Freshwater Health Index
User Manual
Version 1.1

October 6, 2017

TABLE OF CONTENTS

List of Tables

List of Figures

1. INTRODUCTION

The Freshwater Health Index is a decision support tool developed by Conservation International and partners, in collaboration with scientists, water resource and landscape managers, policy makers and the private sector, to help societies manage and conserve freshwater systems. Freshwater health is defined as the ability to deliver water-related ecosystem services, sustainably and equitably, at the drainage basin scale, thus linking the ecological function and condition of upstream areas of service generation with downstream communities. It is implicit that sustainable and equitable long-term delivery of ecosystem services relies on long-term ecosystem function. The Index addresses gaps in the prevailing water indicators by highlighting the relationships between healthy freshwater ecosystems, the flows of services they provide, and the role of governance and stakeholders in freshwater management and use. The Index can be used to evaluate scenarios such as climate variability, land cover change, population growth and water allocation decisions to make trade-offs more explicit and help stakeholders understand what policies and management practices are needed to maintain freshwater systems and service flows into the future. It is also intended to be used to track freshwater health over time. Thus, it requires an iterative process of dialogue between scientists, end-users and stakeholders so that the result is salient, credible and useful.

The Index is intended to measure the full range of benefits of freshwater systems, by making the connections between ecosystem health and service delivery more explicit, and thus helping stakeholders sustain and even enhance these services over time. Benefits of interest include water provision for agricultural, industrial and municipal uses as well as for power generation. But freshwater ecosystems also provide cultural services, including recreation and tourism opportunities, maintenance of biodiversity and habitats, and fisheries. These benefits are underpinned by critical regulating services that occur within the drainage basin, including the moderation of extreme events such as flooding and droughts, waste treatment and nutrient cycling, and erosion control. Inevitably, maximizing a particular suite of benefits entails making trade-offs, in terms of services and beneficiaries, and so the Index is designed to make these trade-offs explicit as well as highlight potential synergies.

The Freshwater Health Index focuses on three main components: ecosystem vitality, ecosystem services, and governance and stakeholders. Each component is assessed with a suite of measurable indicators that are aggregated into an index. Evaluation of the indicators requires using hydrologic and water allocation models as well as ecosystem service models, valuation techniques and stakeholder surveys. The intended scale of application is the drainage basin where resource management decisions have greatest relevance and decision support is likely to be the most useful. However, the framework and indicators are flexible and can be applied to smaller or larger spatial scales depending on stakeholder goals. The indicators can also be tailored to varying socio-political, economic and ecological contexts as well as data availability and informational needs.

This document provides guidelines to the application of the Freshwater Health Index. It explains the conceptual underpinnings of the Index and provides definitions for each of the indicators. It also provides guidance on how to evaluate each indicator, suggestions on data sources helpful to evaluate indicators (for baseline assessments and scenario planning), aggregation of indicators and interpretation of index values. We expect to review and update these guidelines periodically, and input from all users of the Freshwater Health Index is welcome. We particularly welcome new examples that are illustrative of these guidelines.

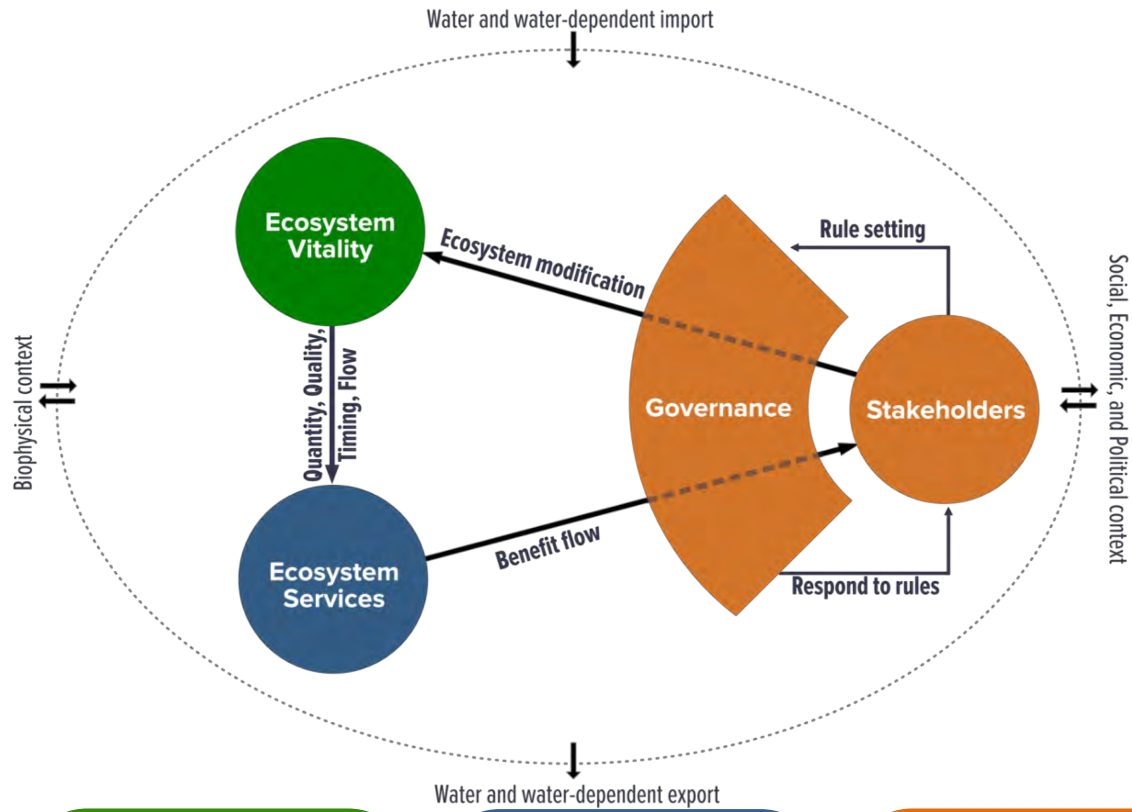
2. CONCEPTUAL FRAMEWORK AND FRESHWATER HEALTH INDEX

2.1 CONCEPTUAL FRAMEWORK

A robust conceptual framework is necessary to provide the foundation for development of a consistent and systematic set of indicators to measure the sustainability of freshwater systems (OECD 2008). A conceptual framework, in this context, is an abstract representation of complex freshwater systems that simplifies and highlights the key components and relationships between the social and ecological systems. The purpose of a conceptual framework is to characterize the multi-faceted and complex nature of these systems to capture generalities to be relevant across a wide range of systems, scales and time periods under a fitness-for-purpose principle (Shields and Rangarjan 2013). Hence, the conceptual framework should provide a clear description and definition of the multi-dimensional phenomenon to be measured, convey a structure of how the various components are connected and nested, highlight how the key components relate to freshwater health, and provide relevant structure and information to guide selection of measurable indicators (OECD 2008).

We invoke the concept of a freshwater social-ecological system here to illustrate the different dimensions that need to be measured to understand how social, hydrologic and ecological systems interact (Vogel et al., 2015; Vollmer et al., 2016). Many conceptual models have been developed for freshwater systems (Binder et al. 2013), however most lack a full accounting of the feedback between the ecosystems that capture, store and deliver water-based services, the beneficiaries of those services and how freshwater systems are governed and managed (Vollmer et al., 2016). We adapted a general framework for analyzing sustainability of social-ecological systems (Ostrom 2009) to tailor it more specifically to the measurement of freshwater sustainability. This general framework is desirable for systems with strong feedbacks between ecosystems, the services they provide and the beneficiaries of those services, as it treats the social and ecological components in almost equal depth and highlights the interactions between them (Binder et al. 2013). It also provides a structure that can facilitate indicator selection to describe the dynamics of social-ecological systems. The conceptual framework described here consists of three main components: “ecosystem vitality,” “ecosystem services” and “governance and stakeholders.” (Figure 1).

Figure 1. Conceptual framework for freshwater social-ecological systems comprised of Ecosystem Vitality, Ecosystem Services, and Governance and Stakeholders. Stakeholders set and adapt rules within governance and market systems and also respond to them. Within the constraints and rules set by water governance, stakeholders modify ecosystems through land-use change or conservation to exploit or manage freshwater ecosystems, and by developing infrastructure and technology to access water-based ecosystem services. Modifications to ecosystems and water withdrawals can alter the flow regime and water quality and thereby affect delivery of ecosystem services to beneficiaries. In basins where there are competing water needs, trade-offs become apparent and may necessitate an adjustment to governance mechanisms that can trigger changes in markets. Freshwater SESs are also impacted by external biophysical influences such as drought or climate change that affect ecosystem service delivery that can feed back to affect governance. Basins also are embedded within a broader social, political and economic context that can influence governance systems and, thus, management of fresh water. While we recognize that water and water-based goods and services may also be imported into or exported from a basin, our focus is primarily on interactions within the basin.



Ecosystem Vitality

Water quantity
- Deviation from natural flow regime
- Groundwater storage depletion

Water quality
- Suspended solids
- Total nitrogen
- Total phosphorus
- Indicators of major concern

Basin Condition
- Extent of channel modification
- Land cover naturalness

Biodiversity
- Change in number and trends of species of concern
- Change in number and trends of invasive and nuisance species

Ecosystem Services

Provisioning
- Water supply reliability relative to demand
- Biomass for consumption

Regulation and support
- Sediment regulation
- Deviation of water quality from benchmarks
- Flood regulation
- Exposure to water-associated diseases

Cultural
- Conservation/cultural heritage sites
- Recreation

Governance and Stakeholders

Enabling environment
- Water resource management
- Rights to resource use
- Incentives and regulations
- Capacity (financial and technical)

Stakeholder engagement
- Information access and knowledge
- Engagement in decision-making processes

Vision & adaptive governance
- Strategic planning and adaptive management
- Monitoring & learning mechanisms

Effectiveness
- Enforcement and compliance
- Distribution of benefits from ecosystem services
- Water-related conflict

“Ecosystem vitality” refers to the maintenance of “ecosystem structure and processes that underpin the capacity of an ecosystem to provide [water-based] goods and services" in the long term (MEA 2005, Turkelboom et al. 2014). Freshwater ecosystems include aquatic as well as terrestrial ecosystems linked within a watershed, encompassing both surface and groundwater.

Ecosystems produce a range of benefits to stakeholders ("ecosystem services") such as water provision, hazard mitigation and cultural services such as recreation opportunities (Haines-Young and Potschin 2010; 2013). Stakeholders operating within a governance system modify and manage the ecosystem to obtain certain services. Modifications might include channel and flow manipulations, pollution and remediation activities, as well as changes to the terrestrial ecosystem that have an impact on water-related services, such as land-use change that accelerates runoff downstream or habitat restoration to improve catchment and filtration. Hence, the structure and function of the ecosystem affects, and is affected by, the delivery of ecosystem services. Stakeholders operating within a governance system also build hard infrastructure to improve the delivery of ecosystem services or compensate for losses of naturally-provided services. These modifications to, and withdrawals from, the freshwater system can involve trade-offs among different objectives, different ecosystem services, beneficiary groups and generations (Rodríguez et al., 2006, Cai et al., 2002).

