampling ranges. The advantage to automatic calibration is that the process is N times more efficient than manual calibration and can give optimal results in a single run.

Automatic Parameter Sampling

To increase calibration efficiency, it is necessary to define ranges within which ***"optimal"*** values of the parameters are expected to be found. If the ranges are too wide, the process loses efficiency because it must consider many parameters that are already unsuitable. However, if the ranges are too limited, there is a risk of leaving out a combination that represents the ***"global optimum"*** and instead end up in a ***"local optimum"***. Local optimums are those combinations of parameters that might have acceptable performances; however, there could be better parameter combinations that were excluded from the sampling range. For this reason, parameter sampling must balance these challenges and reduce biases – a tight range for more model efficiency, but not so tight that it creates a hyper localized result that excludes other realistic outcomes for the study area.

To create matrix X, each of the M parameters must be identified, along with their maximum (MAX) and minimum (MIN) range limits. The random sampling of a parameter is carried out automatically, given a probability distribution within the range of **MIN - MAX** values in which the sampling is carried out. The probability distribution of the parameter may be normal, gamma, asymmetric, etc. However, since the actual probability distribution of each parameter is unknown, as is the optimal value, using a uniform distribution is the most common choice when setting up automatic parameter sampling **(Figure 26)**.

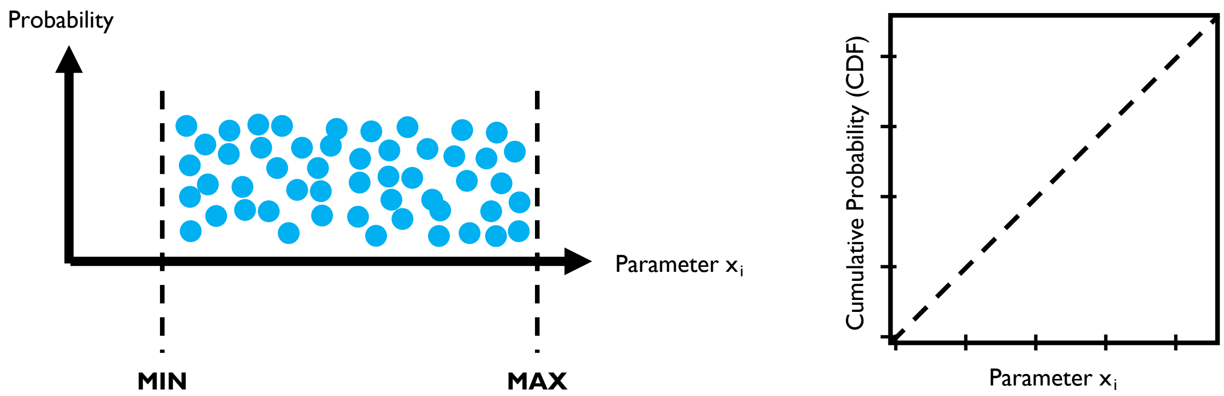


Figure 26. *Concept of automatic parameter sampling. Each blue circle represents a random value within the MIN - MAX range. The probability distribution shown is uniform, that is, any value within the range has an equal probability of being selected. Source: Ochoa–Tocachi et al. (2020a).*

The next step is to define the sampling strategy. When there are multiple parameters, automatic random sampling should produce unique combinations, which "tests" the model by producing wide ranges of outputs. The most common sampling strategies are simple random sampling (i.e., Monte Carlo) and Latin Hypercube Sampling (LHS). **Figure 27** illustrates these two strategies for two parameters. Simple random sampling produces combinations of parameters without considering previously generated values. In this case, it is not necessary to know the number of target combinations (N) beforehand, but there is a risk of repeating values within parameters **(Figure 27)**.

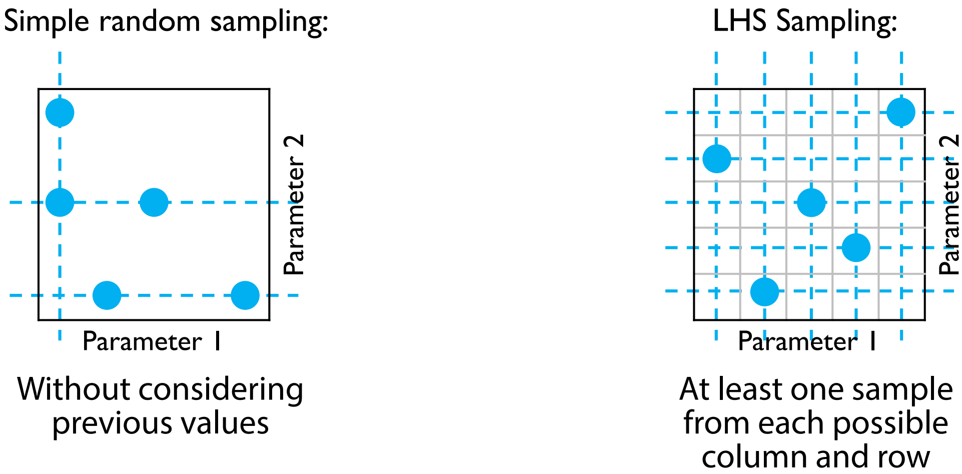


Figure 27. *Comparison of sampling strategies. Each blue point is a combination of parameters 1 and 2. Source: Ochoa–Tocachi et al. (2020a).*

Latin Hypercube Sampling stores the previously sampled values in memory, so parameter values are not repeated. In this case, teams need to know the number of target combinations (N) in advance, so that the parametric space of each parameter is divided proportionally to its probability distribution. For example, for two parameters (M=2, **Figure 27**), means that each sample combination occupies a column (parameter value 1) and a single row (parameter value 2), which can be compared with a chess board, where there are N rooks located so that no one threat or is threatened by another rook. For three parameters (M = 3), this can be represented in three dimensions, using a cube. More than three parameters is conceptually similar, but difficult to represent in a graph.

The configuration generally recommended for automatic parameter sampling is a uniform probability distribution with an LHS strategy.

Performance Evaluation

There are several indicators that can be used to quantify simulation performance. The indicators identified here are widely used objective functions of hydrologic models (Moriasi et al., 2007; Gupta et al., 2009):

- Nash–Sutcliffe efficiency (NSE) **(Table 14)**
- Root mean square error (RMSE) **(Table 15)**
- Ratio between the RMSE and the standard deviation (RSR) **(Table 16)**
- Percentage bias (PBIAS) **(Table 17)**
- Kling–Gupta efficiency (KGE) **(Table 18)**
- Combination of several indicators (OF) **(Table 19)**

Table 14. *Nash–Sutcliffe Efficiency (NSE). Source: Ochoa–Tocachi et al. (2020a).*

INDICATOR	1. NASH–SUTCLIFFE EFFICIENCY (NSE)
Formula	$NSE=1-\left[\frac{\sum_{t=1}^{p}\left(Y_{t,obs}-Y_{t,sim}\right)^2}{\sum_{t=1}^{p}\left(Y_{t,obs}-\mu_{Y,obs}\right)^2}\right]$
Explanation	NSE is a comparison between time series Y_t of observed data $Y_{t,obs}$ and simulated data $Y_{t,sim.}$ The numerator consists of the sum of the squared differences between each observed datum minus each simulated datum. The denominator consists of the sum of the squared differences between each observed datum minus the average (μ) of the observed data. Subtract the division results from one.
Interpretation	The optimal value of NSE is 1 and the worst value tends to –inf (infinite). That is, an NSE value of one (1) implies a perfect fit between the observed data and the simulated data. The value of zero (0) serves as an indicator of how the simulations compare to the average of the observed data. In other words, an NSE value of 0 implies that the simulation is no better than if the observed data had simply been averaged. A hydrologic model is considered satisfactory if NSE > 0.50.
Reference	Nash & Sutcliffe (1970).

Table 15. Root mean square error (RMSE). Source: Ochoa–Tocachi et al. (2020a).

INDICATOR	2. ROOT MEAN SQUARE ERROR (RMSE).
Formula	$$RMSE=\sqrt{\frac{\sum_{t=1}^{P}\left(Y_{t,obs}-Y_{t,sim}\right)^{2}}{P}}$$
Explanation	RMSE estimates the error between time series Y_t of observed data $Y_{t,obs}$ and simulated data $Y_{t,sim}$. The sum of the squared differences between each observed datum minus each simulated datum is obtained. This sum is divided by the number of data P and the square root of the result is calculated.
Interpretation	The optimal value of RMSE is 0 and the worst value tends to +inf. An RMSE value of 0 implies a perfect fit between the observed data and the simulated data. The RMSE units are the same as for the variable in question. There is no standard threshold, but in general, a hydrologic model is considered satisfactory the smaller the RMSE is.
Reference	Singh et al. (2004).

Table 16. Ratio between the RMSE and the standard deviation (RSR). Source: Ochoa–Tocachi et al. (2020a).

INDICATOR	3. RELATIONSHIP BETWEEN RMSE AND STANDARD DEVIATION (RSR)
Formula	$$RSR=\frac{RMSE}{STDEV_{obs}}=\frac{\sqrt{\left[\sum_{t=1}^{P}\left(Y_{t,obs}-Y_{t,sim}\right)^{2}\right]}}{\sqrt{\left[\sum_{t=1}^{P}\left(Y_{t,obs}-\mu_{Y,obs}\right)^{2}\right]}}$$
Explanation	RSR combines the RMSE with the standard deviation of the observed data. This allows normalizing the value and granting relative ranges that can be considered satisfactory.
Interpretation	The optimal value of RSR is 0 and the worst value tends to +inf. That is, an RSR value of 0 implies a perfect fit between the observed data and the simulated data. A hydrologic model is considered satisfactory if RSR <= 0.70.
Reference	Singh et al. (2004).

Table 17. *Percentage Bias (PBIAS). Source: Ochoa–Tocachi et al. (2020a).*

INDICATOR	4. PERCENTAGE BIAS (PBIAS)
Formula	$PBIAS = \left[\frac{\sum_{t=1}^{p} (Y_{t,obs} - Y_{t,sim}) * 100}{\sum_{t=1}^{p} (Y_{t,obs})} \right]$
Explanation	PBIAS measures the average trend of simulated data $Y_{t,sim}$ to be relatively higher or lower than the correspondent observed data $Y_{t,obs}$. It is obtained from the sum of the differences between the observed data and the simulated data multiplied by 100 (to convert them to a percentage) and divided by the sum of the observed data.
Interpretation	The optimal value of PBIAS is 0 and the worst values tend to both –inf and +inf. A PBIAS value of 0 implies a perfect fit between the observed data and the simulated data. Positive PBIAS values indicate underestimation bias of the model. Negative PBIAS values indicate overestimation bias of the model. A hydrologic model is satisfactory if PBIAS <= ±25 % for flow simulation, ±55 % for sediment simulation, and ±70 % for water quality simulation, such as nutrients (nitrogen, phosphorus).
Reference	Gupta et al. (1999).

Table 18. *Kling–Gupta Efficiency (KGE). Source: Ochoa–Tocachi et al. (2020a).*

INDICATOR	5. KLING–GUPTA EFFICIENCY (KGE)
Formula	$KGE = 1 - \sqrt{(r-1)^2 + \left(\frac{\sigma_{Y,sim}}{\sigma_{Y,obs}} - 1\right)^2 + \left(\frac{\mu_{Y,sim}}{\mu_{Y,obs}} - 1\right)^2}$
Explanation	KGE was conceptualized as an enhancement to NSE. When deconstructing the NSE, three elements are obtained corresponding to the linear correlation (Pearson's coefficient r), the relationship between standard deviations (σ) and the relationship between the means (μ). Each component is subtracted one and squared. The root of the sums of these differences is obtained. The result of the square root is subtracted from one.
Interpretation	The optimal value of KGE is one and the worst value tends to –inf. That is, a KGE value of one implies a perfect fit between the observed data and the simulated data. The value of 0 serves as an indicator of how the simulations compare with the average of the observed data. A KGE value of 0 implies that the simulations are no better than if the observed data had simply been averaged. A hydrologic model is considered satisfactory if KGE > 0.50.
Reference	Gupta et al. (2009).

Table 19. *Combination of several indicators (OF). Source: Ochoa–Tocachi et al. (2020a).*

INDICATOR	6. COMBINATION OF SEVERAL INDICATORS (OF)
Formula	$OF = \sqrt{Ind_{norm_1}^{\ 2} + Ind_{norm_2}^{\ 2} + \cdots + Ind_{norm_n}^{\ 2}}$ Where: $Ind_{norm_i}^{\ 2} = \dfrac{Ind_i - min(Ind_i)}{max(Ind_i) - min(Ind_i)}$
Explanation	It is not easy to choose which is the best simulation when there are several performance indicators. Sometimes a simulation performs very well for one indicator and poorly for another. In this case, the result of the different indicators can be combined into one with the indicated formula. For the case of NSE and KGE, whose optimum values are one, we would use (1 – NSE) and (1 – KGE). For PBIAS, which can be positive or negative, the absolute value abs(PBIAS) is used. For indicators that have zero optimum (RMSE, RSR) the indicator is used directly. Before combining the indicators, it is necessary to normalize them so that their different magnitude does not generate greater weight for one or the other.
Interpretation	The optimal value of this combination is 0 and the worst value tends to +inf. An OF value of 0 implies that the same set of parameters is the best for all performance indicators. The best simulation is the one with the lowest OF.
Reference	Ochoa–Tocachi et al. (2019c).

Simulations that comply with defined limits for each indicator can be deemed acceptable **(Table 20)**. According to Moriasi et al. (2007), a simulation model can generally be judged as satisfactory if NSE > 0.50 and RSR < 0.70, and if PBIAS < ±25% for flow modeling, PBIAS < ±55% for sediment modeling, and PBIAS < ±70% for water quality modeling, such as nitrogen and phosphorous. It can also be considered satisfactory if KGE > 0.50, due to its similarity to NSE.

Several combinations of parameters can meet the acceptability criteria and any can be used (principle of ***"equifinality"***, Beven, 2000). By using a set of results, rather than a single result, it is possible to obtain confidence/uncertainty intervals. If a single combination of parameters is required, the best simulation is the one with the lowest OF value (Ochoa–Tocachi et al., 2019c).

Table 20. *Simulation acceptability criteria based on performance. Performance indicators are calculated for each of the N parameter combinations. Source: Moriasi et al. (2007); Gupta et al. (2009); Ochoa–Tocachi et al. (2020a).*

E	NSE	RMSE	RSR	PBIAS	KGE	OF	Interpretation
1	< 0.00	high	> 1.00	> ±50%	< 0.00	high	very poor
2	< 0.50	medium	> 0.70	> ±25%	< 0.50	medium	unacceptable
3	> 0.50	low	< 0.70	< ±25%	> 0.50	low	acceptable
4	~1	~0	~0	~0	~1	~0	best simulation
…	…	…	…	…	…	…	
N	…	…	…	…	…	…	

Calibration in Ungauged Basins

In those cases where there are no local flow data or other variables to calibrate a model, it is possible to use methods for "predictions in ungauged basins" (PUB). The PUB methods have been researched for over a decade in hydrologic science and support the modeling process and hydrologic variable estimation in the face of data scarcity (Ochoa-Tocachi *et al.*, 2016b). Common methods include the regionalization of model parameters, where a model is calibrated for one or several catchments located nearby or similar to the objective catchment, and then the calibrated parameters are used to model the ungauged catchment. Another method consists of the regionalization of hydrologic indices, where indices (or streamflow signatures), such as the runoff ratio, baseflow index or the flow duration curve can be used as the objective functions to measure model performance (instead of indicators such as NSE, PBIAS, etc.) and calibrate the model to adjust the results to the estimated values to the hydrologic indices. Also, hydrologic indices can be calculated or calibrated for nearby or similar catchments to the objective catchment, and then use them as objective functions for the ungauged catchment. Several other PUB methods exist to deal with the issue of calibration in ungauged basins. Lastly, it is necessary to consider than the calibration in ungauged basins using PUB methods is more uncertain than the more conventional hydrologic calibration. It needs with be reported if this is the case to clarify the level of uncertainty associated with the hydrologic model results.

Validation

Whereas calibration is a process of systematically adjusting parameters values, model ***validation*** is used to evaluate the confidence of those calibration results (i.e., calibrated model performance). In general, the model is validated using a different period than the one used in calibration (Vaze et al., 2011). For example, the hydrologic station with the lowest elevation or located at the closure of the modeled catchment is generally used to calibrate and validate a flow model.

According to Klemeš (1986), the separation of the daily data period available for evaluation (short time series) correspond to 60% for calibration and 40% for validation, regardless of the length of the flow time series used **(Figure 28)**. It is important that the time series show the existence of both wet and dry periods. At the start of model execution, a ***"warm up"*** period is discarded. After this period, the output variables of the model become independent or attenuate their effect due to the initial conditions assumed (Mazzilli, Guinot, & Jourde, 2012).

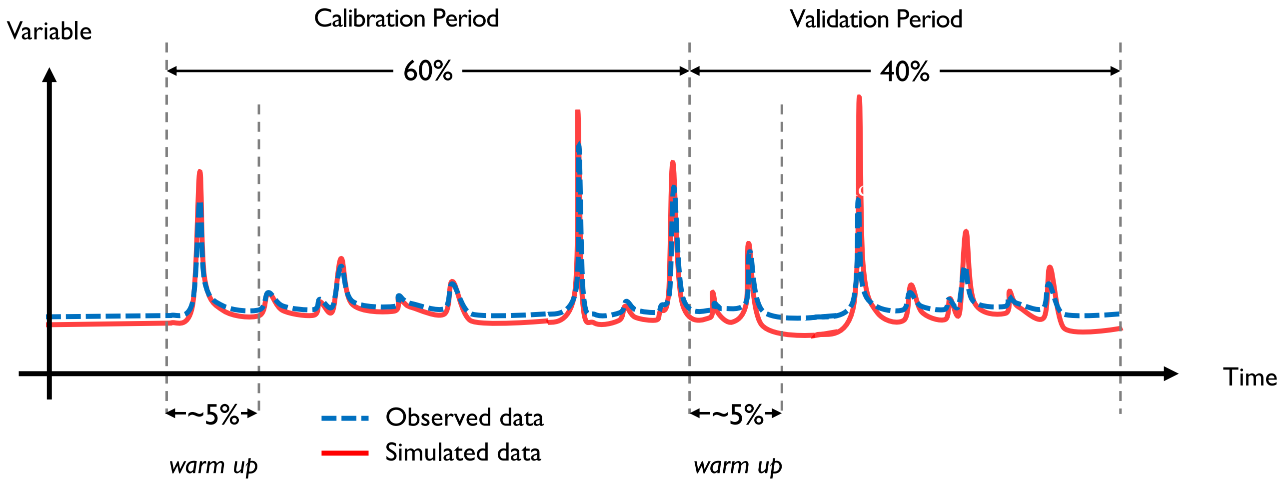


Figure 28. *Division of the time series into calibration and validation periods and warm-up periods. Source: Prepared by the authors.*

Calibration and validation can only be performed on the baseline scenario, which reflects historical conditions and is the actual counterpart to the observed data. It is necessary not to confuse the use of hypothetical scenarios during the calibration and validation exercise, to avoid biases and errors in the hydrologic modeling process. It is important to emphasize the need to calibrate and validate hydrologic models during project design and evaluation applications. If the models are not calibrated, it is possible that the results are not realistic and could generate confusion and poor recommendations during the decision-making process. It is common to encounter situations where data quality is insufficient to proceed with calibration and validation. However, even with scarce, specific, or weak data, it can substantially improve the simulations obtained with hydrologic models (Winsemius et al., 2009). In the absence of complete data, a hydrologic model can be calibrated using data for another geographic area (ideally nearby) that has conditions as similar as possible to the desired study area. This option should be considered for the most challenging situations.