"Governance and stakeholders" is defined as "the structures and processes by which people in societies make decisions and share power, creating the conditions for ordered rule and collective action, or institutions of social coordination" (Schultz et al. 2015). This definition encompasses multiple tiers of governments, their formal rules and informal norms (e.g., community-established guidelines) and market mechanisms. It also encompasses a range of stakeholders comprising decision makers and the human beneficiary population (from individual citizens and community groups to municipalities, corporations and international organizations), as well as other stakeholders such as donor agencies, who may not directly benefit from the ecosystem services in a particular location, but nonetheless have an interest in, and influence over, decisions that affect a particular basin. The geographic range and makeup of stakeholders also changes according to the ecosystem service, e.g., beneficiaries of water-related recreation may live far outside of the basin generating the service. Stakeholders operate within the constraints of the governance system, which affects the behavior of stakeholders. In turn, stakeholders may influence or shape the governance system by modifying rules or changing the makeup of the system. While stakeholders and governance systems can be regarded as separate entities, for practical purposes they are combined to form a single set of indicators because of the heavy reliance of each on the other and the tight feedback that connects them.

Various forms of governance collectively provide the constraints and opportunities within which decisions are made, and then shape the consequences of these decisions (McGinnis, 2011). Here, we distinguish between governance systems directly related to water versus the broader social, economic or political context in which water governance lies. All variables (and their indicators) for the relevant governance system should be directly related to water, and these variables should be under the direct influence of at least some of the stakeholders. While general indicators, such as political stability, may be helpful as context and may indirectly influence water governance, they are not sufficiently specific to be tracked as a characteristic of the governance system of a basin.

Additionally, the freshwater system is affected by external biophysical stressors, e.g. climate change, drought and floods, as well as social, economic and political contexts, which operate at a scale larger than the watershed. Water or water-dependent products can be imported or exported to beneficiaries within and outside of the watershed. These aspects provide additional context for evaluating, monitoring and managing freshwater systems while not necessarily influencing indicator selection explicitly. This social-ecological conceptual framework is, in our view, the most appropriate for characterizing freshwater health because it provides an integrative conceptualization of the complex dynamics in social-ecological problems relating to sustainability. Its underpinnings are based on theories of collective choice, common-pool resources and natural resource management, and the ecological and social systems are treated in equal depth (Ostrom 2009; Binder et al. 2013). The social system (i.e., governance and stakeholders) operates at both the micro and macro level in a feedback loop: The micro level includes individual decision making whereas the macro level depicts the social system at the level of a population or society. Furthermore, the conceptual framework explicitly depicts the reciprocity, or feedback, between the social and ecological systems through specified interactions. It can be applied to multiple spatial scales (McGinnis and Ostrom, 2014), including watersheds and nations as well as the global freshwater system (Vogel et al., 2015).

The conceptual framework formed the basis upon which the indicators for the Freshwater Health Index were developed. Selection criteria are typically used to ensure that indicators are relevant and meet the overarching purpose of the conceptual framework. Numerous criteria have been proposed for a variety of indicator frameworks (e.g. Smith and Zhang 2004; SWRR 2005; OECD 2008), and these were used as a starting point for the development and refinement of criteria for indicator selection to measure freshwater health. The following criteria were ultimately applied to ensure relevance, accessibility and soundness of the resultant set of freshwater health indicators:

- indicators must be measurable, unbiased and defensible,
- choice of indicators must be relevant and guided by the conceptual framework,
- indicators must be relatively easy to understand,
- indicators must be based on information that can be used to compare different geographical areas and contexts,
- indicators must be distinct, i.e., an indicator does not measure the same process or quantity as another indicator,
- indicators or their combinations should be limited in number to provide a clearer signal of progress,
- indicators must be sensitive to changes over time and space to detect change.

Using the conceptual framework and the criteria for indicator selection above, three sets of indicators were identified. These are defined in section 2.3, and guidelines on their application are provided in sections 4, 5 and 6.

2.2 SCALE OF APPLICATION

Spatial scale

The Freshwater Health Index can accommodate a range of spatial scales: sub-basins, basins, regions of adjacent basins, nations and even global assessments. For greatest utility for management, we recommend that the Index be applied to basins represented as a network of connected sub-basins (see section 3 below). In some contexts, some indicators may be most appropriately considered at a different spatial scale than the scale of assessment. For instance, some governance indicators may be best considered at the national scale even when the Index is being evaluated at the basin scale. In such cases, relevant national information can be applied to the smaller scale if local data is unavailable or local governance is dictated by national processes. For transboundary basins, assessments may need to consider information across multiple nations to determine the best and most relevant information to use for a basin-scale assessment. In many cases, however, data will be available and most relevant at a scale smaller than the assessment scale (e.g., water quality data at particular point source locations), in which case data and/or indicator values can be represented at the disaggregated scale through maps but will need to be aggregated to the basin scale for final numeric values. Moreover, scales finer than the sub-basin may need to be considered to detect any changes in the indicator over time. For instance, land cover naturalness may need to be evaluated at a 30m scale to detect meaningful changes from one assessment period to the next.

Temporal scale

The indicators that comprise the Freshwater Health Index measure the status and/or trends of freshwater system attributes. For calculations on the status of an attribute, the most recent year for which data is available should be used. For current status calculations, the dates of the most recent available data do not have to be consistent across all indicators, although care should be taken to use data sets that reflect current conditions. For instance, data collected on water quality 10 years prior to the assessment date is unlikely to be indicative of current water quality, but 10-year-old data might be reflective of current land cover if it is known that land-use change has been negligible in the intervening years. Hence, it is important to use the most current data wherever available. For calculations of trends in freshwater system attributes, the current value needs to be compared with an historic value. For the first iteration of trend calculations, we recommend using five years prior to the assessment date as the reference point to which current values are compared, or as close to this as available data allows. For example, a first assessment of decline in a species of concern undertaken in 2020 should use 2015 as the reference year to which 2020 population size is compared. All subsequent iterations of trend calculations should compare the current value with the previously calculated value. We recommend that the Freshwater Health Index be re-evaluated at least every five years.

2.3 INDICATORS AND DEFINITIONS

2.3.1 Ecosystem Vitality

Table 1. Ecosystem Vitality indicators

Major indicators	Sub-indicators
Water quantity	Deviation from natural flow regime Groundwater storage depletion
Water quality	Suspended solids in surface water[1] Total nitrogen in surface and groundwater[1] Total phosphorous in surface and groundwater[1] Indicators of major concern[2]
Drainage-basin condition	Bank modification Flow Connectivity Land cover naturalness[3]
Biodiversity	Changes in number (i.e., species number) and population size of species of concern

Changes in number and population size of invasive/nuisance species

1. Deviation of concentration from environmental benchmark related to local historic natural conditions.
2. Optional; depends on local conditions and could include salinity, dissolved oxygen, pH, electrical conductivity, total dissolved solids, heavy metals and coliforms, as well as pharmaceuticals and other contaminants.
3. Naturalness here is measured on a gradient from completely natural (e.g., primary forest) to completely artificial (e.g., urban areas).

Water quantity assesses the stock and flow of water through the drainage basin and changes in water-storage capacity.

> *Deviation from natural flow regime* measures the degree to which current flow conditions have shifted from historic natural flows. The greater the deviation from natural flow indicates higher risk for the freshwater ecosystem (Poff and Zimmerman 2010). This measure can be derived from a wide range of variables, including deviation in annual mean, minimum and maximum discharge in the basin, proportion of the year that annual mean discharge was exceeded, etc.
>
> *Groundwater storage depletion* measures changes in the availability of water stored in underground aquifers (Konikow and Kendy 2005). This can be directly estimated using records of groundwater level obtained from observation wells or via the indirect proxy of aquifer compression resulting from groundwater over-exploitation or using data derived from Gravity Recovery and Climate Experiment (GRACE) satellites.

Water quality measures the state of water quality in the basin relevant for maintaining healthy aquatic ecosystems, rather than for human consumption.

> *Suspended solids in surface water*, *total nitrogen and total phosphorous* are all critical parameters that provide a measure of water quality with respect to its impact on biodiversity and ecosystem health in a basin (UNEP 2008a). These should be measured as a deviation from an established environmental baseline, which may be derived from the basin's historic natural conditions or the physiological tolerances of native aquatic species of concern.
>
> *Other indicators of major concern* for water quality can include temperature, salinity, dissolved oxygen, pH, electrical conductivity, total dissolved solids, heavy metals and coliforms, pharmaceuticals and other contaminants. Choice of what to measure can be based on local requirements and capabilities, following recommendations by the United Nations (UNEP 2008a). However, they should be selected judiciously as and when data and analytical capacity are available, selecting those variables that are known to have the greatest potential impact on freshwater ecosystem health. All measures selected in this sub-indicator category should be weighted with weights summing to 1.0.

Drainage Basin Condition measures the extent of physical modifications to the drainage basin and flow network. Such changes result in habitat degradation that impacts biodiversity.

> *Bank Modification* measures what is known as floodplain (lateral) connectivity. Lateral connectivity affects how the streams reach land and thus how materials such as nutrients and sediments are exchanged. Changes to this pattern, either through channelization or inundation

through impoundments, affect the suitability for native vegetation and wildlife (including spawning fish and water birds), the biogeochemistry of the streams, as well as the extent of floodplains.

Longitudinal or *flow connectivity*, also known as fragmentation, is particularly important to the movement of aquatic life such as fish, but also affects the flow of materials. It is affected by natural obstructions such as waterfalls, and engineered structures such as dams and weirs. Decreased longitudinal connectivity can negatively impact fish migration and reproduction, and may prevent sediment and other nutrients from being delivered downstream.

Land cover naturalness measures the amount of human-induced transformation of the landscape, using a gradient ranging from completely natural to completely artificial (Angermeier 2000). A basin in its undisturbed state, with intact forests and wetlands, generally maintains a sufficient quantity and quality of water to support indigenous flora and fauna. Human conversion of lands and waterways are associated with increases in pollutant loads (non-point source from agriculture, point-source from urban and industrial), changes to infiltration and runoff regimes and losses of regulating services (e.g., flood mitigation, erosion prevention, water purification, etc.).

Biodiversity assesses potential shifts in ecosystem functioning by measuring changes in the constituent biota that are integral components of freshwater ecosystems. The status and trends of biodiversity in a basin signify ecosystem health, with declining populations of native species and increasing populations of invasive and nuisance species indicating a deteriorating ecosystem (Dudgeon et al. 2006). This biodiversity indicator is comprised of changes in the following:

Species of concern consist of threatened aquatic or riparian species and species of interest (such as keystone or umbrella species) that will be affected by changes in habitat condition. The number of such species, the change in this number over time and their population trends over time are of interest here.

Invasive and nuisance species in lakes, waterways and the riparian zone indicate anthropogenic alteration of ecological conditions, as these are the circumstances which allow alien species to thrive at the expense of native species. The number and changes in the number of species present and their population trends are of interest.

2.3.2 Ecosystem Services

Table 2. Ecosystem Services indicators

Major indicators	Sub-indicators
Provisioning	Water supply reliability relative to demand
	Biomass for consumption[1]
Regulation and support	Sediment regulation
	Deviation of water quality metrics from benchmarks[2]
	Flood regulation
	Exposure to water-associated diseases
Cultural/aesthetic	Conservation/cultural heritage sites
	Recreation

1. Optional; include depending on local conditions
2. Refers to ability of the freshwater ecosystem to deliver water of the expected water-quality standards for different sectors.

Provisioning measures the material outputs from freshwater ecosystems that are used for human benefit.

> *Water supply reliability relative to demand* is calculated as the net water demand from various sectors (municipal, industry, agriculture, hydropower), the environmental and, where relevant, navigational flow requirements with respect to total freshwater availability. This indicator takes into account the reliability and variability, or seasonality, of freshwater supply relative to demand (Brown and Lall 2006; Grey and Sadoff 2007). The probability that an ecosystem meets demand is dependent on a combination of system attributes, inflows and demands whereby a system may, for example, be water-rich during wet seasons but water-deficient in dry seasons.
>
> *Biomass for consumption* measures the availability of fisheries, wild food, fiber and other materials from freshwater systems for human consumption (TEEB 2011). The availability of these ecosystem services relies on the availability of adequate quantities and quality of fresh water, and may very likely be affected by seasonal patterns of flow (e.g., fisheries yields can be affected by the extent of flood-plain inundation).

Regulation and support measures the regulating, maintenance and support aspects of freshwater ecosystems that provide benefits to people beyond provisioning (de Groot et al. 2002).

> *Sediment regulation* measures the degree to which drainage basins regulate erosion and control sediment dynamics (transportation and deposition) as well the nutrients that may be bound to transported particles. This has important implications for agricultural productivity (especially of floodplains) and supply of particles to deltas and nutrients to coastal waters (Le et al. 2007). Conversely, high rates of erosion, siltation and sedimentation have negative impacts on in-stream habitat conditions, biodiversity and river infrastructure (Ward and Stanford 1995).
>
> *Deviation of water quality metrics from benchmarks (water regulation)* indicates the ability of the freshwater ecosystem to deliver water of the required water-quality standards for different sectors and, thus, can be used as a proxy measure of the regulatory service of filtration and water purification (de Groot et al. 2002).
>
> *Flood regulation* measures the extent to which the condition and functioning of a river basin is damaged through exposure to floods (MEA 2005).
>
> *Exposure to water-associated diseases (disease regulation)* measures the prevalence of water-associated diseases such as typhoid fever, cholera and infection by parasites (e.g., schistosomiasis and malaria) (Prüss-Üstün et al. 2008). Disease risks can be increased by modifications to freshwater habitats such as dam construction, canalization, stagnation due to altered flow and human waste contamination) (Naiman and Dudgeon 2011; Dickin et al. 2013).

Cultural and aesthetic indicators measure the cultural, aesthetic, spiritual and other socio-cultural values of a freshwater system that are important to people (Daniel et al. 2012).

Conservation and cultural/heritage sites represents the water-related natural resources and structures that are under protection or formal management for science, culture, religion or other values (e.g., World Heritage Sites, biodiversity/national parks). This represents the societal importance of water-related features for scientific, cultural, religious and aesthetic or existence values (TEEB 2011; Tengberg et al 2012).

Recreation measures the degree to which fresh water has societal value in the form of recreational and tourism opportunities such as hiking, camping, boating, angling, etc. These can be measured by the number of tourists/recreational visits to water-related sites or the amount of revenue generated within a basin by such activities.

2.3.3 Governance and Stakeholders

Table 3. Governance & Stakeholders indicators

Major indicators	Sub-indicators
Enabling environment	Water resource management Rights to resource use Incentives and regulations Financial capacity Technical capacity
Stakeholder engagement	Information access and knowledge Engagement in decision-making processes
Vision and adaptive governance	Strategic planning and adaptive governance Monitoring and learning mechanisms
Effectiveness	Enforcement and compliance Distribution of benefits from ecosystem services Water-related conflict

Enabling environment refers to the constraints and opportunities that are enshrined by the existing institutional framework (policies, regulations, market mechanisms and social norms) and the financial and technical capacity available to carry out mandates (Moglia et al. 2011).

Water resource management measures the degree to which institutions (formal and informal) are responsible for performing the water resource management functions of monitoring and coordination, planning and financing, developing and managing infrastructure, and resolving conflict (Global Water Partnership, 2009). Closely aligned with measures of Integrated Water Resource Management (IWRM) implementation (e.g. WWAP, 2015), it acknowledges that multiple entities may be involved in fulfilling these functions (Hooper, 2010).

Rights to resource use measures the coherence of the existing rights to resource use, including the coexistence of customary and formal rights. Clear resource use rights are recognized as a precondition to efficient use of scarce resources and as a means of forestalling or settling disputes (Gleick, 1998). Relevant rights include those that govern the various uses and users of surface and groundwater, water emissions and pollution, fishing permits and land-use zoning to safeguard waterways (e.g., riparian buffer zones).

Incentives and regulations measures the availability of different management instruments, including conventional command and control regulations, investment screening criteria, tax incentives and market-based instruments (payments for ecosystem services, water rights trading). In principle, more diverse instruments equate to more flexibility and efficiency in maximizing societal benefits at lower cost (Lemos and Agrawal 2006).

Financial capacity measures the investment gap in water resource protection measures as well as the capacity of skilled professionals working in water resource management fields (Ivey et al. 2004). The investment gap refers specifically to actual budget allocations versus official estimates of investment needed for water distribution networks, treatment, and wetland and ecosystem conservation. Even where financial resources may be available, there may be a shortage of qualified, adequately trained people to carry out the water resource management functions outlined above.

Technical capacity measures the number and skill level of professionals working in water resource management.

Stakeholder engagement refers to stakeholder interactions, their ability to engage in decision making processes and the degree of transparency and accountability that govern these interactions.

Information access and knowledge measures the access (and uptake) all stakeholders have to information including data on water quantity and quality, water resource management and development documents and relevant financial information. Even where data and information are abundant, to be effective, they must be analyzed and applied in decision making processes and made accessible (across agencies, with citizens, etc.) in ways that are understandable to various stakeholders (Burroughs 1999).

Engagement in decision-making processes measures the degree to which all stakeholders have a voice within the cycle of policy and planning. Engagement can be evaluated on a continuum where stakeholder influence increases, from unidirectional communication to consultation, representation and, eventually, co-decision and co-production (OECD 2015). Increased engagement is associated with improved information transfer, better targeted and more equitable plans and policies, improved transparency and accountability, and reduced conflict.

Vision and adaptive governance aims to measure stakeholders' capacity to collect and interpret information, and then use this information to set goals for the basin and adapt to changing circumstances.

Strategic planning and adaptive governance measures the degree to which stakeholders engage in comprehensive strategic planning at the basin or sub-basin scale, and whether they have the capacity to adapt plans to new information or changing conditions. For instance, the basin-specific goals of IWRM are articulated in a river basin management plan and such plans should have well-defined objectives, mutually agreed to goals, and long-term resource development priorities to foster sustainability of freshwater systems (Hooper 2010).

Monitoring and learning mechanisms measures the adequacy and uptake of monitoring programs and information. Assessments of freshwater status and decisions about water resource development projects are best based on sound data and information that allows comparison through time (Pahl-Wostl et al. 2013). Monitoring includes physical, chemical, and biological properties of water resources, along with water-related socioeconomic and financial data.

Effectiveness examines the governance components that are being implemented and whether they are leading to the expected outcomes. A narrow definition of effectiveness is applied here, focusing on key indicators of governance failures (Rogers and Hall 2003): implementation gaps, inequitable distribution of benefits and the presence of water-related conflicts.

Enforcement and compliance considers the degree to which laws are upheld and agreements are enforced. In many societies, a gap exists between laws and their enforcement, reflecting either insufficient capacity or a lack of accountability (and possibly engendering corruption (Tropp 2007).

Distribution of benefits from ecosystem services measures the impact of decisions about water resource management with special attention to vulnerable populations, gender and resource-dependent communities. It is a proxy indicator for equity, which is commonly ascribed to principles of "good governance" but is subject to local interpretations (UN-Water 2015; Pahl-Wostl 2015).

Water-related conflict measures the presence of conflicts over water services, including allocation and diversion decisions, infrastructure development and access to resources. Tensions among stakeholders, particularly in transboundary settings, are to be expected when competition for water services and complex interactions occur within a basin. An effective governance system should prevent tensions from escalating into conflicts (UN-Water 2013).

2.4 INDICATOR AGGREGATION

Calculation of the indicators for a basin or sub-basin will rely on established hydrologic models, local knowledge and data, as well as global datasets and stakeholder surveys (see sections 4, 5 and 6 for recommended calculations and data sources). Once indicator values are calculated, they should be normalized to a common scale (we recommend 0 to 100, with 100 being more "sustainable" and 0 being less "sustainable").

Aggregation occurs in two steps: 1) spatial aggregation of individual indicators across sub-basins (where relevant) to provide basin-level indicator values, and 2) aggregation across all evaluated basin-level indicators to provide the index value for a given component. In the first step, where individual indicators evaluated at the sub-basin level are aggregated to the scale of the basin, sub-basin indicator values should be weighted according to the proportion the sub-basin area makes up of the total basin area. These weights should then be normalized to sum to 1. Sub-basin values are then aggregated as the weighted geometric or arithmetic mean of all sub-basins for which the indicator was evaluated.

In the second step, where indicators are aggregated at the basin level to form a component index, weights can be applied to each indicator to denote greater or lesser importance of the role of the indicator for assessing freshwater health in the basin. There are a variety of methods for assigning weights including, but not limited to, expert elicitation, the Delphi method (Brown 1968, http://www.rand.org/topics/delphi-method.html) or the Analytic Hierarchy Process (Saaty 1990). Each method encourages stakeholder participation. It is not necessary to apply weights to the indicators at this aggregation step; they should only be applied if there is good reason to believe that the indicators play disproportionate roles in measuring freshwater sustainability. In particular, we emphasize caution in applying them to the Ecosystem Vitality indicators. For these indicators, weights should only be applied if there is strong evidence that some ecosystem processes or attributes play a greater role in ecosystem functioning. This is an empirical question rather than a subjective one.