STEP 9a. SENSITIVITY ANALYSIS

Manual and Automatic Sensitivity Analysis

A sensitivity analysis is used to evaluate how model outputs change in response to variations in the inputs. The degree of this change is known as ***"sensitivity".*** A traditional method is to make variations (±5%, ±10%, ±20%, etc.) to one parameter and observe how the results change because of those variations without modifying the other parameters **(Figure 29).** However, this method is very "expensive" computationally because variations are defined discretely and must usually be performed manually.

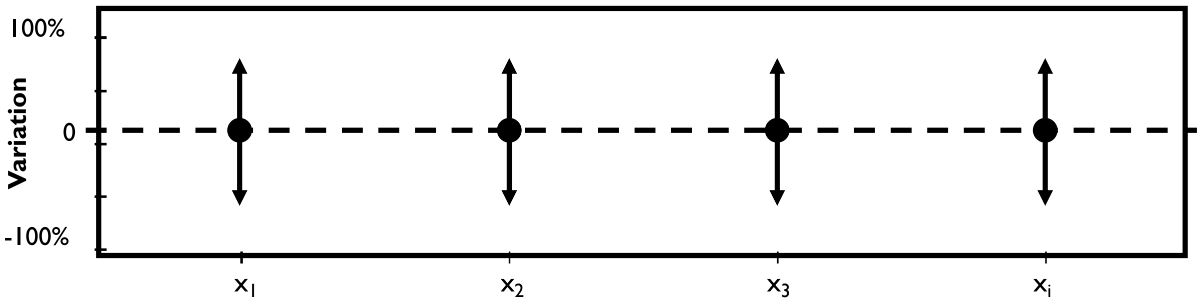


Figure 29. *Manual sensitivity analysis of a simulation model. Source: Ochoa–Tocachi (2020).*

Following the automatic modeling scheme proposed before **(Figure 25)**, the sensitivity analysis process is performed with the simulation results of the model (stage 3); that is, the sensitivity analysis replaces the performance analysis **(Figure 28).**

Matrix Y of simulated results is entered to evaluate how the performance of each simulation line varies against the set of observed data, or to observe how variations (or restrictions) in parameter X influence the outputs of the model. Instead of analyzing the results of each parameter variation one by one, the N results are analyzed together to produce vector **S** of size 1 x M, where M is the number of parameters analyzed and S contains the ***"sensitivity indicators"*** for each parameter. These indicators show: a) the parameters for which the model is ***"most sensitive"***; or b) the parameters that are the ***"most influential"*** in the model outputs. The process can be carried out one time, or iterated several times to obtain confidence intervals or ranges for the sensitivity indicators and increase confidence in how sensitive or influential parameters are. The advantage of an automatic sensitivity analysis is that it is N times more efficient than manual analysis and can yield optimal results in a single run. It also makes use of the same data (parameter matrices and simulations) that are used in automatic calibration, helping guarantee continuity and consistency during a modeling process.

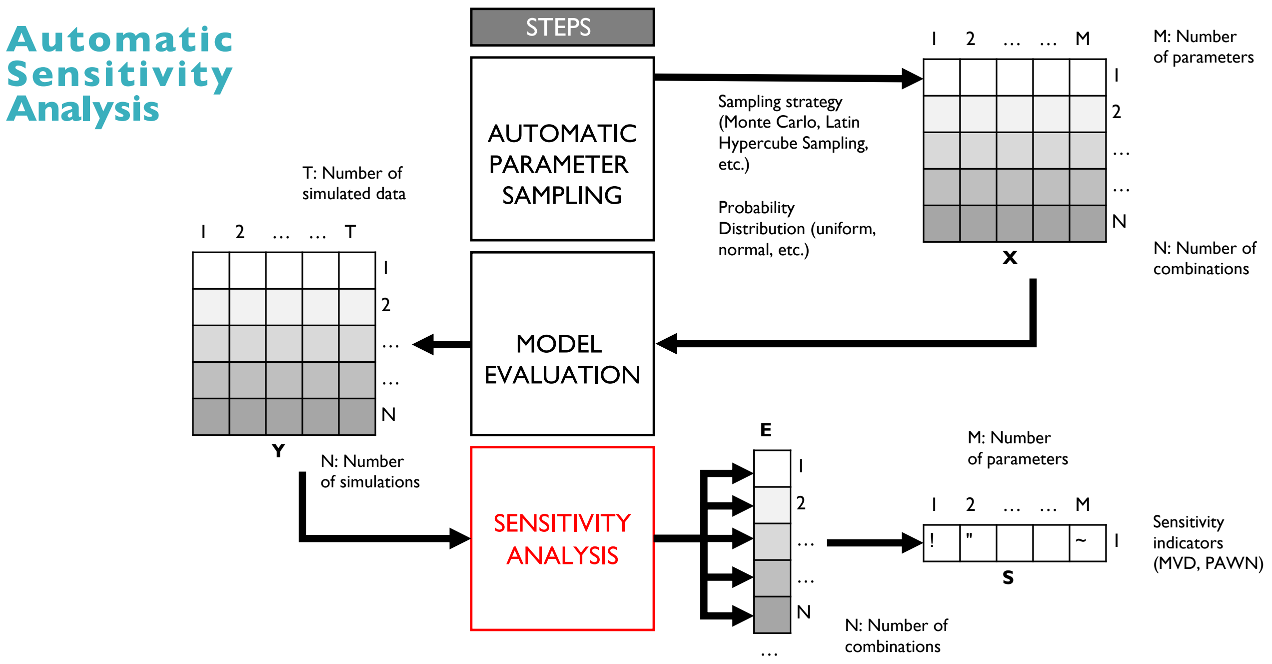


Figure 30. *Automatic sensitivity analysis of a simulation model. Source: Ochoa–Tocachi et al. (2020b)*

"A sensitivity analysis is used to evaluate how model outputs change in response to variations in the inputs."

There are several sensitivity analysis methods. This document focuses on two complementary methods: regional sensitivity analysis (RSA: Sieber & Uhlenbrook, 2005; Spear & Hornberger, 1980) and sensitivity analysis based on output distribution (PAWN: Pianosi & Wagener, 2015, 2018).

RSA: Regional Sensitivity Analysis

The ***regional sensitivity analysis*** method seeks to identify the influence of the parameters on the performance of the model. As indicated above, in stage 1 of the modeling scheme, random sampling of a parameter is performed within a range of MIN–MAX values, usually with a uniform probability distribution. This corresponds to a 1:1 line in the probability graph **(Figure 26)**. These different N parameter values produce better or worse results in a performance indicator (blue dots in **Figure 31**). To evaluate the performance of a model, we generally use thresholds that determine when a simulation is acceptable (***"behavioral"),*** or unacceptable (***"non-behavioral"***) (**Table 20**). The initial set of parameters is divided between these two groups, depending on how each of the performance rows of matrix E is better or worse than the threshold (red dots in **Figure 31**).

$$E_{i,j} < \text{threshold}(j)$$

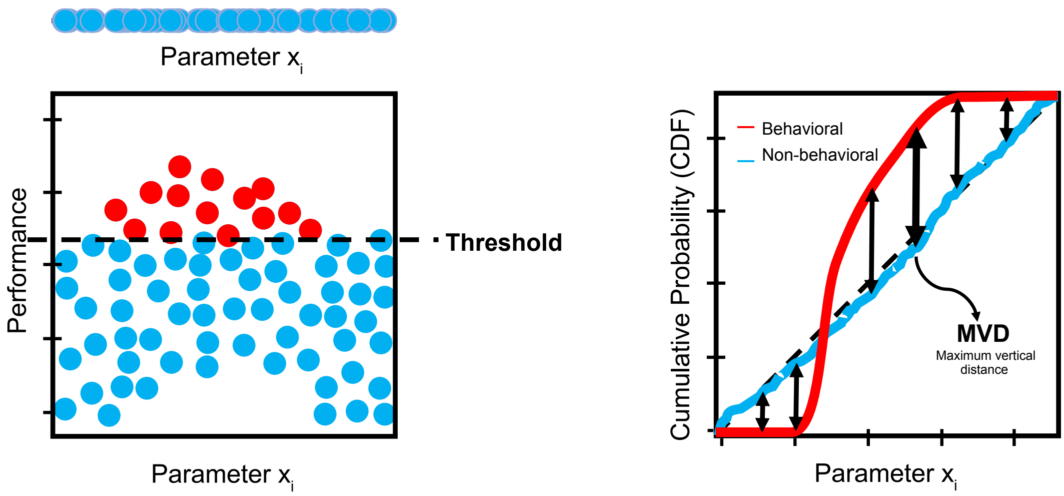


Figure 31. *Conceptual regional sensitivity analysis (RSA). Source: Ochoa–Tocachi et al. (2020b).*

These two subsets of parameters also have their own probability distributions, which can be visualized using cumulative probability curves (CDF curves, **Figure 31**). This method evaluates the differences between the CDFs of the "behavioral" versus the "non-behavioral" parameter combinations. If CDFs differ significantly, then it is understood that "the performance of the model is sensitive to the parameter." To measure this objectively, a quantitative measure is used, such as the Kolmogorov-Smirnov index (Kolmogorov, 1933, Smirnov, 1939), which determines the ***maximum vertical distance*** (MVD) between the two CDF curves.

$$MVD(x_i) = \max_{x_i} \left| CDF_{xi|Ybehavioral}(x_i) - CDF_{xi|Ynon\text{-}behavioral}(x_i) \right|$$

This exercise is performed for each parameter (xi), which reflects how sensitive the performance of the model is to the parameter. The exercise is repeated using resampling or bootstrapping (Efron, 1979; Efron & Tibshirani, 1993). ***Bootstrapping*** is a method that approximates the uncertainty in the calculation of matrix S of sensitivity indicators, using subgroups of parameter combinations that are used iteratively when calculating the indicator. If during each recalculation, the indicator changes its value, then a range of uncertainty is generated around that indicator. Generally, this is performed to generate indicators with a confidence interval of 95% **(Figure 32)**.

MVD is an absolute measure, which means it has a meaningful per se value, regardless of the units of the matrix X (parameters) or E (performance objective function, such as NSE, PBIAS, etc.). By definition MVD is between zero and one **(Figure 32)**:

- The higher the MVD indicator of a parameter, the higher the sensitivity of the model for that parameter;
- If MVD is equal to 0, the model is not sensitive to the parameter; in other words, the two CDF curves (between the behavioral and non-behavioral parameter combinations) are exactly the same;
- If MVD is equal to 1, the model is very sensitive to the parameter; the two CDF curves are "mutually exclusive" (the same value of a parameter has a probability of 0 in one CDF and 1 in the other).

If no values are found using a certain threshold, the threshold should be changed **(Figure 32)**.

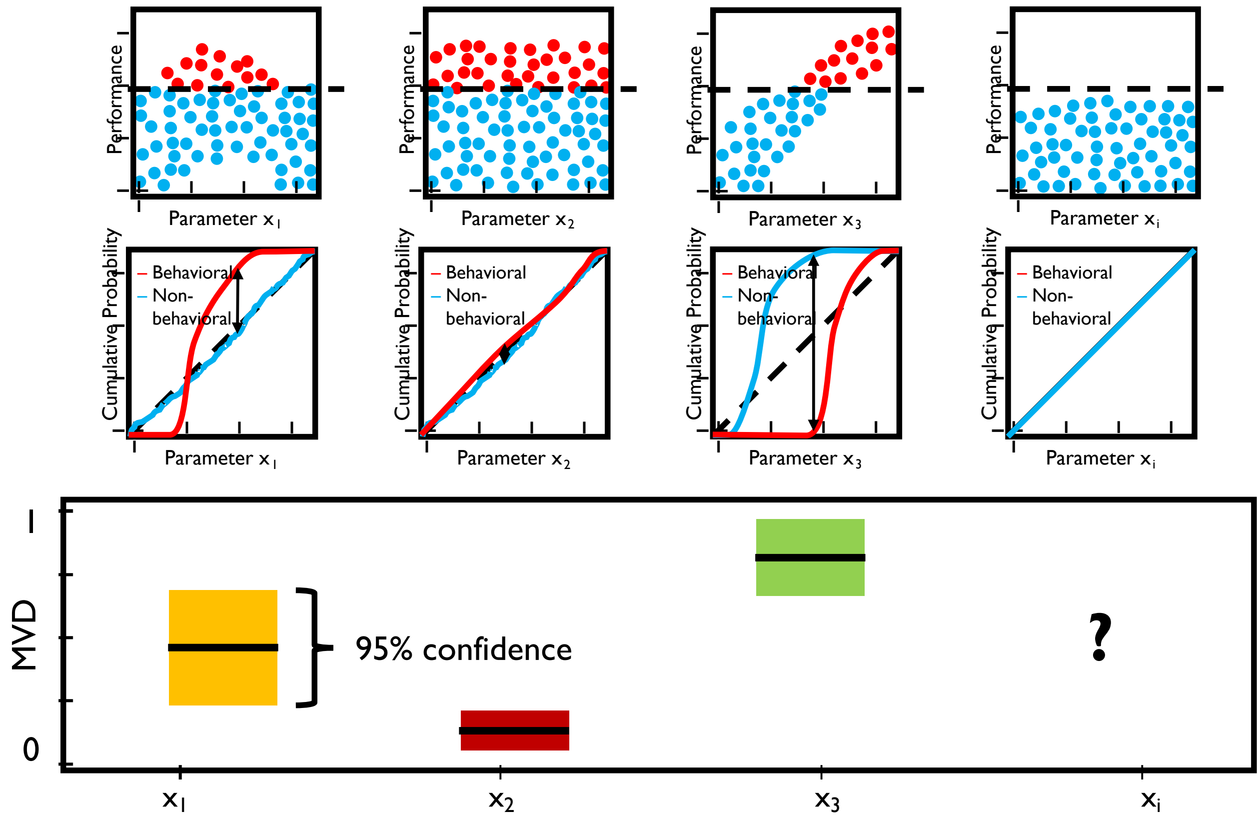


Figure 32. *Examples of MVD sensitivity indicator results & product of the regional sensitivity analysis. Source: Ochoa–Tocachi et al. (2020b).*

It is not always possible to have a performance indicator ("behavioral" and "non-behavioral") to proceed to a sensitivity analysis based on performance. For example, when there are no observations against which to validate the model. Sometimes, it is only required to know which parameters influence the result of the model, regardless of whether they are good or bad. Another factor is that the RSA produces an analysis between results that are "behavioral" and "non-behavioral," resulting from the combined interactions of all the parameters. However, sometimes it is required to simply know the impact of modifying a single parameter or not, which can be done using the PAWN method.

PAWN: Sensitivity based on Output Distribution

The ***PAWN method*** (Pianosi & Wagener, 2015, 2018) divides the outputs of the N combinations of parameters into groups: unconditional (red dots in **Figure 33**) and conditional (gray subsets in **Figure 33**) to a parameter xi analyzed. The ***"unconditional"*** subset is a subsampling of the original matrix X that seeks to maintain the variability in the combinations of the M parameters, so that the cumulative distribution function (CDF) of the model outputs is as close as possible to the total cumulative CDF. Note that this method involves the probability distribution of the output of the model **(Figure 33)** not that of the input parameters (as is the case with the RSA method).

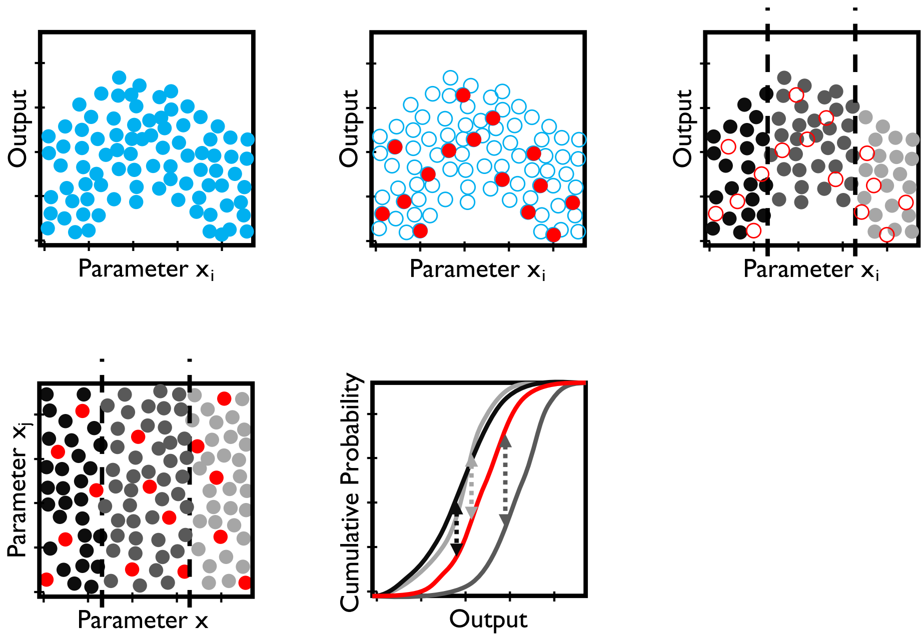


Figure 33. *Conceptual sensitivity analysis based on output distribution (PAWN). Source: Ochoa–Tocachi et al. (2020b).*

After removing the "unconditional" subset, the remaining combinations are divided into groups ***"conditional"*** to the parameter xi analyzed. These groups represent what would happen if the value of a parameter were specifically known and "fixed" in the known value, or to represent the effect of reducing the sampling rate of the parameter to a much narrower size. These conditional subsets are obtained only for the parameter xi analyzed, while leaving the other M-1 parameters to vary freely within their possible range **(Figure 33)**. In other words, "uncertainty is eliminated" around parameter xi to observe how the outputs of the model change without said uncertainty.

The probability distributions of the outputs of the model are determined next, resulting from groups unconditional and conditional to xi **(Figure 33)**. Likewise, to quantify we use the maximum vertical distance (Kolmogorov – Smirnov index) between the unconditional CDF curve and the conditional CDF curves.

$$KS(x_i) = \max_{y} \left| CDF_y(y) - CDF_{y|xi}(y) \right|$$

Finally, this procedure results in a set of KS values equal to the number of conditional subgroups generated (e.g., n = 10). Pianosi & Wagener (2015, 2018) recommend using a statistic (e.g., median or maximum) of this KS set on all possible values obtained by conditioning xi to derive the PAWN sensitivity indicator.

$$PAWN(x_i) = \underset{x_i}{stat}[KS(x_i)]$$

By definition PAWN is in a range between 0 and 1 **(Figure 34)**:

- The higher the PAWN indicator of a parameter, the higher is the influence of that parameter for the model.

- If PAWN is equal to 0, then the parameter has no influence on the output of the model.

The impact of errors in numerical approximations is estimated using a ***"dummy"*** parameter that, in principle, should not affect the variability of the model output. The PAWN sensitivity value for the dummy parameter is used to provide context for the rest **(Figure 34)**:

- If the PAWN for a parameter xi is much higher than the dummy PAWN, the parameter is "influential."