Values should then be aggregated to give a central tendency (e.g., arithmetic or geometric mean) to provide a separate index for each of the three components of Ecosystem Vitality, Ecosystem Services, and Governance and Stakeholders. We recommend using the geometric mean over the arithmetic mean as it has the desirable property of being more sensitive to changes in multiple indicator values. For instance, under the arithmetic mean, a change for the better in one indicator can be offset by a change for the worse in another, resulting in no change in the aggregated index. Under the geometric mean, such a change would be reflected in the index value. The indices should not be further aggregated across the three components due to differences in their resultant interpretations and methods of evaluation. For instance, we recommend conducting surveys to derive subjective values for the Governance and Stakeholders indicators, whereas the indicators within Ecosystem Vitality and Ecosystem Services should be based on empirical data and models wherever possible. Furthermore, treating the indices for the three components separately can highlight where the greatest problems, or the greatest contributors to sustainability, lie.

2.5 INTERPRETATION OF INDEX VALUES

High values across all three of the aggregated indices are indicative of a more sustainable freshwater ecosystem (or greater freshwater health) than indices at the lower end of the scale. Temporal changes in index values upon repeated application reflect either improvement toward, or deviation from, freshwater sustainability. Such changes can also show the effects of management interventions intended to improve service delivery or ecosystem health in scenario analysis. The value of presenting indices for each of the three components is that it clarifies which components are functioning better toward freshwater sustainability and which need further attention to increase overall sustainability.

By calculating these individual indices, a baseline evaluation of freshwater health can be established, along with a mechanism to highlight the trade-offs between services and beneficiaries that are likely to occur at any scale or location. A snapshot of current conditions offers a clear picture of freshwater health that can be repeated over time, updating results and examining scenarios such as interventions intended to improve service delivery or ecosystem health. The indicators can also be used to evaluate future scenarios of land-use change, infrastructure development and climate change as well as other drivers that threaten ecosystem function and service delivery.

2.6 DOCUMENTATION

All assessments should be documented. A narrative overview of the basin should be provided, outlining the basin classification (see Section 3.1), ecosystem services, relevant stressors, main stakeholders and any other pertinent information that illuminates the context of freshwater health and use in the basin.

This will enable people to understand the primary freshwater issues in the basin and may aid in management decisions. Values for each sub-indicator should be reported (pre- and post-normalization), along with the data and models (where relevant) used to evaluate the indicator. In cases where a sub-indicator cannot be evaluated, the reasons should be clearly stated, e.g., the indicator is not relevant for the basin or data is unavailable for the basin. Clearly documenting the calculation of the indicator values provides the reasoning behind an overall index value and, when warranted, the reasoning can be updated or used as the basis for future assessments. In sections 4, 5 and 6, further guidance is provided on the specific documentation expectations for each indicator.

2.7 OVERVIEW OF ASSESSMENTS

Months 1-3

- Review existing datasets
- Establish contact with technical collaborators (e.g., universities, regional government agencies' tech staff)
- Preliminary review of stakeholders within the basin and draft "theory of change"

Months 4-6

- Perform initial calculations of indicators based on existing data
- Establish technical collaborations to calculate indicators with local data
- Organize 1-2 technical meetings with collaborators to review initial results

Months 7-9

- Interviews with individual stakeholder agencies/organizations to administer governance questionnaire
- Stakeholder consultation forum to introduce the Freshwater Health Index
- Review collaborators' calculations and model scenarios

Months 10-12

- Prepare assessment report and policy summary
- Stakeholder forum to discuss results
- Prepare plan for technical collaborators to conduct subsequent assessment (e.g., in 3-5 years)

3. BASIN CHARACTERIZATION

3.1 CLASSIFYING BASIN ATTRIBUTES AND CONTEXT

Characterizing the basin is a first important step to understanding the relationships between ecosystems, ecosystem services and governance and to contextualizing the indicators. There is no globally accepted drainage basin classification system. Based on the components of Ecosystem Vitality, Ecosystem Services and Governance and Stakeholders, the following breakdown of dominant freshwater resource features, management priorities, IWRM implementation and climatic classes in Table 4 can be used to classify the basin. In many cases, the classes within each category are expected to overlap. For example, "Human consumption focused" and "Resource generation focused" in the category *Management priority* will invariably overlap depending on the context. In such cases, the sub-categories should be ranked in order of most to least prominent.

Table 4: Basin categories and classification

Category	Classification	Rationale
A. Dominant freshwater resource feature	A1 River A2 Groundwater A3 Lake A4 Wetland	The dominant water resource feature in the basin will directly influence the "ecosystem vitality" indicators of a study.
B. Management priority	B1 Human consumption focused (specify the sectors) B2 Resource generation focused (specify the sectors) B3 Extreme event focused (specify the events)	The management priority of the Freshwater Health Index implementing body as well as local stakeholders will be heavily influenced by the main concerns of these agencies and the ecosystem services they want to secure. "Human-consumption focused" basins have as a priority provision of water and food for residents **within** the drainage basin itself, while "Resource-generation focused" basins invoke primarily economic reasons for hydropower, tourism, industrial and agriculture production in the basin (e.g., plantations in Indonesia). "Extreme-event focused" reflects the management priority to manage extreme events (e.g., floods).
C. Level of IWRM implementation	C1 Moderate to high C2 Non-existent to low	Governance systems and stakeholder actions impact the sustainability of freshwater in a basin. The level of implementation of IWRM plans can indicate the alignment of governance and stakeholders processes with mechanisms to foster freshwater sustainability. Here class C1 corresponds to UNEP-DHI IWRM implementation study scores >3.5.
D. Climatic classes	D1 Tropical D2 Dry D3 Mild Temperate D4 Snow	Climate provides the broad context of the drainage basin, the extreme weather events the region may be prone to and the variability in weather and therefore water availability. The major classifications by Koppen may be used for this hanschen.org/koppen/.

Table 5 below lists examples of drainage basin of each type. The global maps from the River and Groundwater Basins database can help identify basin classes for category A. Classification under categories C and D can be identified from the Status of IWRM implementation database and the Koppen Climate Classification, while A and B can be refined from initial investigation of the nature and processes in the drainage basin. Thus, a drainage basin can be classified as one of among 96 types (4x3x2x4) although some combinations will be non-existent. The Dongjiang basin, for example, is of type [A1, B1, C1, D3]. This information should be documented, along with a narrative of the types of ecosystem services, relevant water governance and stakeholder processes and a description of the ecosystem and its stressors, as a systematic way to provide contextual information about the basin. As assessments are conducted around the world, they can be catalogued according to this classification system so that users can identify, learn from and compare experiences with basins in a similar context.

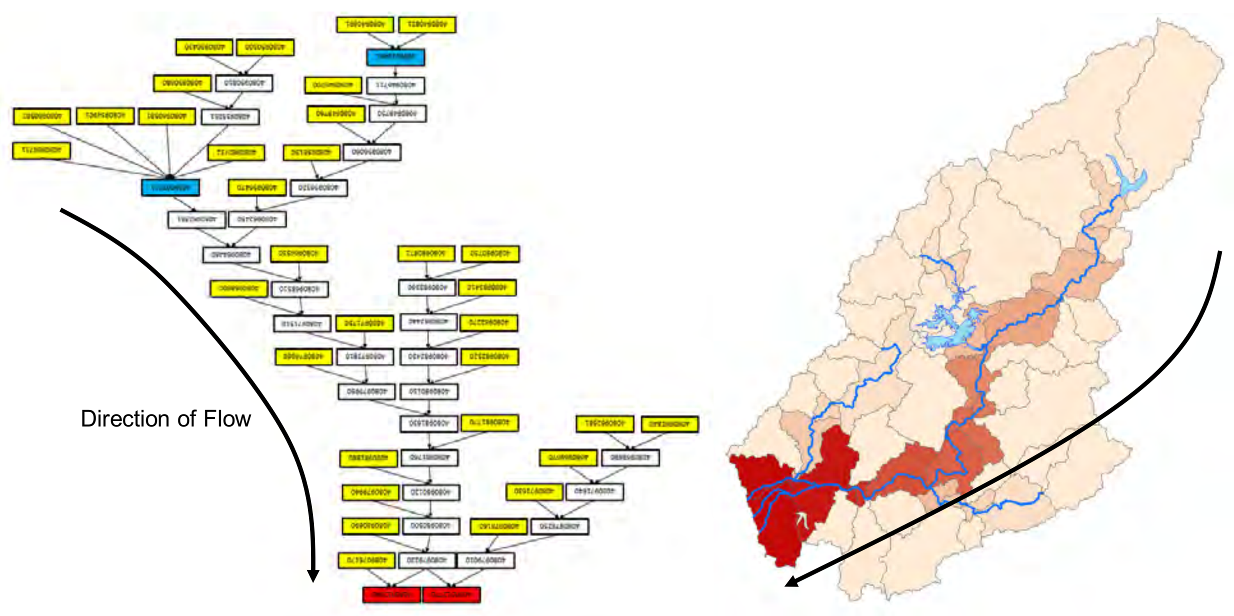
Table 5. Examples of basin classifications

Classification	**Examples**
A1 River	Mekong (Transboundary), Dongjiang (China)
A2 Groundwater	Small Island nations (Pacific)
A3 Lake	Lake Victoria (Africa)
A4 Wetland	Everglades (USA)
B1 Human consumption focused	Dongjiang (China)
B2 Resource generation focused	Amazon (Transboundary), Small watersheds in Indonesia
B3 Extreme event focused	Ciliwung (Indonesia)
C1 Moderate to high	Elbe (EU), Thames (UK)
C2 Non-existent to low	Ganges (India) , Indus (Pakistan)
D1 Tropical	Ciliwung (Indonesia), Amazon (Transboundary)
D2 Dry	Murray-Darling (Australia)
D3 Mild Temperate	Dongjiang (China)
D4 Snow	The Great Lakes (North America)

3.2 BASIN MODELING FRAMEWORK

Carrying out an assessment will require not only defining a spatial and temporal scale for implementation (Section 2.2), but also assessing the distribution of and connection between the indicators of interest (listed in Section 2.3.1 and 2.3.2) within the area of implementation. To reflect the upstream and downstream flow connectivity that drives the physical processes within a basin, delineation of the basin into a hierarchy of sub-basins based on terrain and flow network is an appropriate representation of the physical network underpinning the basin. The indicators of interest are then intended to be calculated at, and aggregated from, information available at this sub-basin level, if possible. Selecting the level of sub-basin delineation is a subjective decision, with higher levels of delineation allowing emergence of a more detailed picture of processes and trade-offs operating within the basin. This comes at a cost of increased effort required for higher resolution monitoring and modelling to calculate the indicators. An example of a straightforward but coarse delineation would be delineating the basin into upstream areas contributing to runoff and downstream areas of water-supply consumption. We recommend using the HydroBASIN database (Figure 2; Lehner and Grill, 2013) to construct the sub-basin network within the basin.

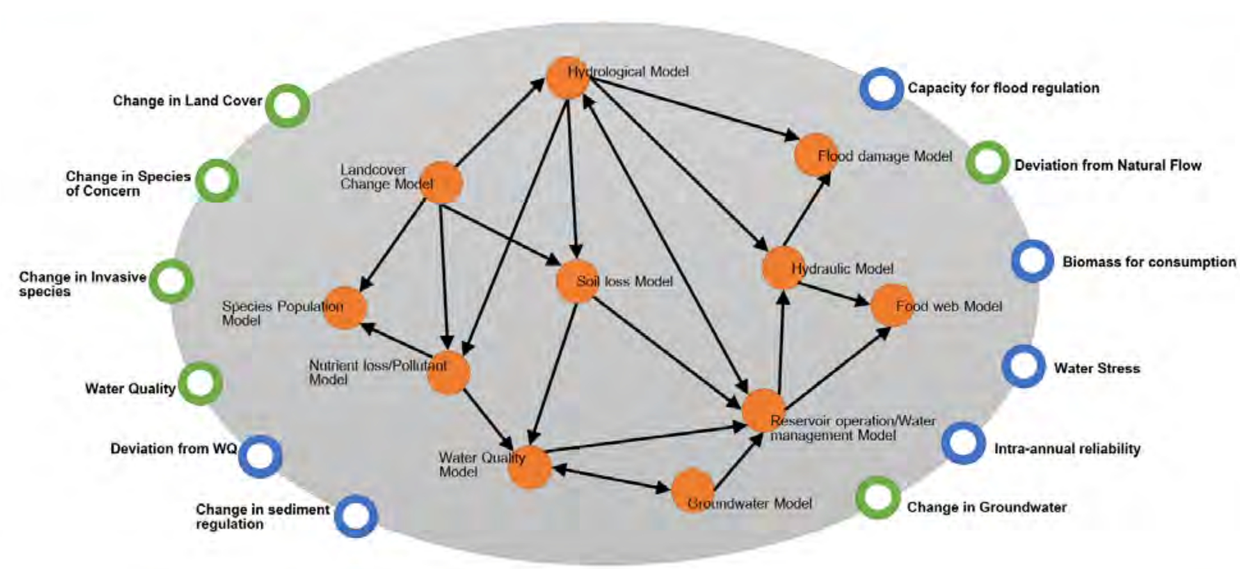
Figure 2. Flow network for the main river channel of the Dongjiang Basin constructed using the HydroBASIN (level 8 modified) database. In left, sub-basins shaded red drain to the sea, shaded yellow have no sub-basin upstream and shaded light blue are the two main reservoirs.



Network within Basin with Sub-basins as nodes **Basin and Sub-basins Outline**

The sub-basin network can then form the foundation for process models informing calculation of the Ecosystem Vitality and Ecosystem Service indicators. Many of the models considered for calculating the indicators will necessarily be coupled due to the interdependency of the processes they model. Figure 3 shows such a possible chain of models, results from which are summarized by the indicators for Ecosystem Services and Ecosystem Vitality. While these links between processes may not be apparent in the indicators themselves, they capture some of the direct quantifiable trade-offs within the basin. T ability to adjust these trade-offs in response to scenarios will be of value when identifying the models and monitoring processes required for Freshwater Health Index assessment.

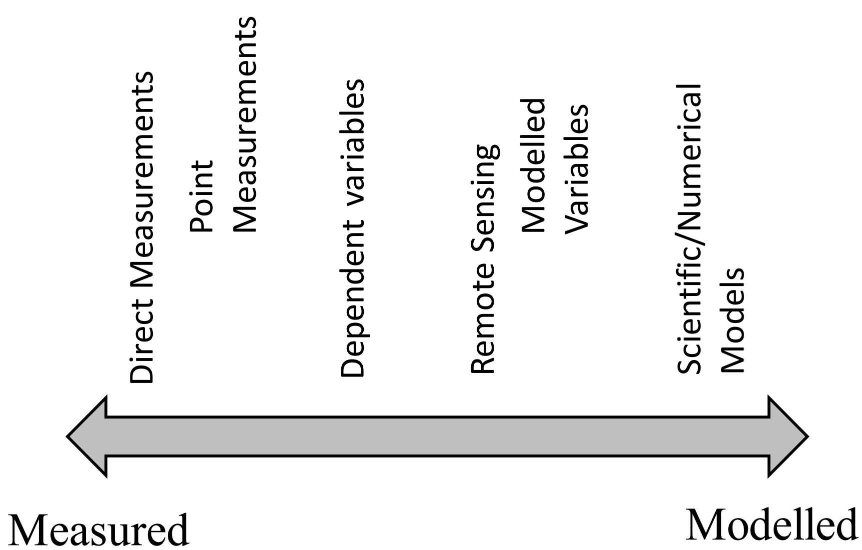
Figure 3. Model chain and indicators. Green circles represent Ecosystem Vitality indicators and blue circles represent Ecosystem Services indicators.



3.3 DATA SOURCES

Data required to calculate the indicators at the sub-basin and basin level is expected to come from a variety of sources combining in-situ measurements, remotely-sensed information and modelled outputs (Figure 4). While direct or in-situ measurement can claim to have the advantage of being the "real" value of variable being measured, such data is both labor intensive to collect and spatially sparse. Remotely-sensed information, on the other hand, may have poorer point accuracy, but have larger and more consistent spatial coverage and ability to identify spatial and temporal patterns. Data from numerical models helps fill in the gaps on 'measured' data availability and provide information on inferred or derived variables, but are dependent on the quality of inputs and the characterization of the complex processes they are attempting to simulate.

Figure 4. Data sources and types. The graph ranges from direct empirical data that is collected on the ground to scientific and numerical models which use direct measurements or variables derived from remote sensing. Direct measurements are data usually in the form of point estimates; however, when numerous such points are collected through space and time (or with respect to some other variable), they can be used to create a distribution. Dependent variables are simple mathematical or statistical functions of direct measurements. Remote sensing data typically needs to be converted via more complex functions or statistical methods into a useful metric. Scientific/numerical models refers to complex models that might use any of the above forms of input.



Each basin is expected to have its own suite of data monitoring sources and models based on a number of local, regional and global factors – such as capability of local authorities, institutional importance for certain physical variables (e.g., water quality), scale of the basin being studied, etc. Table 6 tabulates some examples of local or remotely-sensed data, alongside hydrological, groundwater, hydraulic, water quality, and ecosystem service models (among others) that may be relevant to, or used by, institutions in the basin. Adopting the guidelines documented in Sections 5 and 6 should inform the user of the underlying data required and permit the calculation of the indicators.

Table 6. Local and global data sources, models and metrics for evaluating Ecosystem Vitality and Ecosystem Services indicators

Major indicator	Sub-indicator	Metrics/models	Local and site-scale datasets and models	Global and regional datasets and models
Ecosystem Vitality				

Water Quantity	Deviation from Natural Flow Regime	AAPFD (Gehrke et al., 1995), Hydrologic Deviation (Ladson et al., 1999)	River gauges, hydrological models such as SWAT, HSPF, GSFLOW, etc.	Calibrated instance of Global Hydrologic Models/Land Surface Models such as VIC, WaterGAP, etc.
	Groundwater Storage Depletion	% Area affected	Monitoring wells	GRACE satellite data, land subsidence studies using SAR
Water Quality	Water Quality Index (from TSS, TN, TP and others)	Aggregate of parameter missing WQ targets with frequency and amount with which targets are not met	Local monitoring station, Water quality models such as QUAL, WASP, etc.	MODIS and VIIRS water quality parameters
Drainage Basin Condition	Bank Modification	Percent of bank/shoreline modified	Aerial Photography	LandSAT imagery, SAR (like Sentinel 1) imagery
	Flow connectivity	Dendritic Connectivity Index (Cote et al. 2009)	Aerial Photography; government database on dam and weir locations	GRanD (Global Reservoir and Dam) Database
	Land cover naturalness	Naturalness Index based on land cover, 0-100 scale	Aerial Photography, Local survey for land use	MODIS land cover, Global Forest Change database, ESA CCI land cover products
Biodiversity	Change in richness and population size of species of concern	% Change in number of species and abundance	Local survey	IUCN Red List, national and regional threatened species lists, Global Population Dynamics Database; Global Invasive Species Database
	Change in richness and population size of invasive and nuisance species	% Change in number of species and abundance		
Ecosystem Services				

Provisioning	Water supply reliability relative to demand	Aggregate of sites affected, frequency and amplitude of gap between water supply and demand	Government regulation records, Water supply and demand models such as WEAP	Water availability information from Global Hydrologic Models/Land Surface Models. Demand estimates based on changes in soil moisture, evapotranspiration, etc. (Nazemi and Wheater, 2015)
	Biomass for consumption	Amount of production or area contributing to biomass, frequency and amplitude of gap between biomass supply and demand	Local monitoring data	N/A
Regulation and Support	Sediment regulation	Aggregate of areas affected, frequency and amount of changes in sediment deposition and erosion thresholds	Reservoir operation and regulation records, hydrological models, Ecosystem service models such as InVEST, ARIES	LandSAT or other high resolution imagery, SAR surveys
	Water quality regulation	Aggregate of parameter missing WQ targets with frequency and amount with which targets are not met	Local monitoring stations and authorities	MODIS and VIIRS water quality parameters
	Flood regulation	Aggregate of sites affected, frequency and amplitude of floods compared to demand	Hydrological models and hydraulic models such as HEC-RAS,etc	NRT Global flood mapping, Global flood risk models (Ward et al, 2015)

	Disease regulation	Aggregate of areas affected, incidence ratio and case-to-fatality ratio	Local monitoring and authorities; WADI modelling approach	Resources such as complied by WHO, Global Infectious Disease and Epidemiology Network (GIDEON), generalized global models from Yang et al (2012)
Cultural	Conservation/Cultural Heritage sites	Area (can be weighted by perceived value)	Government regulation records	World Database on Protected Areas
	Recreation	Person-use days or travel costs	Local survey	Geotagged photographs from social media sites

3.4 UNCERTAINTY

3.4.1 Types of uncertainty

Uncertainty in indicator evaluations can manifest in a variety ways, including but not limited to measurement and systematic errors in data, natural variability, missing or incomplete data sets for indicators, simplified representation of complex processes in models used to quantify indicators, and subjective interpretations of language used to define the indicators (e.g. ambiguity, vagueness or lack of specificity; Regan et al. 2002). Perhaps the most serious source of uncertainty in moving forward with an evaluation of the indicators is missing or incomplete data sets. In cases where data are missing for numerous indicators within a component, a decision should be made about whether to proceed with an evaluation. We recommend that if more than 40% of the indicators within one of the three components cannot be evaluated due to lack of data, then aggregation of the indicators into an index for that component not be pursued. However, we encourage the evaluation of indicators for which data is available as this can provide useful information on change of certain aspects of the basin. Lack of data within one of the components (Ecosystem Vitality, Ecosystem Services or Governance and Stakeholders,) should not prevent indicator evaluation and aggregation into indices for the other components, if there is data to do so. Since we do not recommend aggregation across the three components, failure to evaluate and aggregate all the indicators in one component should not preclude evaluation of the others.