- If the PAWN for a parameter xi is equal to or less than the dummy PAWN, the parameter is "not influential."

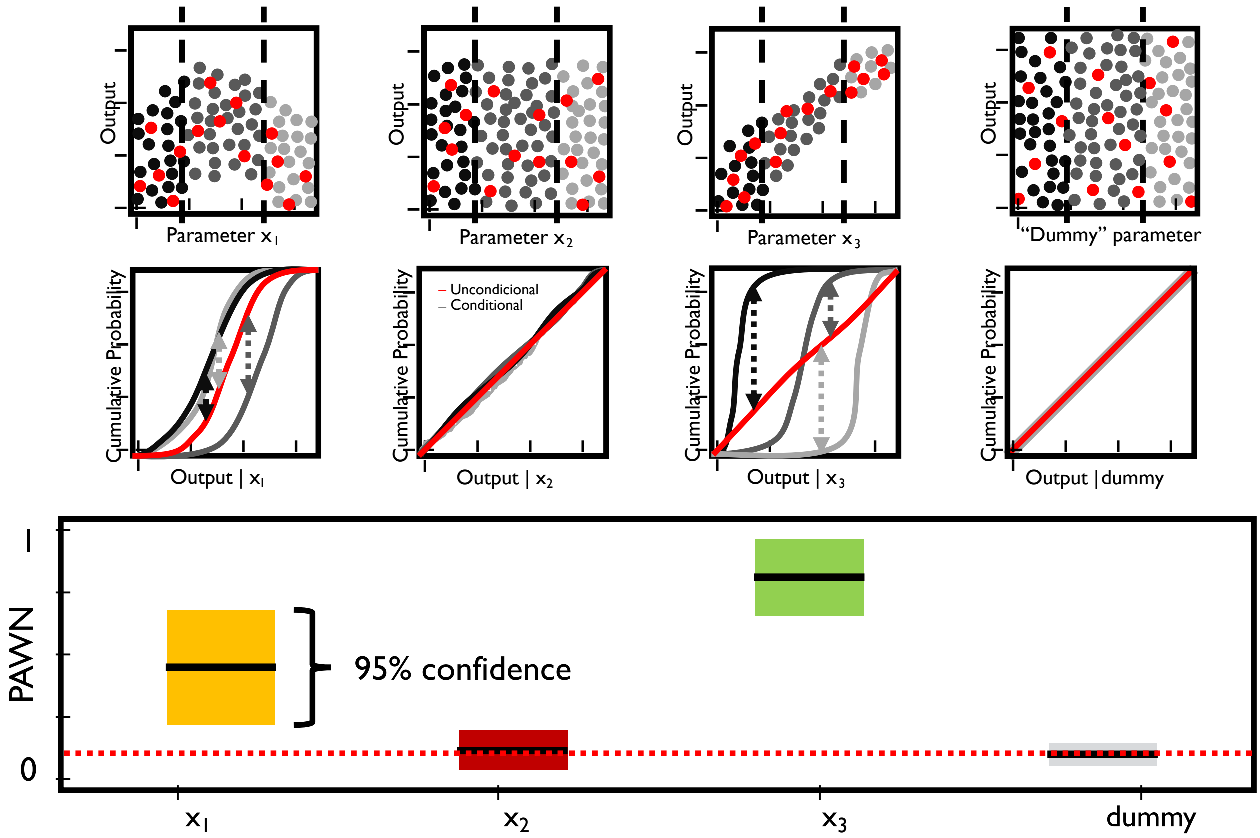


Figure 34. *Examples of results of PAWN sensitivity indicators & product of the sensitivity analysis based on output distribution. Source: Ochoa–Tocachi et al. (2020b).*

This exercise is repeated for each parameter xi, which reflects how influential parameter xi is on the valued output of the model. The exercise is repeated using bootstrapping to provide indicators with a confidence interval, usually 95% **(Figure 34)**.

Differences between the RSA and PAWN Methods

Both RSA and PAWN use CDF for sensitivity analysis, but differently **(Table 21)**.

Table 21. *Differences between the RSA and PAWN sensitivity analysis methods. Source: Ochoa–Tocachi (2020).*

RSA	PAWN
RSA determines how sensitive the performance of the model is for a parameter xi.	PAWN determines how influential a parameter is on the resulting outputs of the model.
RSA focuses on how the probability distributions of an input parameter xi changes while varying the output of the model ("behavioral" or "non-behavioral"). $E_{i,j} < \text{threshold}(j)$	PAWN focuses on how the probability distribution of the model output changes while conditioning an input parameter xi (by removing the uncertainty around the parameter).:
RSA quantifies differences in the CDFs of the model inputs (model parameters), using the equation: $MVD(x_i) = \max_{x_i} \lvert CDF_{xi\lvert Ybehavioral}(x_i) - CDF_{xi\lvert Ynon\text{-}behavioral}(x_i) \rvert$	PAWN quantifies differences in the CDFs of the model outputs (output variable or objective function), using the equation: $KS(x_i) = \max_{y} \lvert CDF_y(y) - CDF_{y\lvert xi}(y) \rvert$ The PAWN sensitivity indicator is obtained from a statistic (e.g., maximum, median, minimum) of all the KS values obtained for parameter xi. $PAWN(x_i) = \mathrm{stat}_{x_i}[KS(x_i)]$
RSA requires performance indicators to determine subsets of parameter combinations that are "behavioral" or "non-behavioral" to a performance criterion.	PAWN does not require performance indicators, but divides the parameter combinations into "unconditional" and "conditional" subsets to a parameter xi.
RSA makes it possible to differentiate between those parameters that effectively produce good simulations, to obtain combinations of parameters that are acceptable in a calibration process.	PAWN does not allow differentiation between which combinations of parameters produce good or bad results. It identifies which parameters are the most influential to put more emphasis on determining or measuring their optimal values in a calibration process.

Usefulness of a Sensitivity Analysis

A sensitivity analysis can serve as a decision-making tool to forecast the success or failure of NI project alternatives in a transparent and efficient manner. Studying how the different variables influence possible outcomes can help decision-makers make better choices between project alternatives, such as which is more beneficial (or less harmful) to the NI and ecosystem services produced. Likewise, the sensitivity analysis makes it possible to multiply the results of the modeled scenarios, so that sub-scenarios can be considered as a product of the individual and group variations of the model parameters for each of the modeling scenarios. The sensitivity analysis also makes it possible to identify the pros and cons of different alternatives, as well as to carry out early prospects of the expected intervention results. This is particularly useful for defining more realistic goals for project success or failure indicators, based on extreme modeling scenarios for NI interventions.

Sensitivity analyses also allow errors and problems to be identified in the modeling process itself. If a model produces counterintuitive results due to changes in a parameter, it is possible that the equations of the model, its structure, or its conceptual or perceptual basis are mistaken. For example, the infiltration and storage capacity of water in the soil is an important factor for hydrologic regulation. If parameters in the model that these properties represent are altered, results are expected to respond along this conceptual understanding. An increase in infiltration capacity and soil storage should result in a reduction in flood flows and an increase in base flows. If not, it is advisable to review what is happening in the water balance calculations that could be producing odd results.

Similarly, a sensitivity analysis makes it possible to identify whether a model is the best tool for the exercise in question. If a NI intervention modifies certain known characteristics of the landscape, the model must be able to represent those changes. If the results of the model are not sensitive to changes in the parameters that represent these characteristics, it is advisable to consider other models in which this sensitivity is higher. This is not to maximize the differences due to the changes produced by the interventions, but to use a model that is evidence-based and capable of robustly linking the changes produced by the interventions with the prioritized ecosystem services.

Finally, a sensitivity analysis allows optimizing the allocation of resources, either in modeling or in the generation of information in the field. If there are parameters that are highly influential on the model results, it is advisable to focus calibration and validation efforts on those parameters to obtain the best estimates of realistic values. If those sensitive parameters are very relevant for the NI modeling exercise, it is advisable to invest in the generation of observed data in the field for those parameters, so that they can be used to improve results. Thus, a sensitivity analysis becomes a useful tool to improve on-site monitoring efforts. Similarly, it is not recommended that teams invest too much time and resources calibrating parameters to which the model is not sensitive.

A sensitivity analysis can be used as:

- **A decision-making tool.** The sensitivity analysis produces forecasts supported by data from different modeling scenarios or sub-scenarios generated by the different values assigned to the variables and input data to the model.

- **A prospecting tool.** The sensitivity analysis can help set realistic hydrologic goals to define reliable expectations resulting from interventions in indicators of proposed results.

- **An error identification tool.** With the sensitivity analysis, it is possible to identify whether a model generates results that are counterintuitive to the hydrologic theory or expectation and to studying the effects that the variations of the parameters have on the expected trends of the results obtained.

- **A quality control tool.** A sensitivity analysis reveals whether the hydrologic model used is truly capable of capturing and representing the impacts of the interventions on the expected indicators. If a model is not sensitive to important features, it may not be as useful.

- **A resource optimization tool.** The sensitivity analysis makes it possible to identify the parameters and data that have the greatest influence or impact on the results. This allows teams to focus efforts on the most influential parameters through calibration, validation, and data collection in the field.


Photo: Carlos Palacios Núñez

STEP 9B. UNCERTAINTY ANALYSIS

An ***uncertainty analysis*** investigates the effect of errors in the variables relevant to decision-making problems and quantifies errors in the relevant variables. The ***"uncertainty"*** refers to anomalies or errors resulting from imperfect or unknown information. It applies to event predictions, physical measurements that have already been made, or the unknown. Uncertainty is also associated with approximation errors or numerical errors in the equations of the model, compared to the actual or theoretical values expected.

> ***"An uncertainty analysis investigates the effect of errors in the variables relevant to decision-making problems and quantifies errors in the relevant variables."***

Due to various sources of uncertainty, it is not possible to determine exact absolute results of the quantity and quality characteristics of the hydrologic ecosystem services produced by the interventions on the NI under future scenarios (time T2 in **Figure 35**). To generate a range of assumptions with a range of potential outcomes, the model must be able to process uncertainty from various sources (e.g., the input information), the structure of the model itself, and the natural variability of the basin. A probability function can be used to assimilate these possible realizations **(Figure 35)**, from which we can estimate what is the most likely result and what the determined variability range with a certain confidence level, usually 90% or 95%. This confidence level is determined by the probability that the estimated value of the analyzed variable is within a certain range or ***"confidence interval"***.

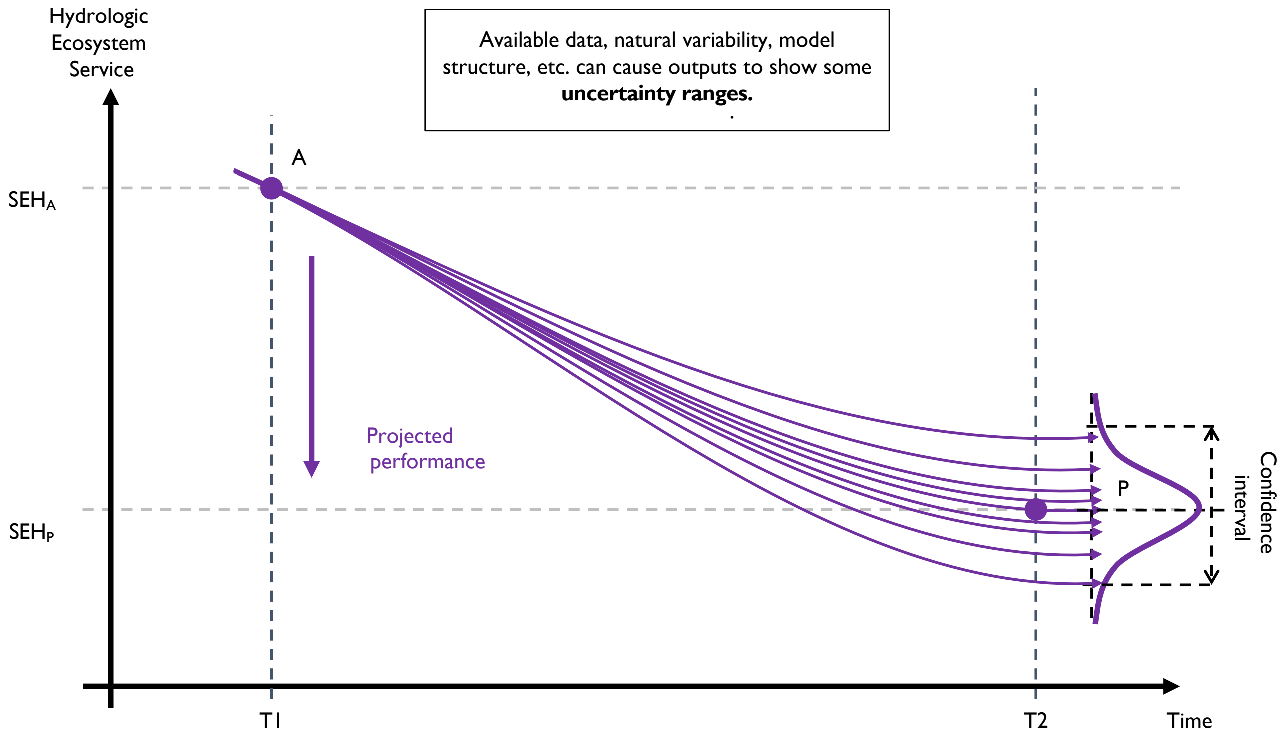


Figure 35. *Uncertainty in the projections of a hydrologic ecosystem service performance in the future. Source: Ochoa–Tocachi et al. (2018b).*

It should be clarified that uncertainty is not unique to NI modeling. Often, conventional gray infrastructure analyses use results from manually calibrated models, sometimes ignoring or assuming that there is no uncertainty. Hydrologic modeling should be seen as a useful but uncertain tool. To use it wisely in decision-making, it is necessary to clarify sources of error, critically analyze the results, and recognize that there is a chance that actual values are outside the confidence interval obtained during modeling.

Sources of Uncertainty

Carrying out simulations inherently involves managing uncertainty. It is impossible to eliminate uncertainty completely because it comes from several sources. Model projections or ex-ante evaluations require assumptions to be made about how the variables of the hydrologic system will change in the future as a result of the proposed NI interventions – each of which can be informed by different sources and are associated with different levels of uncertainty (e.g., climate, types and extent of vegetation, hydrologic parameters, human intervention and management). For example, an ex-ante simulation may consider that the water infiltration rates into the soil will increase by 10% as a result of the proposed interventions in an area of interest. This assumption may have different degrees of confidence and be based on several sources: 1) qualitative information based on ***"expert criteria"*** and experience; 2) secondary quantitative information obtained from published documents, reports from other project sites, or scientific publications; or 3) primary information obtained at the site of interest using standardized methodologies for data collection. No matter the sources and soundness of assumptions, scenarios are uncertain projections of the future conditions of a study area. Any prediction of the future is inherently uncertain.

"Any prediction of the future is inherently uncertain."

There are still large gaps in scientific knowledge about the changes produced by the effect of NI interventions (e.g., see systematic reviews by Bonnesoeur et al., 2019; Locatelli et al., 2020; Willems et al., 2021; Molina et al., 2021; Mosquera et al., 2022). For many of the hydrologic ecosystem services, this is partly because resulting changes or initial baseline conditions have not been sufficiently monitored or studied. A limited set of studies indicate whether an impact causes an increase or decrease in the production of ecosystem services, but not the exact magnitude of the changes. Even when the studies do quantify impacts, there are so many complex social, physical, and biological interactions in the study area that it can be impossible to distinguish cause and effect relationships. In these cases, it is necessary to generate primary information about the site of interest across factors to attempt to increase clarity. Finally, it is difficult to extrapolate information to sites of interest, even if there is sufficient data to complete a hydrologic model. Studies are only able to collect information under specific time and space constraints and controlled scales and conditions, such as vegetation cover, soil type, and climate.

Even when there is confidence in the information entered in a hydrologic model, the model itself can introduce errors in the calculations that originate from two main factors: 1) errors inherent in the formulation of the problem; and 2) consequence of the method used to find the solution to the problem. The equations used are simplified representations of the actual real system and thus sacrifice accuracy in such representation. Teams must compromise between more complex equations that may be a closer representation of the system and equations that are more versatile and user-friendly to accommodate computational, technical, or time constraints they are operating under. However, as previously discussed, equations of any kind will not be perfectly accurate due to approximation errors or numerical errors during the calculation process. The approximation error or numerical error is a measure of the adjustment or calculation of a magnitude with respect to the actual or theoretical value of that said magnitude. The numerical stability of approximation errors refers to how the error is created within the algorithm itself. Monitoring this source of error within the model itself is essential to generating a "confidence interval" or ***"degree of uncertainty"*** or the process.

Formulation error is the difference between the actual value of a variable and its approximate value (the product of the model simulation). This could stem from imprecision in the physical data (e.g., physical constants, measured data not being accurate due to the instruments used). For example, the exact measurement of the water level in a river can be 24.5 centimeters (cm), but if a ruler does not measure decimals, the value will be rounded to 25 cm (i.e., rounding error). Similarly, if an observer introduces human errors into the measurements, such as typos or omissions, they will be assimilated into the model. These types of errors are generally random in nature, and their analytical treatment is essential to contrast the result obtained computationally.

There are three main sources of ***computational error:***

a) Mistakes in carrying out operations where the modelling system calculates a wrong result and recognizes it as correct. The presence of undetected ***"bugs"*** in the model software can cause errors of this type. These errors are called **"bulk errors"**.

b) Solving the problem by means of some type of approximation rather than how it has been formulated can cause **"truncation error"**. For example, the approximation of an integral by means of a finite sum of the values of a function, or the resolution of a differential equation by replacing the derivatives by a finite difference approximation. In other words, interrupting an infinite process to replace it with a finite one.

c) Some model calculations require numbers with infinite decimals to be represented correctly; however, it is often necessary to round them up to use them (e.g., replace the number π with 3.14), causing a **"rounding error"**. Even some arithmetic operations can cause errors; for example, divisions sometimes produce numbers that should be rounded, while multiplications could lead to more digits than can be stored in the memory of the computer.

There are several factors that contribute to the predictions made in a hydrologic model, making it convenient to classify the sources and estimate uncertainty by category. The total uncertainty of the predictions of a model can be shaped by a variety of uncertainty sources (e.g., Gupta et al., 2005) and be classified as follows (Montanari, 2007; Solomatine & Wagener, 2011):

a) **Perceptual model uncertainty.** How the hydrologic system is translated into the conceptual (numerical) model (see section 1.4, "Development of Hydrologic Models"). A common example in the Andes is the way in which bofedales are represented in models. While there might be consensus that a bofedal acts as a reservoir, the water it stores can come from ground or surface water. Modeling with the assumption that the water stored in a bofedal is exclusively surface water could cause problems later in the process or produce inconsistencies with observed values.

b) **Data and scenario uncertainty.** Applies to both observed and unobserved data. Simple examples include uncertainties in precipitation and flow measurement and precipitation interpolation. On the other hand, Uncertainty of unobserved data can stem from the assumptions a team makes when constructing future modeling scenarios.

c) **Parameter estimation uncertainty.** It is the inability to locate a single set of best model parameters based on the available information, also known as equifinality (Beven, 2006; see "GLUE method" below).

d) **Structural model uncertainty.** Introduced through simplifications, inaccuracies, and ambiguities of the description of the real processes during modeling.