Measurement error will likely be the next largest source of uncertainty. It arises from the lack of precise information about the quantities used in the indicator evaluations. This may be due to inaccuracies in estimating values or a lack of knowledge. Measurement error may be reduced or eliminated by acquiring additional data (Regan et al. 2002) and, hence, should be noted wherever possible and practical. Natural variability arises from differences in parameter or indicator values across time, space and other dependent variables (e.g., slope). This will affect indicator values across a basin. For instance, a source of pollution will have greater effects downstream and perhaps no effects upstream. Semantic uncertainty arises from vagueness, ambiguity or lack of specificity in the definition of terms in the indicators or the lack of consistency in different assessors' interpretation of them. Despite attempts to make the definitions of the indicators clear and exact, in some cases this is not possible without the loss of generality needed for broad applicability.

3.4.2 Representing uncertainty

For cases where data is lacking for more than 40% of the indicators within one of the three components, it is recommended that indicator values not be aggregated into an overall index for the relevant component for the basin. When data is available for 60% or more of the indicators within a component (i.e., within Ecosystem Vitality, Ecosystem Services or Governance and Stakeholders), and if no more than one major indicator cannot be evaluated, it is appropriate to aggregate indicators into an overall index using the methods described above. In cases where a major indicator is omitted from the assessment, this must be the result of a consensus decision by the relevant stakeholders and based on a genuine lack of data. Care needs to be taken that when averaging across indicators that only the indicators that can be evaluated are included in the total number of indicators. For instance, if 7 out of 10 indicators can be evaluated, then the aggregated index value should be an average of the 7 assessed indicators without reference to the original 10 in the indicator list for that component. Moreover, if weights are used to denote greater or lesser importance of the indicators for freshwater health in the basin, they should be normalized across the indicators for which data is available, not the full indicator list. In such cases, an uncertainty score should accompany the aggregated index for the relevant component. The uncertainty score is the number of missing indicators from the index over the total number of possible indicators for the component (Alessa et al. 2008).

Natural variability is accommodated, in part, by sub-dividing the basin into sub-basins. Many indicators within components Ecosystem Vitality and Ecosystem Services are then calculated at the sub-basin scale and aggregated to provide an indicator value for the entire basin. In these cases, the same method of aggregating the indicators for the component for the basin should be applied to aggregating the indicator across sub-basins.

Measurement error and natural variability may also be represented by specifying a best estimate and upper and lower bounds on the indicator or on parameters used the calculation of the indicator. The plausible range may be established using various methods – for example, based on confidence intervals, percentiles on distributions of values subject to variability, the opinion of a single expert or the consensus view of a group of experts. The method used should be stated and justified in the assessment documentation. Interval analysis should then be used to calculate the indicator (in cases where parameters in an indicator calculation are represented as intervals) or in the aggregation of indicators represented as intervals (Moore 1966). When the indicator and concomitant indices are represented as intervals, the assessment should be based on the central tendency (i.e., values relying on the best estimates) but reported within the resultant plausible bounds. Representing indicator values as a best estimate within upper and lower bounds will be important in situations where an indicator value lies close to a threshold or objective, e.g., thresholds on acceptable water quality for human consumption. If the best estimate or central tendency for the concentration of a relevant contaminant lies slightly below the threshold of acceptability (by a given standard), this would give an indicator value that would lie on the more sustainable end of the scale. However, if the upper and lower reliable bounds on the concentration span the threshold, this would weaken the evidence for concentrations in compliance of the threshold. It is important to flag such cases and improve assessments by gathering more data to reduce the span of values, and/or initiate management to lower the concentration further.

4. GUIDELINES FOR EVALUATING ECOSYSTEM VITALITY INDICATORS

For each indicator, the following information/guidance is provided:

- **Scale of Calculation**: Identifies the spatial scale at which the calculation of the indicator is applied.
- **Reference:** Publications (where available) that support the methodology behind the indicator calculation
- **Type/Class of Input required:** Lists the types of data necessary to perform the indicator calculation. This list should be carefully considered before starting the calculations and the best available input data should be identified.
- **Suggested source of 'minimum' data:** Specifies potential sources for the minimum data required to calculate the indicator value.
- **Steps for calculation:** Sample calculation process assuming only access to 'minimum' data is available and acts as the process that can be used as default. In cases where better quality local data is available and can be incorporated by changing or modifying the calculation procedure, this should be done.

For each indicator calculated and for different data used for it/versions of calculation, etc., the sheet [**Metadata_[IndicatorName]_[Date]_[version].docx**] is provided to record the metadata.

4.1 WATER QUANTITY (WQT)

4.1.1 Deviation from natural flow regime (DvNF)

Scale of calculation:	Sub-basin, aggregate to basin;
Range of Output:	100 indicates near natural conditions, 0 indicates high deviation; 100-80: Indicates that the seasonal pattern and magnitude of flow/levels resemble a natural flow regime; 80-50: Regulated flow which is likely to maintain seasonal pattern however, magnitudes (especially at peak and ebb points) show marked deviation from natural flow regime; Below 50: Indicates significant deviation in both seasonal pattern and magnitudes from natural flow regime.
Reference:	Ladson et al (1999), Gehrke et al. (1995), Gippel et al (2011)
Type/Class of Input required:	Monthly flow data under present day and natural conditions for the same period
Suggested source of 'minimum' data to enable calculation:	Modeled monthly flow data for 5 years with/without all modifications to basin

Steps for calculation of indicator:

1. *Option 1- Amended Annual Proportion of Flow Deviation (Gehrke et al. 1995, Gippel et al 2011):*

$$AAPFD = \sum_{j=1}^{p} \frac{\sqrt[2]{\sum_{i=1}^{12} \left[\frac{m_i - n_i}{\bar{n}_i}\right]^2}}{p}$$

   Where, m_i is monthly flow data accruing to current condition, n_i is modeled natural flow for the same period. p is the number of years and $\bar{n}_i$ is mean reference flow for month *i* across *p* years (Note: in ephemeral streams, this should be changed to incorporate annual average flow to avoid extremely large values). Values are normalized as follows using thresholds reported in Gehrke et al. 1995 and Gippel et al 2011:

$$DvNF = \begin{cases} 100 - 100 \times AAPFD \text{ for } 0 \le AAPFD < 0.3 \\ 85 - 50 \times AAPFD \text{ for } 0.3 \le AAPFD < 0.5 \\ 80 - 20 \times AAPFD \text{ for } 0.5 \le AAPFD < 2 \\ 50 - 10 \times AAPFD \text{ for } 2 \le AAPFD < 5 \\ 0 \text{ for } AAPFD \ge 5 \end{cases}$$

   In case of lakes, monthly flow data can be replaced with 'level' data (See Liang et al. 2015 as an example).

2. *Option 2- Hydrologic Deviation (Ladson et al, 1999):*

$$HD = \frac{\sum_{i=1}^{12} |m_i - n_i|}{\sum_{i=1}^{12} n_i}$$

   Where, m_i is monthly flow data accruing to current condition and n_i is modeled natural flow for the same period.

   Based on the studies on regulated streams in Australia, Ladson et al. (1999) assigned ratings within lower and upper bounds of 20% and 65% respectively. We use these to normalize hydrologic deviation as follows:

$$DvNF = \begin{cases} 0 \text{ for } HD \ge 0.65 \\ 100 - \frac{100}{0.45}(HD - 0.2) \text{ for } 0.20 < HD < 0.65 \\ 100 \text{ for } HD \le 0.20 \end{cases}$$

4.1.2 Groundwater storage depletion (GwSD)

Scale of calculation:	Sub-basin or basin
Range of Output:	100 indicates no groundwater storage depletion; 0 signs of widespread depletion
Reference:	Vrba and Lipponen (2007)

Type/Class of Input required:	(1) Groundwater(GW) heads data (2) Groundwater extraction information
Suggested source of 'minimum' data to enable calculation:	Identify/outline areas of potential GW depletion using information from GRACE, Land Subsidence studies, identifiable areas with heavy GW exploitation.

Steps for calculation of indicator:

1. *Identify areas with a potential GW depletion problem:*
   Areas with potential depletion problems can be identified by the following methods (following Vrba and Lipponen (2007)):
   - *Areas with a high density of production wells:* Groundwater level declines are strongly associated with an increase of pumping costs or loss of spring or production well yields, which can indicate groundwater depletion in areas where many wells are exploiting an aquifer. Two alternatives for identifying water level declines are: 1) to detect a consistent and gradual downward trend of water level from a well monitoring network (when available) or 2) to compare the groundwater level at wells drilled at different times (i.e., compare water level evolution using nearby wells, but drilled at different times: 1960s, 1970s, etc.). For the latter, it is fundamental to have a well inventory that can provide information about the well construction and hydraulic parameters of the aquifer. (In basins large enough, GRACE data may help establish depletion.)
   - *Change of base flow:* In many areas, rivers and other surface water bodies receive an important proportion of their water from groundwater base flow. Drastic reduction of this groundwater flux and loss of base flow can be associated with groundwater depletion. In this case, the monitoring of river flow is important. An indirect indication of reduction of base flow can be established when phreatic vegetation or wetlands suffer notable changes.
   - *Change of groundwater quality characteristics:* Although the physical-chemical properties of water can vary throughout the aquifer, in conditions of regular exploitation, drastic changes in groundwater quality are not expected (including stable isotope composition). Therefore, changes in age and origin of groundwater at specific locations in the aquifer can be an indication of groundwater depletion.
   - *Land subsidence:* At some localities, groundwater exploitation from thick sedimentary aquifer-aquitard systems has been accompanied by significant land subsidence. In this case, land subsidence can be used as an indirect indicator of unsustainable groundwater exploitation.

   As Vrba and Lipponen (2007) notes, care must be taken when evaluating GW depletion because it is also subject to natural and seasonal fluctuations from the influence of climatic conditions and aquifer characteristics. Sometimes, groundwater storage depletion may also be associated with a long transient evolution from one steady state to another and does not necessarily represent a problem of unsustainable exploitation of an aquifer. The most difficult problem in aquifers that are subjected to exploitation is distinguishing permanent and regional depletions from only temporal and local interferences caused by the proximity of production wells.

2. *Calculate GwSD as:*

$$GwSD = \left(1 - \frac{\sum a}{A}\right) * 100$$

Where *a* is the area with depletion problems identified, and *A* is the sub-basin/basin area being studied.

4.2 WATER QUALITY (WQL)

Scale of calculation:	Basin/Sub-basin
Range of Output:	100-95 indicates excellent water quality; 80-94 indicates good water quality; 79-65 indicates fair; 64-45; <45 indicates poor water quality
Reference:	Canadian Water Quality Index (CCME 2001)
Type/Class of Input required:	Total Suspended Solids (TSS), Total Nitrogen (TN), Total Phosphorus (TP) time series and concentrations of other pollutants of interest
Suggested source of 'minimum' data to enable calculation:	Data requires local input from observation or models for minimum 4 pollutants with at least 4 data points each

Steps for calculation of indicator:

1. *Measure/Model to obtain estimate:*
   Values calculated for the 3 "essential parameters" (SS measured as Turbidity, TP and TN) and an additional local parameter of concern. Each parameter considered needs an objective to satisfy in the form of a threshold or range that the measured value should be below or within. While the threshold may be assigned according to the local context (e.g., effects on sensitive species or ecosystem processes) by the user, initial values may be populated based UNEP (UNEP 2007; UNEP 2008b) recommendations such as:

Parameter	**Recommended Range**
Suspended Solids (measured as Turbidity)	<5 NTU
Total Nitrogen	< 2 mg/L – 6mg/L
Total Phosphorus	< 10 µg/L – 40 µg/L

The ranges reported here reflect differences across ecosystem types, with the intent that users apply a threshold within this range. Values less than this selected threshold are then considered to satisfy the objective. Alternatively, the data-derived ecosystem-based thresholds for total nitrogen and total phosphorous for an Australian case study in the table below can be used as defaults in the same way (Hart et al. 1999). However, we refer users to Hart et al. (1999), CCME (2002),UNEP (2007) and UNEP (2008b) to ascertain whether the default thresholds presented in either table are appropriate for the given context. Thresholds should be chosen because they are the most appropriate for the context and not because they will provide a favorable indicator value.

Ecosystem type	TN (mg/L)	TP (µg/L)
Lowland River	1.60	37
Upland River	0.34	35
Freshwater Lakes and Reservoirs	0.44	50
Estuaries	0.08	45
Coastal and Marine	0.35	55
Wetlands	No data	No data

2. *Scale and aggregate:*
   Use Canadian Water Quality Index method. See WQL Calculator for details.

4.3 DRAINAGE BASIN CONDITION (DBC)

4.3.1 Flow Connectivity

The Combined Dendritic Connectivity Index (cDCI) measures the longitudinal connectivity of the river network for potamodromous and/or diadromous fish species.

Combined Dendritic Connectivity Index (cDCI)

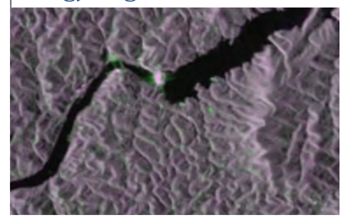
Scale of calculation:	Single value per basin
Range of Output:	100 indicates free flowing river and 0 highly fragmented river. See Figure 6 for graphical depiction of values for a hypothetical river
Reference:	Based on Cote et al. (2009)
Type/Class of Input required:	(1) GIS layer of river network (2) Location of barriers/structures across the river (dams, etc.) (3) Measure of "passability" of each structure in both upstream and downstream directions for fish (4) Information on whether index is being calculated for impact on potamodromous (migrations within freshwater) and/or diadromous (migrations between marine and freshwater) fish species
Suggested source of 'minimum' data to enable calculation:	(1) HydroBASIN river network with manual correction at outflow (2) SAR data for manually locating structures and/or Global Reservoir and Dams database (GRanD)

Steps for calculation of indicator:

1. *Identify and geo-locate barriers fragmenting the river network:*
   This can be compiled from local agencies and/or global databases (if local information is unavailable). The image on the right shows SENTINEL-1 SAR data, which can be used to manually identify "obstructions" that show up as bright patches across water relative to the river network (dark).

2. *Assign "passability" value for each structure:*
   This will be based on local input. Cote et al. (2009) assign barriers an associated passability value, *p*, which ranges from 0 to 1. This value depends on the physical (e.g., dam height), chemical and/or the hydrologic (flow rates, which vary temporally) attributes of the barrier as well as the biology of the organism in question (which can vary by species, age, etc.).
   Note: In the absence of data, following Clarkin et al. (2005), we assign each barrier a binary passability value. That is, either a barrier meets the designated fish passability criteria (*p*=1) or not (*p*=0). We start with *p*=0 for all structures and allow the user to change this to *p*=1. At later stages, functionality can be added for intermediate values if found useful.

Figure 5. SENTINEL 1 SAR image of a dam on the Dongjiang.



3. *Impact on potamodromous fish species:*
   For (*n*-1) structures with *p*=0, dividing the river into *n* fragments, DCIp is calculated as:

$$DCIp = \sum_{i=1}^{n} \frac{l_i^2}{L^2}$$

   where, *L* is the total length of the river, and l_i is the length of i^{th} fragment

4. *Impact on diadromous fish species:*
   For (*n*-1) structures with *p*=0, dividing the river into n fragments, DCId is calculated as:

$$DCId = \frac{l_i}{L}$$

   where, *L* is the total length of the river, and l_i is the length of fragment closest to the mouth of the river system.

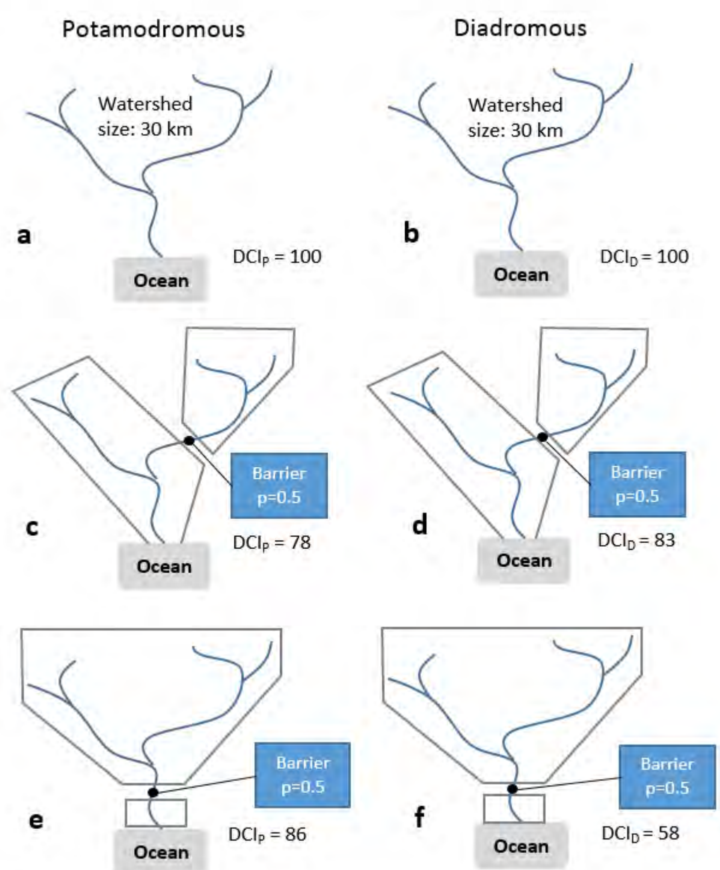
5. *Combine:*
   Finally, cDCI can be calculated as:

$$cDCI = \left(\frac{w_p DCIp + w_d DCId}{w_p + w_d} \right) * 100$$

   where, weights w_p and w_d depend on the nature of the fish species in the freshwater system.

   Notes: 1) Suggested weightings for (w_p, w_d) are (1,0) for potamodromous dominant systems; (0,1) for diadromous dominant systems.
   2) For large sub-basins, like the transboundary 3S system (the Sesan, Sre Pok, and Sekong rivers spanning Cambodia, Lao ODR and Vietnam), which contain potamodromous species, the fish species travelling up from the Mekong main stem will be affected by obstacles to connectivity in same fashion as diadromous species. Hence, calculation of DCId will be appropriate.

Figure 6. Calculation of DCIp and DCId for a hypothetical river network with one barrier having p in both directions as 0.5. (Source: Cote et al. 2009)



4.3.2 Bank Modification

Bank Modification measures lateral connectivity; as a proxy, we focus on the percentage of river channels (could also include lakeshores) affected by human modification, such as channelization or shoreline hardening.

Percent Channel affected by Modification (pCM)

Scale of calculation:	Sub-basin; aggregate to single value per basin
Range of Output:	100 indicates no modification and 0 highly is modified
Reference:	Nil
Type/Class of Input required:	(1) GIS layer of river network (2) Location of structures along the river including dykes, levees, channelization, etc.
Suggested source of 'minimum' data to enable calculation:	(1) HydroSHEDS/HydroBASINS river network with manual correction at outflow (2) LandSAT imagery when appropriate

(3) Aqua Monitor: http://aqua-monitor.appspot.com/ may be useful

Steps for calculation of indicator:

1. *Assign/Calculate percentage length modified for each sub-basin:*
   If local data on location of levees, dykes, channelization, clearance of instream obstructions to navigation, reservoir extent etc., is available, then for each sub-basin, the percentage length affected can be calculated (0 for near-natural, 1 for fully channelized). Otherwise, the following decision matrix maybe used based on visual inspection of LandSAT imagery for year being audited:

Visual inspection of imagery for channel in sub-basin	**Score (pCM_i)**
Almost no modification visible with near natural riparian corridor	0
Some modification is visible including farms/urban areas reaching river edge. But length affected is clearly less than half of the channel length	0.25
Modifications are visible, nearly half the channel in sub-basin is affected	0.5
Large sections of the channel are clearly modified but some patches of natural sections remain	0.75
Most of the channel in sub-basin is modified; natural riparian corridor is entirely absent and all river banks have been affected by human use	1

   Sub-basins that are predominantly lakes/reservoirs should be excluded from this calculation.

2. *Aggregate based on channel length:*
   Scores within a sub-basin can be calculated as:

$$pCM = \left(1 - \frac{\sum_{i=1}^{n} l_i pCM_i}{L}\right) * 100$$

   where, *L* is the river network length, l_i is the length of the *i*th river fragment, and pCM_i is the appropriate score from the table directly above for the *i*th river fragment.

4.3.3 Land cover naturalness (LCN)

The Naturalness Index describes the state and trend of land use/land cover (LULC) within the basin, according to the amount of human-induced transformation present. A basin in its undisturbed state, with intact forests and wetlands, generally maintains a sufficient quantity and quality of water to support indigenous flora and fauna. Naturalness exists on a gradient from completely natural to completely artificial or human dominated (Angermeier, 2000). Human conversion of lands and waterways are associated with increases in pollutant loads (non-point source from agriculture, point-source from urban and industrial), changes to infiltration and runoff regimes, and losses of regulating services (flood mitigation, erosion prevention, water purification). The Naturalness Index is, therefore, a proxy indicator for the degree to which these naturally-occurring functions are preserved within the basin. It is based on similar efforts to categorize and quantify this gradient over landscapes, such as Machado's (2004) index of naturalness and the hemeroby hvb scale (Sukopp, 2004; Paracchini and Capitani, 2011; Walz and Stein 2014). More detailed investigations of specific LULC types, such as forest conversion to agriculture, may be warranted as a secondary step and can be calculated using the same data.