Uncertainty Propagation

Measurements play an important role in the calibration and validation of hydrologic models. Scenarios play an equally significant role by predicting how the system might behave in the future. Uncertainty in hydrologic data typically ranges from 10% to 40%, but may be greater than 100%, depending on the type (McMillan, 2018). Each scenario is built from assumptions about how land use, water demand, or climate might change in the future, which can amplify uncertainty during the modeling process **(Figure 36)**.

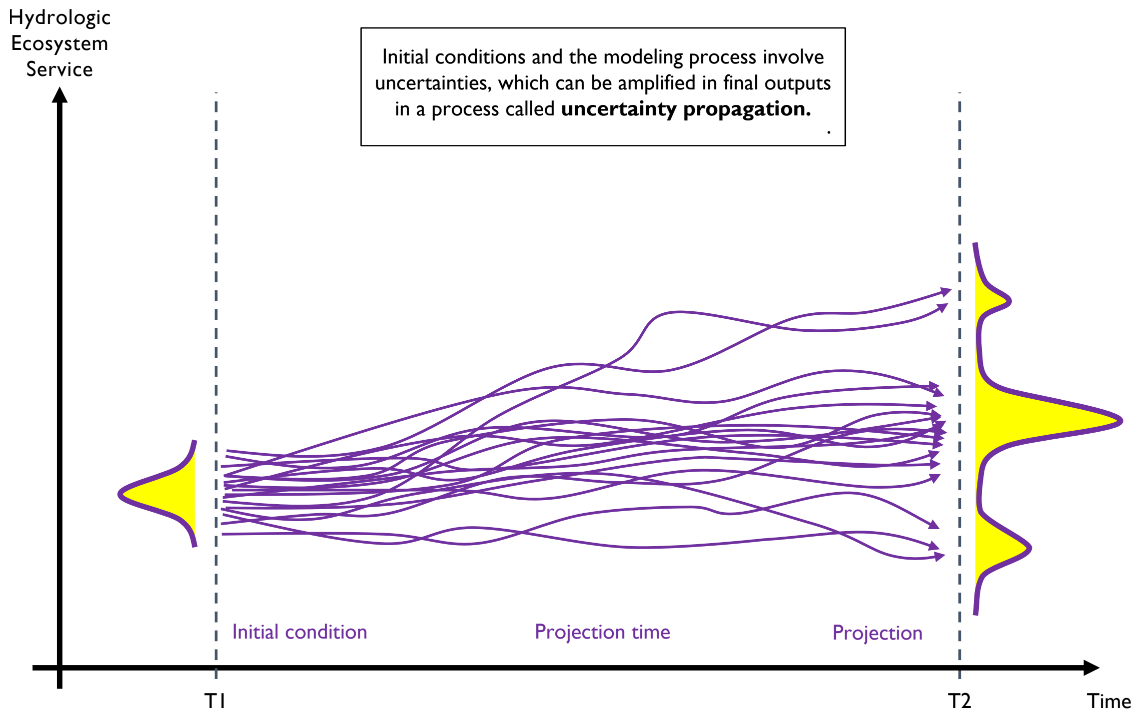


Figure 36. *Uncertainty propagation of a variable from initial conditions to results. Adapted from: ECMFW (2017).*

It is not common to evaluate the impact of all uncertainties combined, because the total uncertainty of the prediction would be so high that it would not be useful to decision-making. However, it can be useful to understand the impact of the quality of input information on the total uncertainty of results, even if the latter is not calculated. For example, the measurement uncertainty of a hydrometric station can be quantified in the results of a hydrologic model. Triangular V-notch weirs have high precision for the measurement of low flows, but measurement uncertainty increases during high flows (Herschy, 2009; ISO 18365, 2013). Extreme flow observations, which are used to calibrate flood models, can be used to evaluate the effect of the uncertainty of these types of stations. Sometimes, ***uncertainty propagation*** of input variables may be greater than the uncertainty associated with the estimation of parameter values (Kavetski et al., 2006).

"Uncertainty propagation of input variables may be greater than the uncertainty associated with the estimation of parameter values."

Uncertainty Analysis Methods

No matter which hydrologic models are used, decision-makers must be aware of the uncertainty inherent in the predictions of how NI interventions and ecosystem services might change in various scenarios. Bullock & Ding (2018) propose three approaches for uncertainty analysis:

- Obtain actual data on the production and flow of modeled ecosystem services. Even if there is little available information, test model outputs against these. If the models results are not accurate enough, alternative models can be used or other calibrated parameters to improve the fit of the model.

- Run multiple models at once (i.e., model ensembles). The mean of the resulting values from each can be used to make a decision or to estimate uncertainty.

- Treat modeling as an ongoing process. Running models and using them to make necessary decisions in real time. The model can be adjusted as additional information becomes available to refine estimates at each iteration.

This guide also recommends applying rigorous methods of uncertainty analysis to evaluate the outputs of hydrologic models to test hypotheses. Hypotheses are research questions that can be proved or disproved through the modeling process through experimentation and repetition.

- **Frequentist approaches.** Use the probability of an event as the limit of its relative frequency after multiple repetitions. Probabilities can be determined, in principle, by a repeatable objective process and are therefore ideally devoid of subjective opinions. In this approach, probabilities are determined only around random experiments or random samples (Neyman, 1937). The total set of all possible outcomes of a random experiment is called the *experiment sample space*. An event is defined as a particular subset of the sample space considered. For a given event, there are only two possibilities: i) the event occurs; or ii) it does not occur. The relative frequency of event occurrence is calculated as the number of occurrences of the event over the number of repetitions of the random experiment. As the number of repetitions of the experiment increases, the changes in the relative frequency decrease and tend to a certain limit (the probability of the event). Hypothesis testing, following frequentist inference, determines whether a hypothesis can be accepted or rejected with a certain level of statistical significance, which is calculated as the probability of observing something at least as extreme as what was objectively observed in an experiment (probability known as ***"p-value"***), under the assumption that the null hypothesis is true.

- **Bayesian approaches.** Assign probability to a hypothesis. These approaches are based on the interpretation of probability as a reasonable expectation that represents a state of knowledge or the quantification of a personal belief (Cox, 1946). It can be seen as an extension of propositional logic that allows reasoning with hypotheses. Bayesian probability is probationary (with evidence), in which an *a priori* probability is assigned to evaluate a hypothesis. This probability is updated to an *a posteriori* probability as new and relevant data (evidence) becomes available. Bayesian inference derives the *a posteriori* probability from two antecedents: i) the *a priori* probability; and ii) a function of ***"likelihood,"*** which is derived from a statistical model with the observed data using Bayes' Theorem on conditional probability. In Bayesian inference a probability is assigned to a hypothesis, whereas in frequentist inference, a hypothesis is tested without a probability being assigned.

"No matter which hydrologic models are used, decision makers must be aware that uncertainty is inherent in modeled predictions."

GLUE Method *(Generalized Likelihood Uncertainty Estimation)*

The ***GLUE Method*** (Beven & Andrew, 1992) is an informal Bayesian method used in hydrology to quantify the uncertainty in model predictions. The basic idea of this method is that, given our inability to represent accurately how nature works on a mathematical model, there will always be several models that could simulate equally well the observed natural processes, such as the generation of flow. Similarity in model acceptability is known as **equifinality** (Beven & Freer, 2001). Equifinality rejects the idea that there is a single optimal solution, given that the knowledge we have of the study area is imperfect, and estimates that several sets of models, parameters, and variables can be considered equally (or nearly equally) acceptable simulations of the system.

The GLUE Method deals with models whose results are expressed as probability distributions of possible outcomes, often in the form of Monte Carlo simulations or LHS **(Figure 37).** The problem can thus be seen as an evaluation and comparison of models to find out how good these representations of uncertainty are. There is an implicit understanding that the models used are approximations of what could be obtained by a Bayesian analysis of the problem, if a fully adequate model of real-world hydrological processes existed. The GLUE Method is equivalent to the approximate Bayesian calculation for some choices of statistical evaluation functions and performance thresholds. However, it tends to be criticized by professionals in more formal statistics.

Performance analysis can be replaced with an uncertainty analysis (**Figure 37**) by following the scheme of automatic modeling shown before (**Figure 25**). The uncertainty analysis using the GLUE Method can be performed with simulation results (stage 3). That is, we replace the performance analysis with an uncertainty analysis (**Figure 37**).

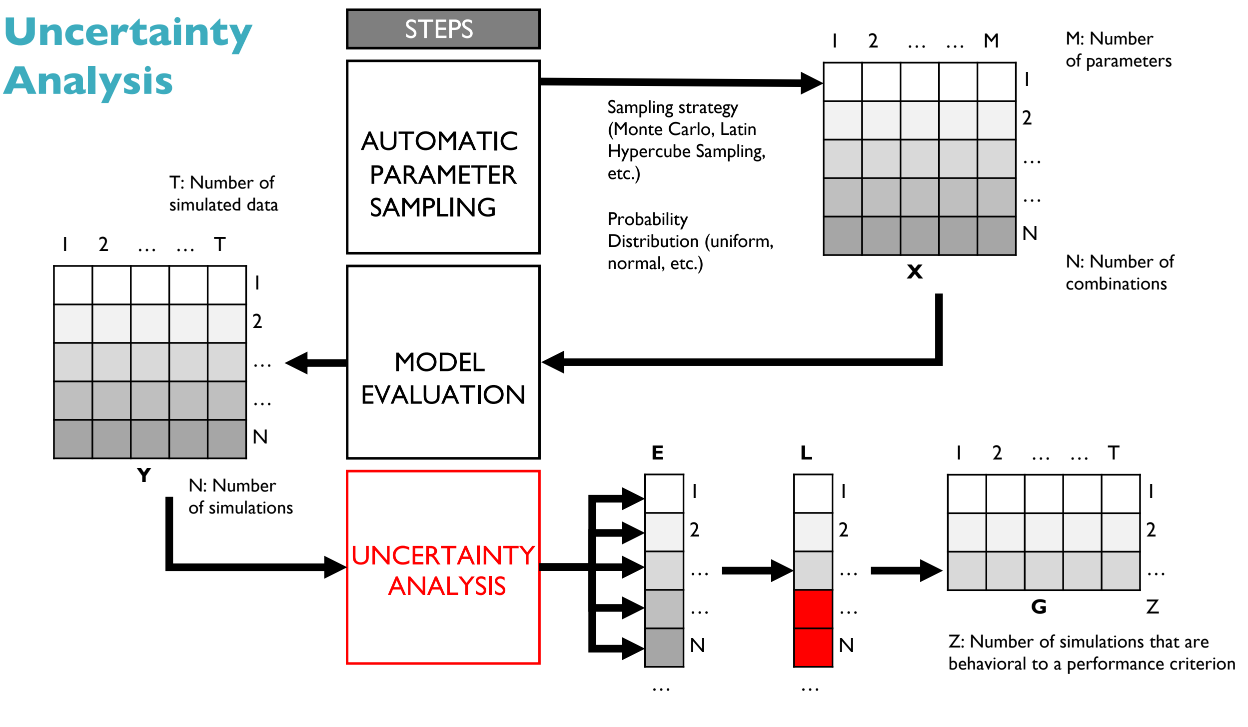


Figure 37. Uncertainty analysis of a simulation model. Source: Prepared by the authors.

Matrix Y of simulated results is analyzed to evaluate how the performance of each simulation line varies compared to the set of observed data, which is used to obtain performance indicators (matrix E). This list of indicator results is ordered from "best performance" (closest simulation to observed data) to "worst performance." In the GLUE Method **(Figure 37)**, the performance indicators are used as a proxy to approximate a likelihood function **L**. In other words, a higher probability is assigned to a simulation that has a better performance than to a simulation that has a bad performance. It also produces a performance threshold equivalent to the "behavioral" or "non-behavioral" criteria of the RSA sensitivity analysis **(Figure 31)**. Simulations that do not meet the threshold are discarded (likelihood is equal to zero (L = 0)), and only the simulations that are better than said threshold are considered. Then, using the likelihood function for those simulations that meet the threshold, a subset of G simulations is obtained, from which a probability function of the outcomes is derived. Processing this probability function for a confidence level (e.g., 90%, 95%, 99%) produces the minimum and maximum ranges enclosing observations **(Figure 38)**. These ranges represent the uncertainty associated with the subset of simulations generated by the model at the established level of significance, which includes multiple sources of error, such as the equifinality of the model parameters and variables.

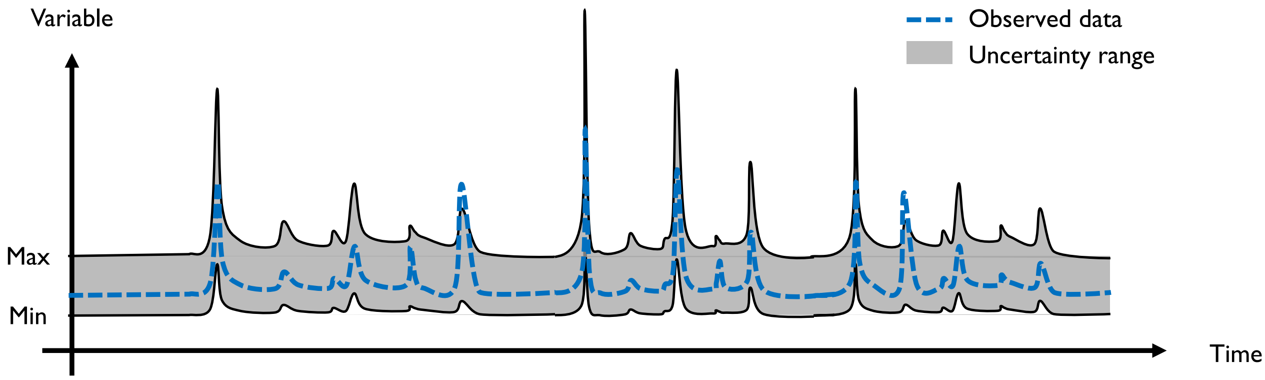


Figure 38. *Conceptual representation of the uncertainty bands determined with the GLUE Method. Source: Prepared by the authors.*

"Equifinality rejects the idea that there is a single optimal solution and estimates that several sets of models. Parameters and variables can be considered equally (or nearly equally) acceptable simulations of the system."

Other Methods for Estimating Uncertainty

GLUE is the most popular method in hydrology, frequently available with other hydrologic models. Because of this, GLUE is accepted for most hydrologic model applications. However, there are other options, all with pros and cons; and thus uncertainty analysis in hydrologic predictions is under ongoing development. There are different methods based on a specific variation of Monte Carlo simulations, known as Markov Chain Monte Carlo, or MCMC. These methods are used to sample distribution parameters *a posteriori*. Among them is the Metropolis-Hastings (MH) algorithm (Metropolis et al., 1953; Hastings, 1970) and the DREAM algorithm, which is a more efficient version applied to hydrologic models (Vrugt et al., 2008a; Vrugt et al., 2008b). Some other methods based on Monte Carlo simulations include the Kalman filter and its extensions (Kitanidis & Bras, 1980), the DYNIA approach (Wagener et al., 2003), the BaRe approach (Thiemann et al., 2001) and the SCEM-UA algorithm (Vrugt et al., 2003). Some approaches for specific case studies include, flood risk analysis (Beven & Hall, 2014; Hall & Solomatine, 2008), hydrologic forecasts (Montanari & Grossi, 2008) and ungauged basins (Blöschl et al., 2013).

One final point to consider is that the GLUE Method and most Monte Carlo simulation-based methods only address the uncertainty associated with parameter estimation. New methods are being developed to treat various sources of uncertainty collectively, including the Bayesian multi-model approach (Ajami et al., 2007) and data assimilation (Lui & Gupta, 2007).


Photo: Carlos Palacios

STEP 10. INTERPRETATION OF RESULTS

As outlined above, the use of models for the evaluation of NI is subject to various sources of error. Simulation results need to be critically and thoroughly evaluated to reduce the likelihood that they will have an unexpected or detrimental impact on decision-making. For example, if a decision depends on whether a variable is above or below a given critical value, and the modeling result is dangerously close to that value, it is necessary to quantify and analyze the errors associated with the result obtained. If the estimated total error is not enough to cross this critical variable value, the decision can be made with confidence. But if the estimated error is large enough, then the decision will be associated with only a certain level of confidence.

For example, if a model produces 100 simulations of flow, of which 90 generate results between 3 and 7 m^3/s, then we can say that the estimated value of flow rate is 5 m^3/s, with a 90% confidence interval between 3 and 7 m^3/s. Still, it is important to recognize that, in this case, results indicate that there is still a 10% chance that the actual flow rate is outside this range. Usually this means that there is 5% probability that the flow is less than 3 m^3/s and another 5% probability that it is greater than 7 m^3/s, although this depends on the type of probability distribution that best fits the data analyzed. A normal probability distribution is most commonly used, which has a bell shape and is symmetrical around the mean. Asymmetric, finite, or discrete distributions can also be appropriate, depending on model and decision context.

"Hydrologic modeling must be seen as a useful but uncertain tool."

Uncertainty must be considered in all possible evaluation scenario results, as shown **Figure 39**. It is possible that, at certain levels of uncertainty, the results of the BAU scenario overlap with those of the NI or SEM scenario. This is not necessarily because the impacts of the interventions are not effective, but because of errors introduced by data availability, model structure, or natural variability. It is therefore necessary to include uncertainty analysis in the model calculations and in analyses of water benefits of the NI with a defined confidence interval. It is widely accepted that the overlaps between the confidence intervals must be less than 5% to be declared ***"statistically significant"*** changes (similar to a p-value <0.05). In other words, if the "tails" of the probability distributions between two scenarios (e.g., BAU vs. IN) overlap by more than 5%, the decision poses a higher risk that the observed effects are artifacts of modeling, not caused by the NI interventions.

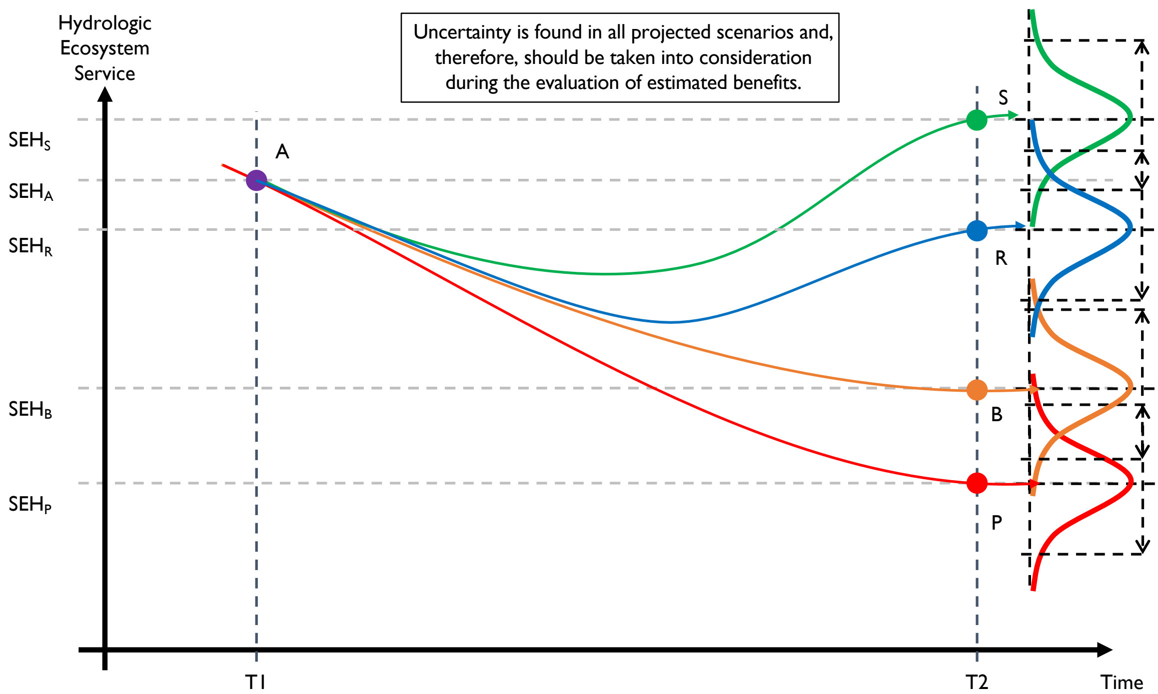


Figure 39. *Uncertainty in the projections of all possible future scenarios. Source: Ochoa–Tocachi et al. (2018b).*

Clarifications on the Use of the p-value

The ***p-value*** is commonly misunderstood and misused. While it can be a useful tool, it is not a replacement for critically thinking about your data, model design, or decision context. The American Statistical Association states that p-values can indicate how incompatible the data are with respect to a specific model (Wassertein & Lazar, 2016). In Neyman-Pearson hypothesis testing, the data obtained by comparing the p-value with a significance level will yield one of two results: i) the null hypothesis is rejected (which does not prove that the null hypothesis is false); or ii) the null hypothesis cannot be rejected at that significance level (which does not prove that the null hypothesis is true). From a Fisher statistical inference approach, a low p-value means that the null hypothesis is true and that a highly unlikely event has occurred, or that the null hypothesis is false. The following table clarifies some common misunderstandings about p-values (Schervish, 1996; Sterne & Smith, 2001; Wassertein & Lazar, 2016):

- **The p-value is not the probability that the null hypothesis is true or the probability that the alternative hypothesis is false.** A p-value can indicate the degree of compatibility between a data set and a particular hypothetical explanation, such as the null hypothesis. Specifically, the p-value can be taken as the *a priori* probability of obtaining an effect that is at least as extreme as the observed effect, in case the null hypothesis is true. This should not be confused with the *a posteriori* probability that the null hypothesis is true given the observed effect. Professionals in frequentist statistics do not assign probabilities to hypotheses.

- **The p-value is not the probability that the observed effects were produced by chance alone.** The p-value is calculated under the assumption that a certain model, usually the null hypothesis, is true. This means that the p-value is a statement about the relationship between the data and the hypothesis.

- **The 0.05 level of significance is a convention.** The ***"alpha" (α)*** level of significance of 0.05 (95% confidence) is often used as the limit between a statistically significant and a non-significant p-value. However, this does not imply that there is a general scientific reason to consider that results that are outside this threshold are qualitatively different.

- **The level of significance represents a balance of decisions.** The level of significance chosen represents the type I error of a hypothesis test or the probability of rejecting the null hypothesis when it is true. In the ***"beta"*** level (***β***), 1-β is the power of a statistical test. ***β*** represents the type II error of a test or the probability of accepting a null hypothesis when the null hypothesis is false. Analysis of type I and type II errors in a decision-making scenario can lead to different associated costs. For example, investing in NI when a desired threshold value will not be reached, versus not investing in NI when a desired threshold would be reached eventually without NI. This last example will have different associated costs for extreme flood events or water scarcity scenarios. The cost and probability associated with each error type could be balanced, depending on the complexity of the analysis and needs of the decision-makers.

- **The p-value does not indicate the size or importance of the observed effect.** An effect that is not considerable nor important can be associated with a small p-value. In fact, the larger the sample size, the smaller the minimum effect necessary to produce a p-value that is statistically significant. Visualizing the size of effects is a critical component of a data analysis method called *estimation statistics*.

"The p-value is commonly misunderstood and misused."

Communicating Modeling Results

Time series model results are not usually useful to decision-making – they require context and interpretation, especially for non-experts. Results can be interpreted and clarified using post-processing, such as data visualization or statistical treatment. When choosing post-processing, the team should focus on policy questions and the stated objectives of the modeling process. While this is additional work, it is a necessary step in effectively communicating results in a way that will be accessible and useful to decision-makers. Collaboration and open discussion can allow both technical and decision-making teams to better understand results.

"Collaboration and open discussion can allow both technical and decision-making teams to better understand results."

The treatment and communication of results must have the following objectives: i) to answer policy questions; and ii) to include ranges for all estimates so decision-makers have more context and more space within which to decide. The simplest way to observe hydrologic benefits between two time series (e.g., BAU vs. NI) is by calculating the differences in the results of the variable in question between the scenarios. For example, time series data can be grouped by month or season to obtain values that differentiate water availability or a priority ecosystem service during each year. Grouped data sets will likely contain ***("outliers")*** due to extreme events in the simulation. Depending on the normality of the distribution, the average availability of water throughout the year could be represented using means or medians with standard or quartile deviations. These results can be presented visually monthly or by season as time series, box and whisker plots, as text, or charts.

When it comes to extreme events, it is possible that additional ***post.*** will be required. For example, answering the question of how much proposed NI interventions are capable of reducing flood risk may be impossible with raw time series data. Depending on the number of years simulated, it could be possible to obtain some extreme values. However these values alone may be insufficient to report on the hydrologic benefits of NI in events with high return periods (e.g., 100 years). These time series results would have to be extrapolated by statistically fitting them to an extreme distribution function, such as *generalized Pareto* (GP) or *generalized extreme value* (GEV). Prediction intervals can be constructed through the methods presented in Step 9 and, in all cases, must be presented with an explanation of how benefit intervals were calculated and what sources of uncertainty are included. Prediction ranges can also be derived using post-processing statistical methods; for example, the confidence intervals for analysis of outliers can be obtained by adjusting the modeling results to a distribution of outliers by *bootstrapping* **(Figure 40)**.

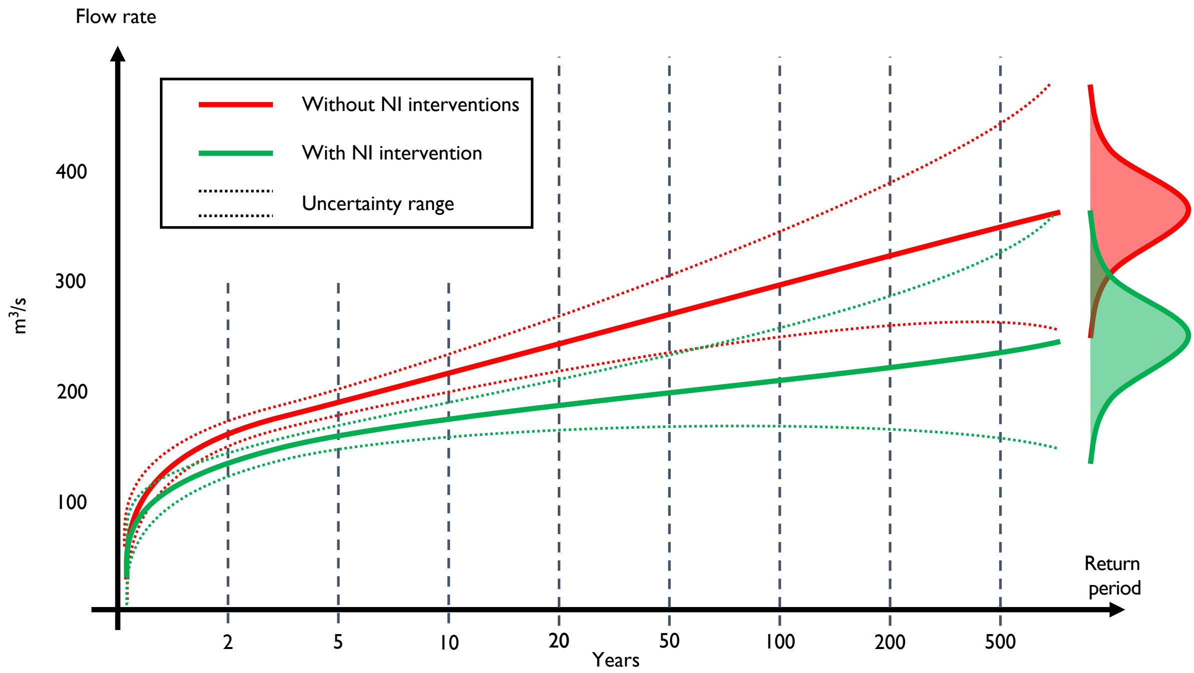


Figure 40. *Presentation of confidence intervals derived from data post-processing. Source: Prepared by the authors.*

Figure 40 illustrates the results of a modeling exercise evaluating the likelihood of flooding without NI interventions (red) and with NI interventions (green). When comparing the average results (solid lines), a clear difference is observed between both alternatives, which suggests that the situation without interventions is more susceptible to flooding than the scenario with NI interventions. For example, in a scenario without intervention, a flow event of up to 240 m^3 /s is expected to occur every 20 years. In comparison, the NI scenario has a flow event of just 180 m^3 /s expected every 20 years. If, for example, the maximum level of the flood defense or the early warning system is 200 m^3 /s, this threshold will be exceeded more frequently in the absence of NI interventions. It is also possible to read and interpret the reverse: a flow of 200 m^3 /s is expected to occur every 10 years in a non-intervention situation, but the same flow is expected to occur barely every 50 years in a scenario with NI interventions. This difference in framing can have very significant implications for disaster risk management.

When an uncertainty analysis (dotted lines) is included, the exercise may appear more complex but this is actually important additional information for decision-making **Figure 40**. During the first couple of years, the uncertainty ranges overlap between the two scenarios, which means that results should not be expected early in the project. This is standard for NI interventions because their effects take time to produce measurable effects on selected hydrologic indicators. In this situation, decision-makers can feel confident in medium and long-term interventions, and should not expect to see results immediately after implementation. In **Figure 40**, from the second year until year 50, uncertainty ranges between scenarios do not intersect, making it possible to strengthen the confidence that measurable and distinct results are indeed expected between situations without intervention versus the scenario with NI intervention. From year 50 onwards, uncertainty ranges cross again between scenarios, which suggests that there is a probability that the effects of the interventions cannot be clearly distinguished. The more we venture into the future, the ranges of uncertainty become wider the farther scenarios extend into the future and the overlaps are greater, reflecting the fact that we cannot perfectly predict how study areas will be affected.

This type of analysis can provide some additional advantages to single analyses of a design storm event with specific return period or a single time series of observed and simulated flow. The results of a statistical adjustment to the distribution of extremes will provide information and confidence intervals of several return periods at the same time, which makes it not only possible to quantify multiple benefits of flood control, but also the probability of hydraulic infrastructure failure, which occurs in shorter return periods (usually 5 to 15 years).

The examples presented here are not comprehensive. The purpose of this section is to clarify that the results of a hydrologic model usually need further processing to make sense in the decision-making context. The specifics of processing, displaying, and presenting results will depend on the policy question, the modeling objectives, and the target audience. Finally, risk management and the existence of backup and contingency plans are elements that must be considered in final decision-making.

> ***"The specifics of processing, displaying, and presenting results will depend on the policy question, the modeling objectives, and the target audience."***

Illustrative Application

Returning to the Illustrative Application of the Tambo-Ilo-Moquegua Basin from Volume I **(Figure 10)**, biophysical information was collected for the selected hydrologic model. **Figure 41** shows a map with the data available in the basin, including digital elevation model, weather stations, flow gauging stations, soil map, and biophysical characteristics. **Figure 42** shows land use and land cover maps for years 2000, 2010, and 2019.

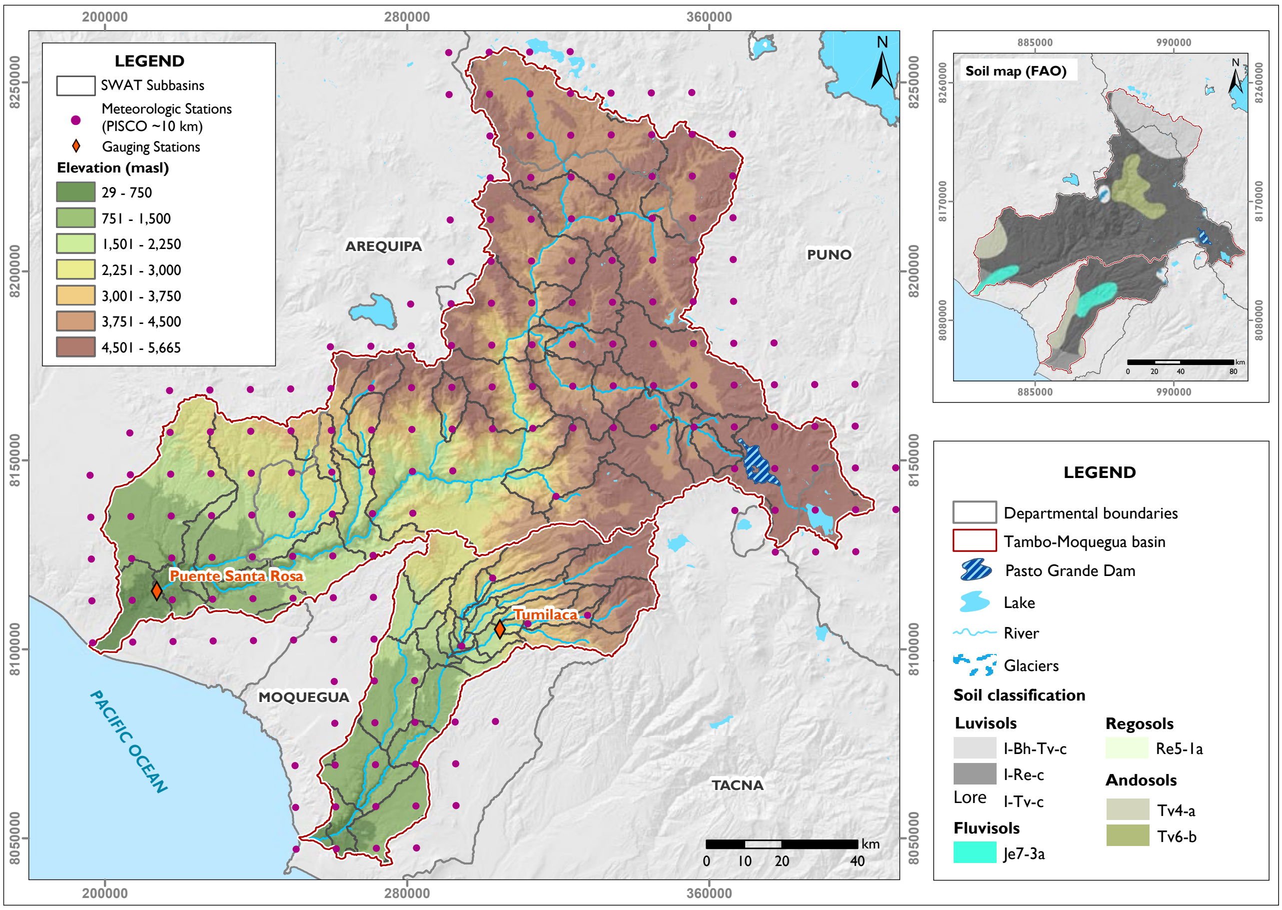


Figure 41. *Data map for the Tambo–Ilo–Moquegua Basin: elevation, meteorological stations, flow gauging stations, soil types, and biophysical characteristics. Source: Prepared by the authors.*

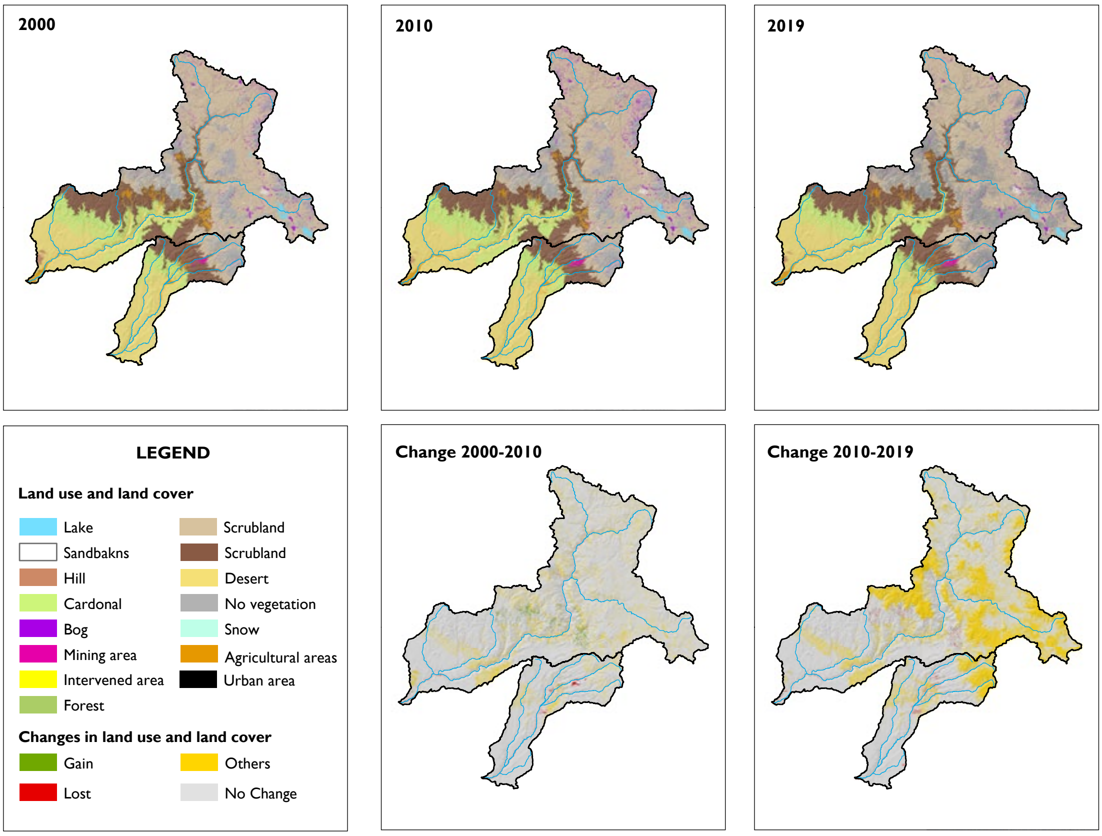


Figure 42. *Land use and land cover maps for the Tambo–Ilo–Moquegua Basin: 2000 (a), 2010 (b), and 2019 (c). Source: Prepared by the authors.*

Using the HIRO tool (Román et al., 2020), 354,933 hectares (ha) have been prioritized for the provision of hydrologic ecosystem services. Of this, 67,745 hectares are priorities for the water regulation service; 175,426 hectares are priorities for the erosion control service; and 111,762 hectares are priorities for both ecosystem services **(Figure 43)**.

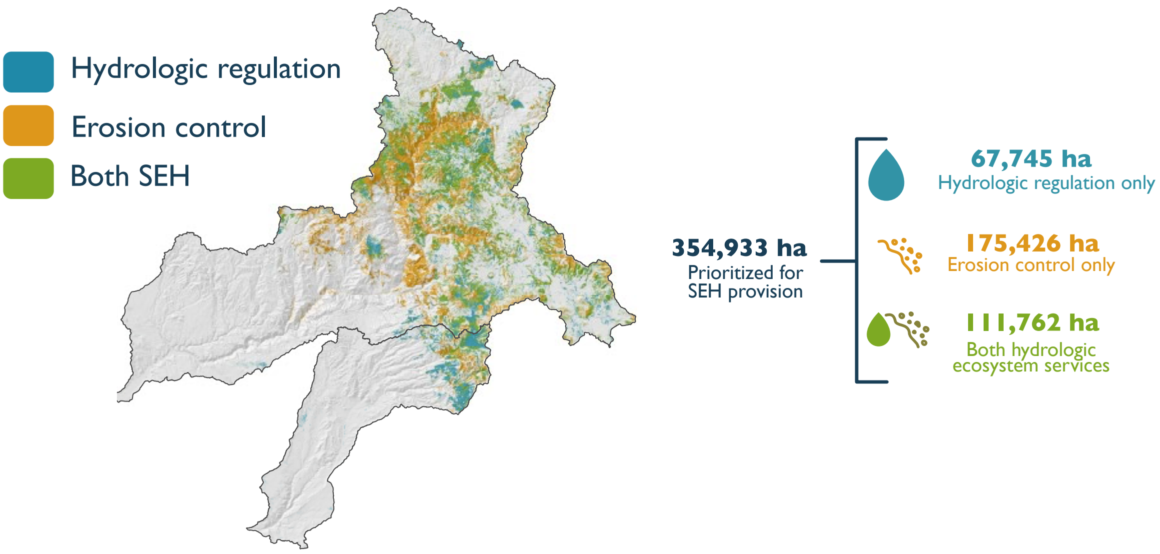


Figure 43. *Areas of importance for hydrologic ecosystem service provision (SEH), as identified using the HIRO tool. Source: Prepared by the authors.*

With this information, four proposed modeling scenarios were built: BASE, BAU, PES, and SEM

- **BASE scenario**
  - o The current state of the system to be modeled.

- **BAU scenario**
  - Projection of land use and land cover change for year 2060, considering anthropic and natural drivers. The projection is based on the rate of cover loss observed in the period 2000–2019.
  - Land use and land cover change of anthropic land, following the trend observed between 2000–2010/2010–2019:
    - Loss rate: 6,300 ha/decade
    - Vegetation cover lost as of 2060: 25,200 ha
  - Land use and land cover change assuming that unfavorable covers broaden their range of distribution due to environmental stressors:
    - Expanding covers: area without high-Andean vegetation and sandbanks
    - Loss rate: 3,300 ha/decade
    - Vegetation cover lost as of 2060: 13,400 ha
  - Losses of vegetation cover and land-use changes are recorded in 138,307 ha.

- **PES scenario**
  - The areas prioritized by HIRO will be totally degraded by year 2060.
  - 354,933 ha prioritized by HIRO are considered "degraded" covers, depending on land use and ecosystems near degraded areas.

- **SEM scenario**
  - The areas prioritized by HIRO will be fully restored and conserved by the year 2060.
  - 354,933 ha prioritized by HIRO are considered "conserved" cover, depending on the ecosystems and ideal uses close to the designated areas.

Scenarios BASE, BAU, PES, and SEM are shown in **Figure 44.**

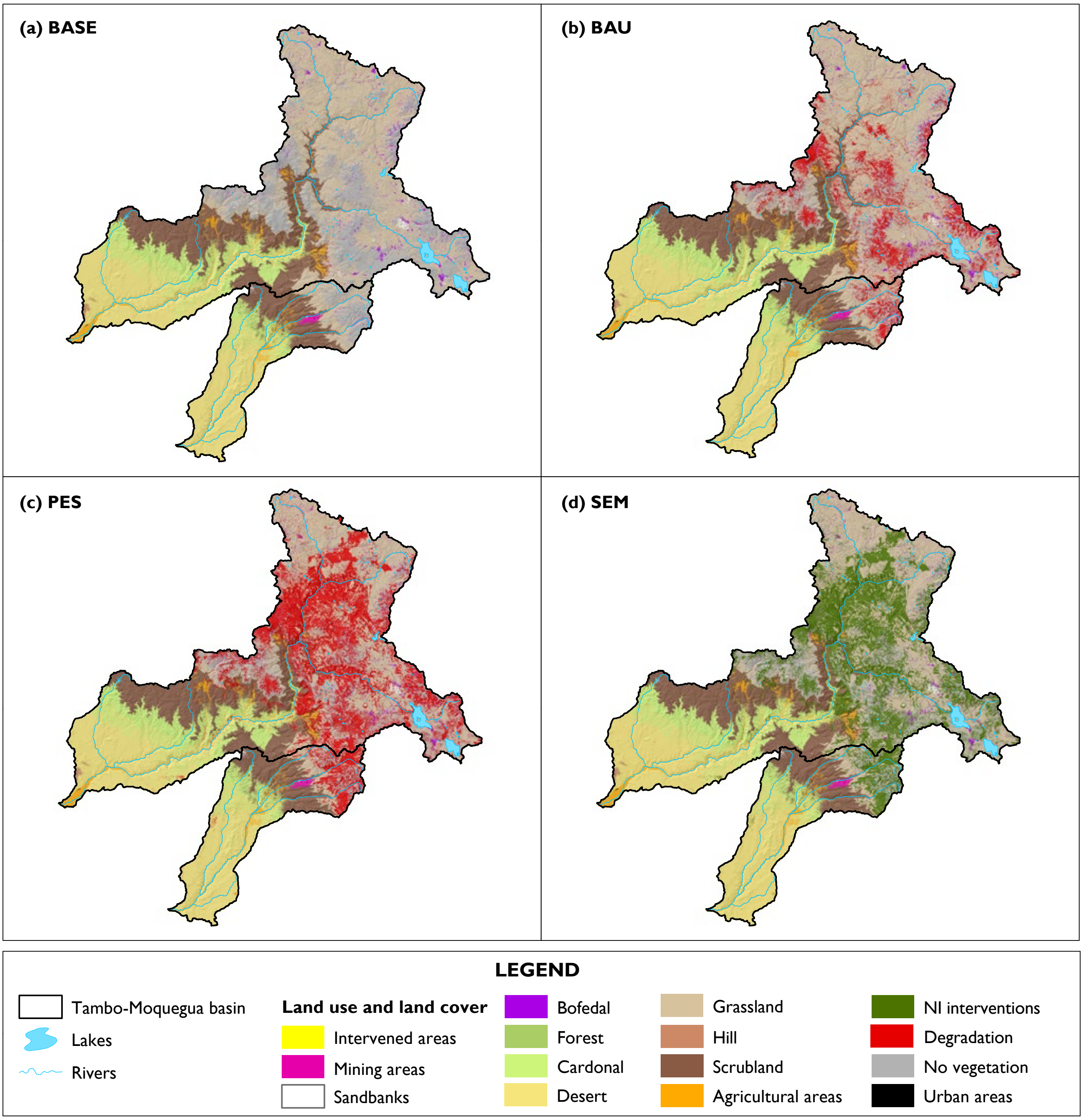
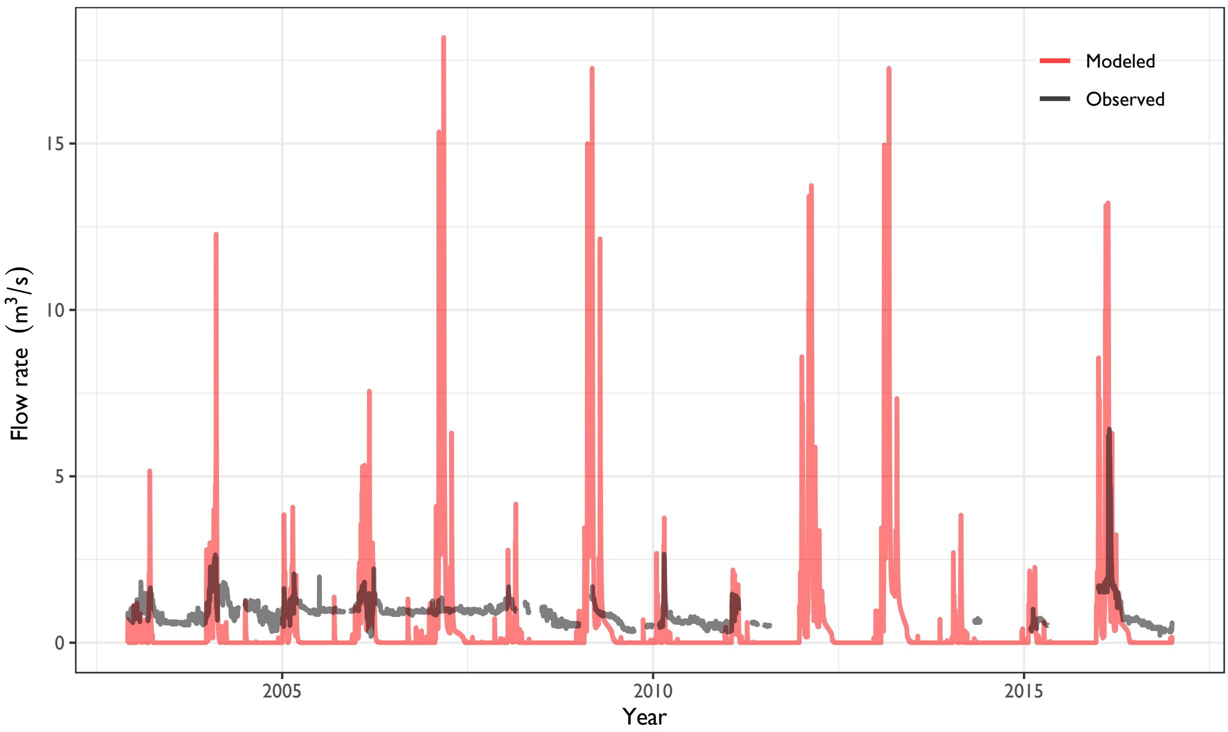


Figure 44. *BASE (a), BAU (b), PES (c) and SEM (d) scenarios for the Tambo-Ilo-Moquegua Basin. Source: Prepared by the authors.*

Figure 45 shows the results of the calibration process and validation for Santa Rosa and Tumilaca sub-basins. As can be seen, the two flow stations generate data with different magnitudes (Santa Rosa is one or two orders of magnitude higher than Tumilaca). The data in Tumilaca seem farther from the observed data, and the data in Santa Rosa seem more approximate, due to the wide range in the magnitude of the flows.

Tumilaca



Santa Rosa

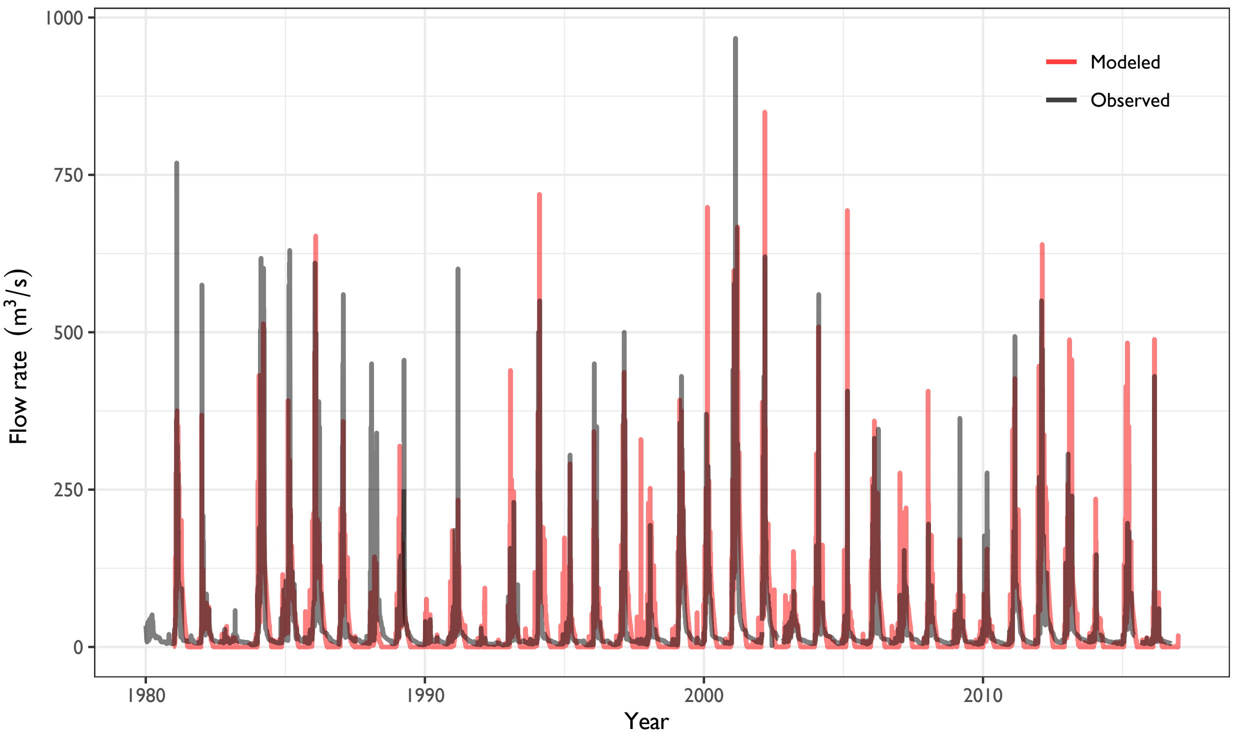


Figure 45. *Flow modeling calibration and validation in the Santa Rosa and Tumilaca sub-basins, within the Tambo–Ilo–Moquegua Basin. Source: Prepared by the authors.*

Tumilaca Outputs

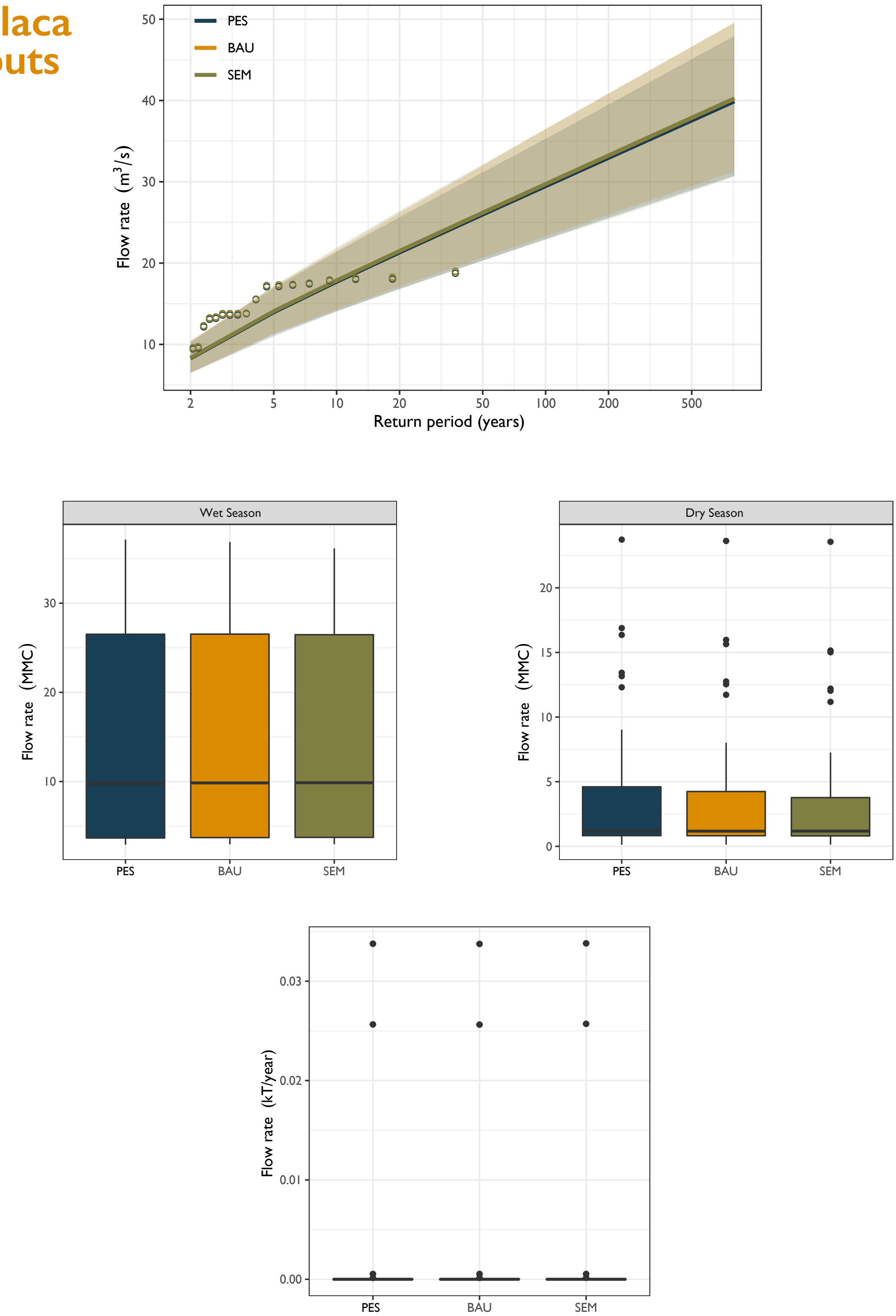
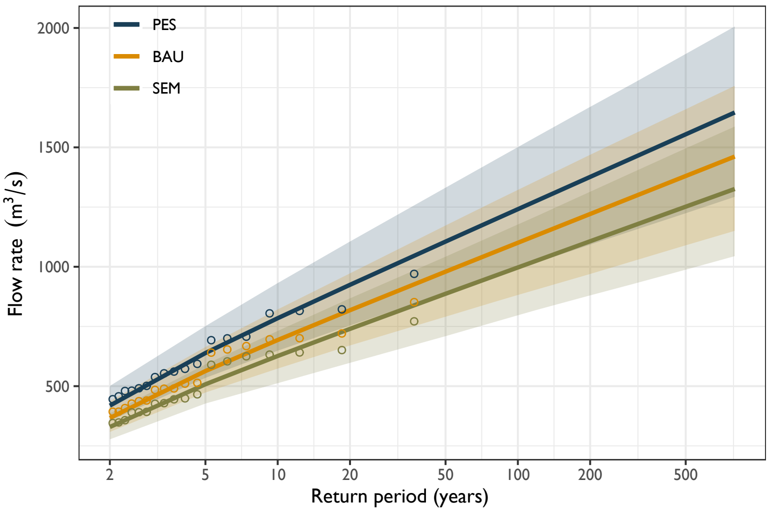


Figure 46. *Tumilaca sub-basin flow results (within Tambo–Ilo–Moquegua Basin). Source: Prepared by the authors.*

Santa Rosa Outputs



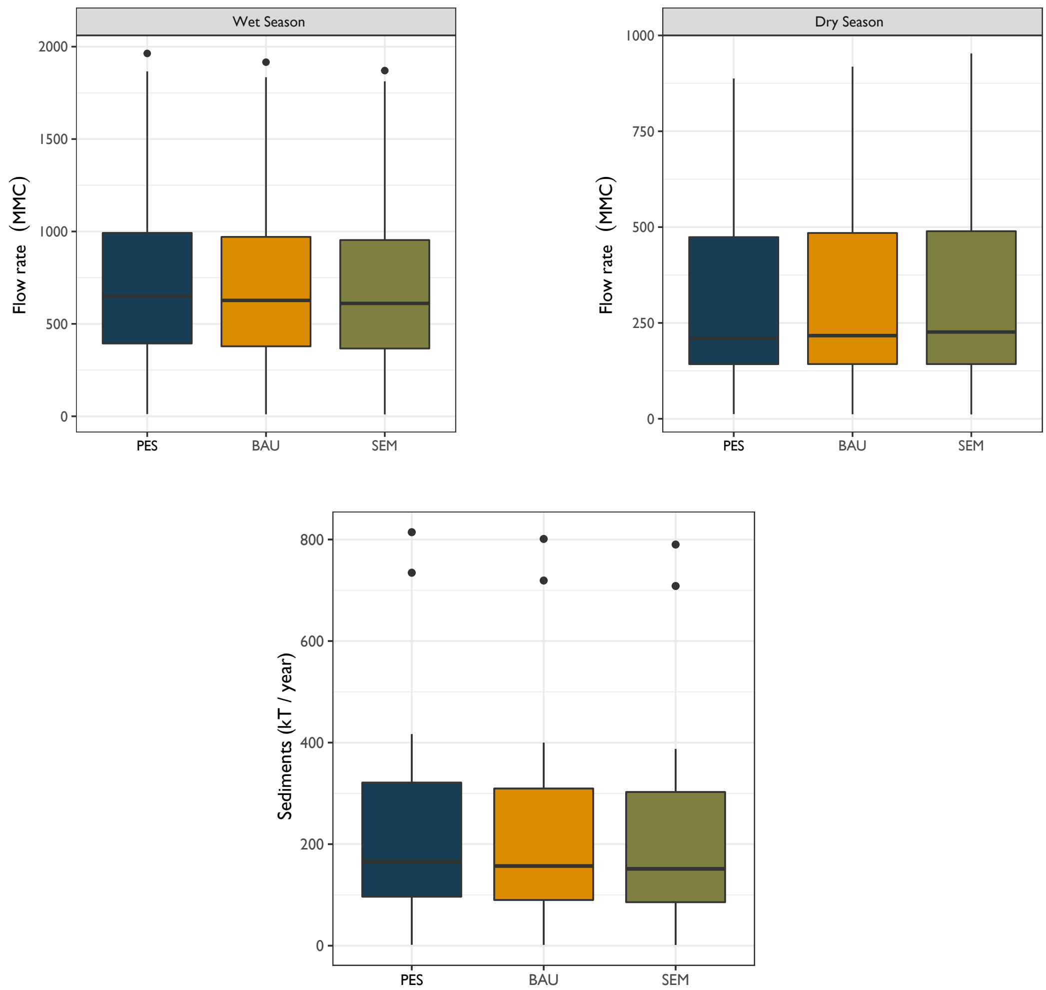


Figure 47. *Santa Rosa subbasin flow results (within Tambo–Ilo–Moquegua Basin). Source: Prepared by the authors.*

As shown in **Figure 46** (Tumilaca) and **Figure 47** (Santa Rosa), the results can be presented differently. For Tumilaca, no significant changes in flow are observed in the three modeled scenarios (the sediment load is also negligible in all three cases). There are changes in Santa Rosa, but not considerable ones. However, greater differences are observed in the return periods that can occur under the different scenarios. The underlying uncertainty is quite high.

Conclusions, Recommendations, and Key Ideas

Hydrologic models are a resource for decision-makers to address a variety of questions related to the impacts of NI interventions and the quantification of their hydrologic benefits. This *Guide to Hydrologic Modeling of Natural Infrastructure* provides guidelines in detailed steps, with a goal to better support hydrologic model selection and implementation for decision-making.

Modeling tools are particularly useful when observed hydrologic data is sparse or uncertain in the region of interest. However, selecting an appropriate model to support decision-making is not an easy task due to a variety of elements: complexity of the available models, local technical capacities, availability of resources, decision-making context, etc. For this reason, decision makers must clearly define the scope of their analysis, evaluate the characteristics of each model for the context in question, and carry out a self-evaluation of their resources and capacities to understand specific needs. It is common for initiatives to have objectives that are either too ambiguous or vague to be useful to decision-making. This often produces results that do not clearly inform policy questions. Clear objectives more soundly direct model selection and operation.

Hydrologic modeling and data generation are closely linked. Monitoring and measurement data can improve modeling results, but the use of poor-quality input data will result in poor model outputs. Use of models can contribute to improving the understanding of ecosystem processes and optimize or inform data collection. For example, sensitivity analysis methods can reveal the variables on which efforts should be focused when collecting data in the field.

In addition, the simple use of popular models does not necessarily lead to better or more useful results for specific applications in NI projects. To obtain the most appropriate

results for decision-making, model selection must be determined by the availability of data and the experience of the team. This guide allows NI project planners to adapt their needs in real time as they design and assess interventions for salience, credibility, and legitimacy. Some key questions decision-makers should feel empowered to ask are: Is this model calibrated? Is uncertainty assessed? Who has been involved in the process?

Decision makers should not expect hydrologic models to provide solutions for immediate application and to adequately consider the associated sources of uncertainty. It is impossible to eliminate uncertainty completely, since it comes from conceptual model design, observed data, the model itself, scenarios projecting future outcomes, model calibration, etc. This does not mean modeling is not important or useful, rather that uncertainty analysis and communicating model limitations is necessary to informed decision-making. The use of a set of different models could increase the robustness of simulation results compared to the use of a single one in cases where a combination would best represent how a study area functions and produce a greater range of possibilities.

Finally, it is necessary to consider a risk management approach and have contingency plans incorporated into decision-making. Models can be useful decision-making support tools, but it is essential to recognize that they are also uncertain. Using models to complement other decision-making processes can help increase the robustness of the conclusions that guide the decisions at hand.

What's next?

Modeling is not a replacement for collecting more data in situ, particularly in remote areas, low-budget regions, or in the face of local technical capacity constraints. In the tropical Andes, for example, much remains to be understood about processes such as evapotranspiration and underground and subsurface water flows, among others. Using low-cost sensors with telemetry, possibly in connection with participatory monitoring schemes, or citizen/community science, could support this process. The NIWS project also promotes the development of ecohydrologic monitoring guides for the evaluation of NI. These data, combined with results from robust hydrologic models, can multiply value for decision-making.

"This guide emerges from wide experience in generating information to support decision-making for natural infrastructure and water resources management."

It is essential to develop models that are as representative of local realities as possible. Although the resources, data, and technical capacity necessary to build models the preferred way are not always available, building original models can be more efficient and useful than using those that already exist in certain contexts. Many popular models used today have been developed in other countries to represent hydrologic conditions that are not present at the application site. A common case is that of the models simulating surface runoff due to excess infiltration (e.g., using curve number) and which are applied in geographical areas with high-infiltration capacity soils and low-intensity rainfall. In cases like this, the models that simulate runoff due to excess saturation would be more representative.

Finally, modeling applications are varied. Analyzing the hydro-socioeconomic benefits of NI projects is a promising front for natural resource management, sustainable development, and helping abate climate change. In this guide, we suggest a flow of analysis where hydrologic modeling plays a leading role in various stages of the process. For many applications it is necessary to go beyond the quantification of water benefits and towards the quantification of social and economic benefits; the consideration of co-benefits, externalities, commitments, and costs; and the corresponding balance of all these elements to support investments that are sustainable, equitable, and profitable. Further developments of this guide will address details related to analysis of economic benefits and cost-effectiveness and the preparation of original models for specific applications.

"Given the scarcity of data observed in various regions of the world, we hope that this document is a valuable contribution to the design and evaluation of natural infrastructure investment projects for water security."

Photo: Natural Infrastructure for Water Security

References

Ajami, N. K., Duan, Q., and Sorooshian S. (2007). An integrated hydrologic Bayesian multimodel combination framework: Confronting input, parameter, and model structural uncertainty in hydrologic prediction. *Water Resources Research* 43: W01403.

Arnold, J. G., Kiniry, J. R., Srinivasan, R., Williams, J. R., Haney, E. B., Neitsch, S. L. (2012). *Soil and Water Assessment Tool (SWAT) Version 2012.* Texas Water Resources Institute TR–439.

Arrow, K., Cropper, M., Gollier, C., et al. (2013). Determining benefits and costs for future generations. *Science* 341, 349–350.

Bergström, S. (1992). *The HBV model – its structure and applications.* SMHI Reports RH, No. 4, Norrköping.

Bergström, S. (1995). The HBV model. In: Singh, V. P. (Ed.). *Computer models of watershed hydrology.* Water Resources Publications, Highlands Ranch, CO., 443–476.

Beven, K. J. (2006). A manifesto for the equifinality thesis. *Journal of Hydrology* 320: 18–36.

Beven, K. J. (2011). Rainfall-runoff modelling: the primer / Keith Beven. – 2nd ed. ISBN 978-0-470-71459-1.

Beven, K. J. (2019). Towards a methodology for testing models as hypotheses in the inexact sciences. *Proceedings of the Royal Society* A. 475, 2018086220180862.

Beven K. J., Binley, A. (1992). The future of distributed models: model calibration and uncertainty prediction. *Hydrological Processes*. 6 (3): 279–298. doi:10.1002/hyp.3360060305. ISSN 0885–6087.

Beven, K. J., Binley, A. (2014). GLUE: 20 years on. *Hydrological Processes*. 28 (24): 5897–5918. doi:10.1002/hyp.10082. ISSN 0885–6087.

Beven, K. J., Freer, J. (2001). Equifinality, data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems using the GLUE methodology. *Journal of Hydrology*. 249 (1–4): 11–29. doi:10.1016/S0022–1694(01)00421–8. ISSN 0022–1694.

Beven, K. J., Hall, J. (2014). *Applied uncertainty analysis for flood risk management.* World Scientific.

Beven, K. J., Kirkby, M. J. (1979). A physically based, variable contributing area model of basin hydrology. *Hydrological Sciences – Bulletin – des Sciences Hydrologiques,* 24, 1.

Beven, K. J., Young, P. (2013). A guide to good practice in modeling semantics for authors and referees. Water Resources Research 49: 5092–5098.

Blöschl, G., Sivapalan, M., Wagener, T., Savenije, H., Viglione, A. (2013). *Runoff prediction in ungauged basins: synthesis across processes, places and scales,* Cambridge University Press.

Boeker, E., van Grondelle, R. (1995). *Environmental physics.* John Wiley & Sons, Chichester, England.

Bonnesoeur, V., Locatelli, B., Guariguata, M. R., Ochoa–Tocachi, B. F., Vanacker, V., Mao, Z., Stokes, A., Mathez–Stiefel, S–L. (2019). Impacts of forests and forestation on hydrological services in the Andes: a systematic review. *Forest Ecology and Management.* 433, 569–584.

Brauman, K. A. (2015). Hydrologic ecosystem services: linking ecohydrologic processes to human well - being in water research and watershed management. WIREs Water 2: 345–358.

Brauman, K.A., Bremer, L.L., Hamel, P., Ochoa-Tocachi, B.F., Roman-Dañobeytia, F., Bonnesoeur, V., Arapa, E. and Gammie, G. (2021b), Producing valuable information from hydrologic models of nature-based solutions for water. Integr Environ Assess Manag.

Brauman, K. A., Daily, G. C., Duarte, T. K., Mooney, H. A. (2007). The nature and value of ecosystem services: an overview highlighting hydrologic services. *Annual Review of Environment and Resources* 32: 67–98.

Brauman, K. A., Gammie, G., Bonnesoeur, V., Ochoa–Tocachi, B. F., Arapa, E., Román, F., Cassin, J. (2021). *Informing natural infrastructure investment decisions with hydrologic modelling: lessons learned from a case study in Chancay–Lambayeque watershed, Peru.* Proyecto Infraestructura Natural para la Seguridad Hídrica, Forest Trends, Lima, Perú.

Bremer, L. L., Hamel, P., Ponette–González, A. G., Pompeu, P. V., Saad, S. I., Brauman, K. A. (2020). Who are we measuring and modeling for? Supporting multilevel decision–making in watershed management. *Water Resources Research* 56(1): e2019WR026011.

Bullock, J., Ding, H. (2018). *A guide to selecting ecosystem service models for decision–making: lessons from Sub–Saharan Africa.* World Resources Institute. ISBN: 978–1–56973–933–1.

Butts, M. B., Payne, J. T., Kristensen, M., Madsen, H. (2004). An evaluation of the impact of model structure on hydrological modelling uncertainty for streamflow simulation. *Journal of Hydrology,* 298, 242–266.

Buytaert, W., Beven, K. J. (2010). Models as multiple working hypotheses: hydrological simulation of tropical alpine wetlands. *Hydrological Processes.* 25, 1784–1799.

Cash, D. W., Clark, W. C., Alcock, F., Dickson, N. M., Eckley, N., Guston, D. H., Jäger, J., Mitchell, R. B. (2003). Knowledge systems for sustainable development. *Proceedings of the National Academy of Sciences* 100(14): 8086.

Cochrane, J. H. (1988). How big is the random walk in GNP? *Journal of Political Economy* 96: 893–920.

Cohen–Shacham, E., Walters, G., Janzen, C., Maginnis, S. (2016). *Nature–based solutions to address global societal challenges.* Gland, Suiza: Unión Internacional para la Conservación de la Naturaleza y de los Recursos Naturales.

Commission of the European Communities (CEC). (2009). *Impact assessment guidelines.* European Commission, Brussels, Belgium. [online] URL: https://ec.europa.eu/info/law/law–making–process/planning–and–proposing–law/impact–assessments_en

Cox, R. T. (1946). Probability, frequency and reasonable expectation. *American Journal of Physics.* 14 (1): 1–10.

Darcy, H. (1856). *Les Fontaines de la Ville de Dijon: exposition et application des principes à suivre et des formules à employer dans les questions de distribution d'eau.* Dalmont, Paris.

Davies, B. (2020). Precision and accuracy in glacial geology. Disponible en: https://www.antarcticglaciers.org/glacial-geology/dating-glacial-sediments-2/precision-and-accuracy-glacial-geology/

ECMWF. (2017). How to pin down uncertainty in weather forecasting. European Centre for Medium–Range Weather Forecasts. Reading, United Kingdom. URL: https://www.ecmwf.int/en/about/media–centre/news/2017/how–pin–down–uncertainty–weather–forecasting

Efron, B. (1979). Bootstrap methods: another look at the jackknife. *The Annals of Statistics.* 7 (1): 1–26.

Efron, B., Tibshirani, R. J. (1993). *An introduction to the Bootstrap.* CRC–Press.

Entekhabi, D., Njoku, G., O'Neill, P. E., et al. (2010). The Soil Moisture Active Passive (SMAP) Mission. *Proceedings of the IEEE.* 98(5): 704–716.

Fick, S. E., Hijmans, R. J. (2017). WorldClim 2: new 1 km spatial resolution climate surfaces for global land areas. *International Journal of Climatology.* 37 (12): 4302–4315.

Fleming, M. J., Doan, J. H. (2000). Geospatial Hydrologic Modeling Extension (HEC–GeoHMS). U. S. Army Corps of Engineers – Institute for Water Resources – Hydrologic Engineering Center (CEIWR–HEC). CPD–77.

Foster, M. E., Chen, D., Kieser, M. S. (2020). *Metodologías de cuantificación de beneficios hidrológicos de intervenciones en cuencas (CUBHIC).* Forest Trends, Kieser & Associates, LLC, Proyecto INSH.

García Hernández, J., Claude, A., Paredes, J., Roquier, B., Boillat, J–L. (2014). *Integrated flood forecasting and management system in a complex catchment area in the Alps – implementation of the MINERVE project in the Canton of Valais*. In A. Schleiss, Speerli, & Pfammatter (Eds.). Swiss Competencies in River Engineering and Restoration (pp. 87–97). London: Taylor & Francis Group.

García Hernández, J., Foehn, A., Fluixá–Sanmartín, J., et al. (2020). *RS MINERVE – Technical manual, v2.25*. Ed. CREALP, Switzerland. ISSN 2673–2661.
https://www.crealp.ch/fr/accueil/le–crealp/telechargement–doc/category/173–manuels–utilisateur.html?download=219:rs–minerve–technical–manual

Grayson, R. & Bloschl, G. (2000). Spatial modelling of catchment dynamics. In: Grayson, R., Bloschl, G. (Eds). *Spatial patterns in catchment hydrology: observations and modelling*. Cambridge: Cambridge University Press. p. 51–81 [chapter 3].

Gupta, H. V., Beven, K. J., Wagener, T. (2005). Model calibration and uncertainty estimation. In: Andersen, M. (Ed.). *Encyclopedia of Hydrological Sciences*, pp. 2015–2031. New York, NY: Wiley.

Gupta, H. V., Sorooshian, S., Yapo, P. O. (1999). Status of automatic calibration for hydrologic models: comparison with multilevel expert calibration. *Journal of Hydrological* Engineering. 4(2): 135–143.

Gupta, H. V., Kling, H., Yilmaz, K. K., Martinez, G. F. (2009). Decomposition of the mean squared error and NSE performance criteria: implications for improving hydrological modelling. *Journal of Hydrology*, 377(1–2), 80–91.

Haines–Young, R. H., Potschin–Young, M. B. (2010). The links between biodiversity, ecosystem services and human well–being. *Ecosystem ecology: a new synthesis*, 110–139.

Hall, J., Solomatine, D. (2008). A framework for uncertainty analysis in flood risk management decisions. *International Journal of River Basin Management,* 6: 85–98.

Hamel, P., Bremer, L. L., Ponette–González, A. G., Acosta, E., Fisher, J. R. B., Steele, B., Cavassani, A. T., Klemz, C., Blainski, E., Brauman, K. A. (2020). The value of hydrologic information for watershed management programs: The case of Camboriú, Brazil. *Science of The Total Environment* 705: 135871.

Harbaugh, A. W. (2005). MODFLOW–2005, The U.S. Geological Survey modular ground–water model – The Ground–Water Flow Process: U. S. Geological Survey Techniques and Methods 6–A16.

Harrison–Atlas, D., Theobald, D. M., Goldstein, J. H. (2016). A systematic review of approaches to quantify hydrologic ecosystem services to inform decision–making. *International Journal of Biodiversity Science, Ecosystem Services & Management* 12: 160–171.

Hastings, W. K. (1970). Monte Carlo sampling methods using Markov chains and their applications. *Biometrika* 57(1): 97–109.

Herschy, R. (2009). *Streamflow measurement*. 3rd edition. CRC–Press.

Hou, A. Y., Kakar, R. K., Neeck, S., et al. (2014). The Global Precipitation Measurement Mission. *Bulletin of the American Meteorological Society* 95(5): 701–722.

IISD. (2018). The multiple benefits of natural infrastructure. Disponible en: https://www.iisd.org/articles/multiple–benefits–natural–infrastructure

IPCC. (2014). *Fifth assessment report (AR5) of the United Nations Intergovernmental Panel on Climate Change* (IPCC).

ISO. (2013). ISO 18365:2013. Hydrometry Selection, establishment and operation of a gauging station.

Kavetski, D., Kuczera, G., and Franks, S. W. (2006). Bayesian analysis of input uncertainty in hydrological modeling: 1. Theory. *Water Resources Research* 42: W03407.

Kitanidis, P. K., Bras, R. L. (1980). Real–time forecasting with a conceptual hydrologic model: analysis of uncertainty. *Water Resources Research* 16(6): 1025–1033.

Klemeš, V. (1986). Operational testing of hydrological simulation models. *Hydrological Sciences Journal*. 31(1): 13–24.

Kolmogorov, A. (1933). Über die Grenzwertsätze der Wahrscheinlichkeitsrechnung. *Bull. (Izvestija) Acad. Sei. URSS*, 363–372.

Kummerow, C., Barnes, W., Kozu, T., Shiue, J., Simpson, J. (1998). The Tropical Rainfall Measuring Mission (TRMM) Sensor Package. *Journal of Atmospheric and Oceanic Technology*. 15(3): 809–817.

Li, Z., Zhu, Q., Gold, C. (2005). D*igital terrain modeling: principles and methodology*. Boca Raton: CRC Press.

Liu, Y., Gupta, H. V. (2007). Uncertainty in hydrologic modeling: toward an integrated data assimilation framework. *Water Resources Research* 43: W07401.

Locatelli, B., Homberger, J. M., Ochoa–Tocachi, B. F., Bonnesoeur, V., Román, F., Drenkhan, F., Buytaert, W. (2020). Impactos de las zanjas de infiltración en el agua y los suelos de los Andes: ¿Qué sabemos? Resumen de políticas, Proyecto Infraestructura Natural para la Seguridad Hídrica, Forest Trends, Lima, Perú.

Mace, G. M. (2008). Millennium Ecosystem Assessment. In: eLS. Chichester: John Wiley & Sons Ltd. Available at: http://www.els.net

Manz, B., Buytaert, W., Zulkafli, Z., Lavado, W., Willems, B., Robles, L. A., Rodríguez–Sánchez, J. P. (2016). High–resolution satellite–gauge merged precipitation climatologies of the tropical Andes, *J. Geophys. Res. Atmos.,* 121, 1190–1207.

Manz, B., Páez–Bimos, S., Horna, N., et al. (2017). Comparative ground validation of IMERG and TMPA at variable spatiotemporal scales in the tropical Andes. *Journal of Hydrometeorology.* 18(9): 2469–2489.

Mazzilli, N., Guinot, V., Jourde, H. (2012). Sensitivity analysis of conceptual model calibration to initialisation bias. Application to karst spring discharge models. *Advances in Water Resources 42,* 1–16.

McMillan, H. K., Westerberg, I. K., Krueger, T. (2018). Hydrological data and its implications. *WIREs Water,* 5: e1319.

Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H. (1953). Equations of state calculations by fast computing machines. *Journal of Chemical Physics.* 26(6): 1087–1092.

Millennium Ecosystem Assessment. (2003). *Ecosystems and human well–being: our human planet.* Washington, DC: Island Press.

Millennium Ecosystem Assessment. (2005). *Ecosystems and human well–being: current state and trends,* vol. 1. Washington, DC: Island Press.

MINAM. (2019a). Resolución No. 178–2019–MINAM. Lineamientos para la formulación de proyectos de inversión en las tipologías de ecosistemas, especies y apoyo al uso sostenible de la biodiversidad. Ministerio del Ambiente, Lima, Perú.

MINAM. (2019b). Mapa Nacional de Ecosistemas del Perú: memoria descriptiva. http://geoservidor.minam.gob.pe/wp–content/uploads/2017/06/MEMORIA–DESCRIPTIVA–MAPA–DE–ECOSISTEMAS.pdf. Ministerio del Ambiente, Lima, Perú.

Molina, A., Vanacker, V., Rosas–Barturen, M., Ochoa–Tocachi, B. F., Bonnesoeur, V., Román–Dañobeytia, F., Buytaert, W. (2021). Natural infrastructure interventions and their effect on soil erosion mitigation in the Andes. EGU General Assembly 2021: Gather Online, 19–30 April 2021. European Geosciences Union.

Montanari, A., Grossi, G. (2008). Estimating the uncertainty of hydrological forecasts: a statistical approach. *Water Resources Research 44,* W00B08.

Moriasi, D. N., Arnold, J. G., Van Liew, M. W., Bingner, R. L., Harmel, R. D., Veith, T. L. (2007). Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. *Transactions of the ASABE.* 50(3): 885–900.

Mosquera, G. M., et al. (2021). Progress in understanding the hydrology of high-elevation Andean grasslands under changing land use. Science of The Total Environment 804: 150112.

Muñoz, R., Huggel, C., Drenkhan, F., Vis, M., Viviroli, D. (bajo revisión). Comparing model complexity for glacio–hydrological simulation in the data–scarce peruvian Andes. *Journal of Hydrology: Regional Studies.*

Nash, J. E., Sutcliffe, J. V. (1970). River flow forecasting through conceptual models. Part I – A discussion of principles. *Journal of Hydrology*. 27(3), 282–290.

Neyman, J. (1937). Outline of a theory of statistical estimation based on the classical theory of probability. *Phil. Trans. R. Soc. Lond.* A. 236 (767): 333–380.

Ochoa–Tocachi, B. F., Buytaert, W., De Bièvre, B., et al. (2016a). Impacts of land use on the hydrological response of tropical Andean catchments. *Hydrological Processes* 30: 4074–4089.

Ochoa–Tocachi, B. F., Buytaert, W., De Bièvre, B. (2016b). Regionalization of land–use impacts on streamflow using a network of paired catchments. *Water Resources Research* 52: 6710–6729.

Ochoa–Tocachi, B. F., Buytaert, W., Antiporta, J., et al. (2018a). High–resolution hydrometeorological data from a network of headwater catchments in the tropical Andes. *Scientific Data* 5: 180080.

Ochoa–Tocachi, B. F., Torres, S., Ochoa–Tocachi, E. (2018b). *Diagnóstico, alcance y propuesta metodológica. Estudio del retorno sobre la inversión de las intervenciones del FONAG.* ATUK Consultoría Estratégica y Fondo de Protección del Agua (FONAG), Quito, Ecuador.

Ochoa–Tocachi, B. F., Bardales, J. D., Antiporta, J., et al. (2019a). Potential contributions of pre–Inca infiltration infrastructure to Andean water security. *Nature Sustainability* 2: 584–593.

Ochoa–Tocachi, B. F., Tenelanda–Patiño, D. O., Ochoa–Tocachi, E., et al. (2019b). *Modelo hidrológico FONAG fortalecido y automatizado. Generar y actualizar insumos y herramientas hidro–informáticas e hidroeconómicas para apoyar la toma de decisiones en la planificación hídrica del FONAG.* ATUK Consultoría Estratégica y Fondo de Protección del Agua (FONAG), Quito, Ecuador.

Ochoa–Tocachi, B. F., Alemie, T. C., Guzman, C. D., et al. (2019c). Sensitivity analysis of the parameter–efficient distributed (PED) model for discharge and sediment concentration estimation in degraded humid landscapes. *Land Degradation and Development 30:* 151–165.

Ochoa–Tocachi, B. F. (2020). *Análisis de sensibilidad de los modelos CUBHIC. Proyecto Infraestructura Natural para la Seguridad Hídrica.* Imperial College London y Forest Trends, Lima, Perú.

Ochoa–Tocachi, B. F., Tenelanda–Patiño, D. O., Ochoa–Pérez, P., et al. (2020a). *Módulo de calibración automática del modelo FONAG 2.1. Generar y actualizar insumos y herramientas hidro–físicas, hidrométricas, hidro–informáticas e hidroeconómicas para apoyar la toma de decisiones en la planificación hídrica del FONAG.* ATUK Consultoría Estratégica y Fondo de Protección del Agua (FONAG), Quito, Ecuador.

Ochoa-Tocachi, B. F. & Cuadros-Adriazola, J. (2022). CUBHIC 2.0: Metodologías de cuantificación de beneficios hidrológicos de intervenciones en cuencas – Manual de Uso. Proyecto Infraestructura Natural para la Seguridad Hídrica, Forest Trends, Lima, Perú.

Ochoa–Tocachi, B. F., Tenelanda–Patiño, D. O., Ochoa–Pérez, P., et al. (2020b). *Módulo de visualización gráfica y dinámica de resultados y análisis de sensibilidad para las salidas del modelo FONAG 2.1. Generar y actualizar insumos y herramientas hidro–físicas, hidrométricas, hidro–informáticas e hidroeconómicas para apoyar la toma de decisiones en la planificación hídrica del FONAG.* ATUK Consultoría Estratégica y Fondo de Protección del Agua (FONAG), Quito, Ecuador.

Ozment, S., Feltran–Barbieri, R., Hamel, P., et al. (2018). *Natural infrastructure in São Paulo's water system (World* Resources Institute: Washington DC, USA).

Palumbi, S. R., Sandifer, P. A., Allan, J. D., Beck, M. W., Fautin, D. G., Fogarty, M. J., Halpern, B. S., Incze, L. S., Leong, J–A., Norse, E., et al. (2008). Managing for ocean biodiversity to sustain marine ecosystem services. F*rontiers in Ecology and the Environment.* 7: 204–211.

Pianosi, F., Beven, K. J., Freer, J., et al. (2016). Sensitivity analysis of environmental models: a systematic review with practical workflow. *Environmental Modelling & Software.* 79: 214–232.

Pianosi, F., Sarrazin, F. & Wagener, T. (2015), A Matlab toolbox for Global Sensitivity Analysis. Environmental Modelling & Software, 70: 80–85.

Pianosi, F., Wagener, T. (2015). A simple and efficient method for global sensitivity analysis based on cumulative distribution functions. *Environmental Modelling & Software.* 67: 1–11.

Pianosi, F., Wagener, T. (2018). Distribution–based sensitivity analysis from a generic input–output sample. *Environmental Modelling & Software.* 108, 197–207.

Potschin, M. B. & Haines–Young, R. H. (2011). Ecosystem services exploring a geographical perspective, *Progress in Physical Geography* 35: 575–94.

Redhead, J. W., Stratford, C., Sharps, K. (2016). Empirical validation of the InVEST water yield ecosystem service model at a national scale. *Science of The Total Environment.* 569–580: 1418–1426.

Refsgaard, J. C. (1996). Terminology, modelling protocol and classification of hydrological model codes. In: Abbott M. B. and Refsgaard J. C. (Eds.). *Distributed hydrological modelling,* 17–39. Dordrecht: Kluwer.

Román, F., Estévez, G., Aste, N., Moles, A. (2020). Guía HIRO – Herramienta de identificación rápida de oportunidades para la infraestructura natural en la gestión del riesgo de desastres. https://www.forest–trends.org/publications/guia–hiro/

Ronald, L. Wasserstein & Nicole A. Lazar. (2016). The ASA's statement on p–values: context, process, and purpose. *The American Statistician,* 70:2, 129–133. doi: 10.1080/00031305.2016.1154108.

Scharffenberg, B., Bartles, M., Brauer, T., Fleming, M., Karlovits, G. (2016). *Hydrologic modeling system HEC–HMS user's manual.* U.S. Army Corps of Engineers – Institute for Water Resources – Hydrologic Engineering Center (CEIWR–HEC). CPD–74A.

Schervish, M. J. (1996). P values: what they are and what they are not. *The American Statistician.* 50 (3): 203–206. doi:10.2307/2684655. JSTOR 2684655.

Scholz, L. (1980). *Generación de caudales mensuales en la sierra del Perú.* Plan Meris II. Cuzco–Perú.

Sharp, R., Tallis, H. T., Ricketts, T., Guerry, A. D., Wood, S. A., Chaplin–Kramer, R., Nelson, E., Ennaanay, D., Wolny, S., Olwero, N., Vigerstol, K., Pennington, D., Mendoza, G., Aukema, J., Foster, J., Forrest, J., Cameron, D., Arkema, K., Lonsdorf, E., Kennedy, C., Verutes, G., Kim, C. K., Guannel, G., Papenfus, M., Toft, J., Marsik, M., Bernhardt, J., Griffin, R., Glowinski, K., Chaumont, N., Perelman, A., Lacayo, M., Mandle, L., Hamel, P., Vogl, A. L., Rogers, L., Bierbower, W., Denu, D., and Douglass, J. (2018). *InVEST user's guide.* The Natural Capital Project, Stanford University, University of Minnesota, The Nature Conservancy, and World Wildlife Fund.

Singh, V. P. (1995). Watershed modeling. In: *Computer models of watershed hydrology,* pp. 1–22. Highlands Ranch, CO: Water Resources Publication.

Singh, J., Knapp, H. V., Arnold, J., Demissie, M. (2004). Hydrologic modeling of the Iroquois river watershed using HSPF and SWAT. *JAWRA Journal of the American Water Resources Association,* 41: 343–360.

Sieber, A., Uhlenbrook, S. (2005). Sensitivity analyses of a distributed catchment model to verify the model structure. *Journal of Hydrology.* 310(1): 216–235.

Sieber, J., Purkey, D. (2015). *Water evaluation and planning system (WEAP) user guide*. Stockholm Environment Institute SEI–US Water Program.

Smirnov, N. V. (1939). On the estimation of the discrepancy between empirical curves of distribution for two independent samples. *Bull. Math. Univ. Moscow* 2: 2.

Solomatine, D. P., Wagener, T. (2011). Hydrological modelling (chapter 2.16). In: Treatise on water science, volume 2: *The science of hydrology*, 435–457.

Spear, R. C., Hornberger, G. M. (1980). Eutrophication in peel inlet—II. Identification of critical uncertainties via generalized sensitivity analysis. *Water Research*. 14(1): 43–49.

Sterne, J. A., Davey Smith, G. (January 2001). Sifting the evidence–what's wrong with significance tests? *BMJ*. 322 (7280): 226–31. doi:10.1136/bmj.322.7280.226.

SUNASS. (2019). Proyecto de estudio tarifario de EPS Moquegua S. A. (https://www.sunass.gob.pe/wp–content/uploads/2020/09/eps_moquegua_revision_tarifa_–2019_2022.pdf). Superintendencia Nacional de Servicios de Saneamiento, Lima, Perú.

Tang, G., Clark, M. P., Papalexiou, S. M., Ma, Z., Hong, Y. (2020). Have satellite precipitation products improved over last two decades? A comprehensive comparison of GPM IMERG with nine satellite and reanalysis datasets. R*emote Sensing of Environment* 240: 111697.

Tapley, B. D., Bettadpur, S., Ries, J. C., Thompson, P. F., Watkins, M. M. (2011). GRACE Measurements of mass variability in the earth system. *Science* 305: 503–505.

Thiemann, M., Trosser, M., Gupta, H., Sorooshian, S. (2001). Bayesian recursive parameter estimation for hydrologic models. *Water Resources Research* 37: 2521–2536.

Tung, Y. K. (1996). Uncertainty and reliability analysis. In: Mays L. W. (Ed.). *Water resources handbook,* pp. 7.1–7.65. New York, NY: McGraw–Hill.

UICN. (2020a). *Estándar global de la UICN para soluciones basadas en la naturaleza* (primera edición). Unión Internacional para la Conservación de la Naturaleza.

UICN. (2020b). *Orientación para usar el Estándar Global de la UICN para soluciones basadas en la naturaleza. Un marco fácil de usar para la verificación, diseño y ampliación de las soluciones basadas en la naturaleza* (primera edición). Unión Internacional para la Conservación de la Naturaleza.

Van Meerveld, HJ(I.), Jones, JPG, Ghimire, CP, et al. (2021). Forest regeneration can positively contribute to local hydrological ecosystem services: Implications for forest landscape restoration. J Appl Ecol. 2021; 58: 755– 765.

Vaze, J., Jordan, P., Beecham, R., Frost, A., Summerell, G. (2011). Guidelines for rainfall–runoff modelling: towards best–practice model application. *eWater Cooperative Research Centre*.

Virtual Terrain Project. (2021). Disponible en http //vterrain.org/Misc/usgs_data.html.

Vrugt, J. A., Gupta, H. V., Bouten, W., Sorooshian, S. (2003). A shuffled complex evolution metropolis algorithm for optimization and uncertainty assessment of hydrologic model parameters. *Water Resources Research* 39(8): 1201.

Vrugt, J. A., ter Braak, C. J. F., Clark, M. P., Hyman, J. M., Robinson, B. A. (2008a). Treatment of input uncertainty in hydrologic modeling: doing hydrology backward with Markov chain Monte Carlo simulation. *Water Resources Research* 44: W00B09.

Vrugt JA, ter Braak, C. J. F., Gupta, H. V., Robinson, B. A. (2008b). Equifinality of formal (DREAM) and informal (GLUE) bayesian approaches in hydrologic modeling? *Stochastic Environmental Research and Risk Assessment* 23: 1011–1026.

Wagener, T., McIntyre, N., Lees, M. J., Wheater, H. S., Gupta, H. V. (2003). Towards reduced uncertainty in conceptual rainfall–runoff modelling: dynamic identifiability analysis. *Hydrological Processes* 17(2): 455–476.

Willems, B., Leyva–Molina, W. M., Taboada–Hermoza, R., Bonnesoeur, V., Román, F., Ochoa–Tocachi, B. F., Buytaert, W., Walsh, D. (2021). Impactos de andenes y terrazas en el agua y los suelos: *¿qué sabemos? Resumen de políticas*, Proyecto Infraestructura Natural para la Seguridad Hídrica, Forest Trends, Lima, Perú.

Winsemius, H. C., Schaefli, B., Montanari, A., Savenije, H. H. G. (2009). On the calibration of hydrological models in ungauged basins: a framework for integrating hard and soft hydrological information. *Water Resources Research,* 45, W12422.

Woolhiser, D. A., Smith, R. E., Goodrich, D. C. (1990). KINEROS, *a kinematic runoff and erosion model: documentation and user manual.* U.S. Department of Agriculture, Agricultural Research Service, ARS–77.

WWAP. (2018). *The United Nations world water development report 2018: nature–based solutions.* Paris, UNESCO.

Photo: Denis Justo Mayhua Coaquira

Photo: Bruno Bernal

www.infraestructuranatural.pe

The Natural Infrastructure for Water Security (NIWS) project promotes the conservation, restoration and recovery of ecosystems across Peru, forming alliances with public and private organizations to reduce water risks such as droughts, floods and water pollution.

The project is funded and promoted by the United States Agency for International Development (USAID) and the Canadian Government; and executed by Forest Trends, CONDESAN, the Peruvian Society of Environmental Law (SPDA), EcoDecisión, and researchers from Imperial College London.

How to cite this document:

Ochoa–Tocachi, B. F., Cuadros–Adriazola, J., Arapa, E., Aste, N., Ochoa–Tocachi, E., Bonnesoeur, V. (2022). Guide for Hydrologic Modeling of Natural Infrastructure. Forest Trends, Lima, Peru.

USAID FROM THE AMERICAN PEOPLE
In partnership with Canada

FOREST TRENDS CONDESAN SPDA EcoDecisión Imperial College London atuk