Scale of calculation:	Sub-basin; aggregate to single value per basin
Reference of indicator (if any):	Based on index of naturalness methods described in Machado (2004)
Input required:	(1) Land cover data (raster or shapefile) for at least 2 time periods to derive change (e.g., 2010 and 2015) (2) Degree of Naturalness classification table (3) Sub-basin shapefile
Suggested sources of 'minimum' data to enable calculation:	(1) ESA CCI land cover (2000, 2005, 2010, 300m resolution); Normalized Difference Vegetation Index (NDVI)

Steps for calculation of indicator:

1. *Review and revise naturalness weights:*
   The Degree of Naturalness classification table (Deg_of_N.csv) contains descriptions of LULC types as well as cultural practices (e.g., irrigation) that correspond to "naturalness" weights on a 0-100 gradient. Sub-classifications are suggested based on three factors:

   - **Management of the water cycle**: manually altering the flow and/or use of water to maintain a particular land-use type
   - **Pollution**: chemical and physical pollutants entering the local water cycle due to cultural practices, such as fertilizer and pesticide use, and increased soil runoff from croplands as well as urban runoff and point-source wastewater loads from urban and industrial lands
   - **Vegetation characteristics**: degree of native vegetation and permanence of vegetative cover

   The proposed weighting includes ranges of values to help highlight transitions from "natural" to "transformed" systems, i.e., from forests and wetlands to cultivated lands or from cultivated lands to urban areas (see Table 7). It is recommended that the default weights in the classification table be reviewed and, based on expert judgment, adjusted to be compatible with local conditions. For example, in some regions, flooded rice paddies may be considered to have a higher degree of naturalness than other irrigated crops, due to their ability to mimic some aspects of wetlands (which they may have replaced). In this case, a different classification and higher relative weight may be appropriate. Similarly, local or region-specific land use datasets may include highly detailed and differentiated classes of land use that will require expert judgment on their relative weight.

Table 7. Proposed "naturalness" characteristics and weights

Degree of naturalness	Management of water cycle	Pollution emissions	Vegetation characteristics	Examples	Weight
Natural and semi-natural	None	None	Native	Forest (primary and secondary); lakes (natural) and wetlands; native grasslands; native shrublands	100
Cultural assisted system	Low	Low	Mixed, high diversity	Mosaic native vegetation (>50%, vegetation cover <50%)	70
	Low	Low	Mixed, moderate diversity	Mosaic cropland (>50%, natural vegetation <50%)	60
Transformed system	Low	Low	Permanent cover with atypical species	Permanent pasture land; agroforestry; tree crops	50
	Low to Moderate	Moderate to High	Seasonal cover with atypical species	Non-irrigated arable land	40
	High	Moderate to High	Seasonal cover with atypical species	Permanently irrigated arable land	30
Completely artificial	High	Moderate to High	Sparse cover with grass	Urban park space; low-density suburban areas; barren land	10
	High	High	None	Urban commercial areas; mining areas	0

2. *Associate land cover codes from input file with categories in the* Degree of Naturalness *classification table*:
   If your LULC input file does not already have numeric codes associated with each cover type, first assign a unique numeric identifier. Then, enter the numeric identifier for each LULC type into the classification table (Deg_of_N.csv) where it is a best fit. Copy rows in the classification table to accommodate multiple LULC types from the input file that should have the same categorization and weight. **NOTE**: It is preferable to have your LULC input file as raster data. If it is in vector data, convert the polygons to a raster data set and specify a reasonable cell size based on your input data, e.g., 30m or a cell size that is consistent with the DEM you have available for the basin. Reclassify grid values for your LULC input file according to the corresponding naturalness weight from the classification table. NOTE: The basin's overall "naturalness index" score will be the mean value of the reclassified raster.

3. *Calculate mean naturalness values at sub-basin scale:*
   Using the recommended sub-basin delineation, calculate an average value over each sub-basin and save as a new file. Use zonal statistics with the sub-basin file as the zones. Calculate the mean and standard deviation using the Deg_of_N raster values as input. Join the resulting output table to the sub-basin vector file to produce a map of these values.

4. *Calculate changes between two time periods:*
   Repeat steps 2 and 3 for LULC input file from an earlier time period (e.g., 5 years ago) to compare with current land use. Join the two zonal statistics tables, then create a new field and subtract the mean Deg_of_N values in the earlier time period from the current time period. Join this new attribute to the sub-basin vector file to produce a map of the change in naturalness scores. NOTE: So some values will be negative.

4.4 BIODIVERSITY (BIO)

Biodiversity assesses potential shifts in ecosystem functioning by measuring changes in the biota that constitute an integral component of freshwater ecosystems. The status and trends of biodiversity in a basin signify ecosystem health, with declining populations of native species and increasing populations of invasive and nuisance species indicating a deteriorating conditions or degradation of an ecosystem. This biodiversity indicator is comprised of the number (expressed as both species richness and abundance) and changes in:

> *Species of concern* consisting of threatened aquatic or riparian (water-dependent) species and other species of interest (such as keystone or umbrella species) that will be affected by changes in habitat condition. Both presence or absence of particular species and their population trends over time are of interest here.
>
> *Invasive and nuisance species* in lakes, waterways and riparia indicate anthropogenic alteration of ecological conditions, as these are the circumstances that allow alien species to thrive at the expense of native species. Numbers of species present and their population trends are of interest.

Scale of calculation:	Sub-basin; aggregate to single value per basin
Range of Output:	100 – 0 where 100 indicates higher biodiversity and 0 indicates lowest biodiversity
Reference:	Living Planet Index (Loh et al. 2005)
Type/Class of Input required:	
Suggested source of ‘minimum’ data to enable calculation:	Local surveys Useful data bases for calculation of the *Biodiversity* indicator http://www.iucnredlist.org/ https://www.iucn.org/theme/species/our-work-ssc/our-work/freshwater-biodiversity

	http://data.freshwaterbiodiversity.eu/ http://www.livingplanetindex.org/data_portal http://www3.imperial.ac.uk/cpb/databases/gpdd http://www.compadre-db.org/ http://www.natureserve.org/ http://www.iucngisd.org/gisd/ https://www.invasivespeciesinfo.gov/aquatics/databases.shtml

Steps for calculation of indicator:
Calculate 'Species of Concern (ISC_i)' and 'Invasive and nuisance species (INS_i)' sub-indicators in year *i* using the process outlined below.

4.4.1 Changes in presence (i.e. species number) and population size of species of concern

Species of concern should consist of native at-risk freshwater species (including but not limited to aquatic invertebrates, amphibians, fish and water birds) listed on the IUCN Red List as critically endangered (CR), endangered (EN), or vulnerable (VU) (IUCN 2012) and nationally and/or provincially listed threatened and endangered freshwater species occurring in the basin. Species of concern should also include carefully selected freshwater species whose status and population trends are linked to the health of the freshwater ecosystem, such that a change in freshwater ecosystem health would result in a change in the status or population trends of the species over time. These species might include umbrella, keystone, flagship or indicator species that might not be under threat but would be sensitive to changes in the freshwater system (Caro 2010). Local ecologists will be the best source for information on the identities and status and trends of such species. Monitoring data on population sizes or other measures of abundance, such as biomass or density, should be collected periodically for such additional species to be considered species of concern. Since this indicator evaluates *changes* in species of concern, species should be carefully selected at the outset – and there should be sufficient of them — so that a change in the number of species of concern adequately reflects genuine changes as the result of threats or beneficial management actions in the basin.

Steps for calculation of indicator:
The index for *species of concern* is calculated in four parts: calculation of the proportion of threatened and endangered freshwater species out of the total freshwater species assessed in the basin, calculation of the change in the number (i.e., richness) of species of concern, and the average population trend across all species of concern for which there is data. These three parameters are then combined to give an overall index for the status and change in *species of concern*.

1. The first step in the calculation estimates the proportion of threatened and endangered freshwater species, of the total freshwater species assessed in the basin. For IUCN Red List species, total assessed species are comprised of all species that have undergone IUCN Red List assessments, excluding those that have been deemed Data Deficient (DD). If information is available, nationally and/or provincially listed threatened and endangered freshwater species should be included, as well as the total number of freshwater species assessed at the national and provincial level, taking care to ensure that species are not represented more than once. Using only the number of listed threatened and endangered species (excluding additional species of concern), the proportion of threatened and endangered species, of the total species assessed, is

calculated as:

$$I_{TE,i} = 1 - \frac{w_{CR}n_{CR,i} + w_{EN}n_{EN,i} + w_{VU}n_{VU,i} + \sum_j w_j n_{j,i}}{(w_{CR}n_{CR,i} + w_{EN}n_{EN,i} + w_{VU}n_{VU,i} + \sum_j w_j n_{j,i} \ + w_{NotT}n_{NotT})}$$

where $n_{CR,i}$, $n_{EN,i}$, and $n_{VU,i}$ are the number of species listed as CR, EN, or VU under the IUCN Red List categories and criteria at time *t = i,* respectively, $n_{j,i}$ is the number of species classified in an endangered or threatened category at the national or provincial level at time *i* (e.g., for regions that classify species as "endangered" or "threatened", *j*=1 refers to the endangered category and *j*=2 refers to the threatened category), n_{NotT} refers to the remaining assessed species that are not classified in a threatened category (e.g. Least Concern [LC], or Near Threatened [NT] in the IUCN Red List), w_{CR}, w_{EN}, w_{VU}, and w_{NotT} are weights applied to the number of CR, EN, VU and not threatened species, respectively, w_j are the weights applied to the number of endangered and threatened species at the national or provincial level. The sum of all $n_{x,y}$ is the total number of species assessed in the basin under the IUCN Red List criteria and/or national or provincial criteria. Weights should be assigned such that $w_{CR} \geq w_{EN} \geq w_{VU} \geq w_{NotT}$ and $w_j \geq w_{j+1} \geq w_{NotT}$. Default values for the IUCN Red List species are $w_{CR} = 3.0$, $w_{EN} = 2.0$, and $w_{VU} = 1.0$ and $w_{NotT} = 0.5$. Default values for the nationally and/or provincially listed species will depend on the number of categories of threat; for two threat categories, "endangered" and "threatened," $w_1 = 3.0$ and $w_2 = 2.0$ are recommended.

2. A change in the number of species of concern is calculated as:

$$\Delta SC_i = \frac{SC_{i-1}}{SC_i}$$

where ΔSC_i denotes the change in the number of species of concern from time $t = i - 1$ to time $t = i$, SC_{i-1} is the number of species of concern at time $t = i - 1$ and SC_i is the number of species of concern at time $t = i$. Note that species of concern here refers to both the threatened and endangered species as well as carefully selected umbrella, keystone, flagship or indicator species that might not be under threat but would be sensitive to changes in the freshwater system. $SC_i = 0$ under two very different circumstances: 1) all of the species of concern have become extirpated in which case $\Delta SC_i = 0$ or 2) all of the species of concern have improved in their status and trends as to be deemed no longer of concern, in which case $\Delta SC_i = 1$. It is strongly recommended that species of concern are selected such that the second circumstance is not achieved, e.g., while species may improve from threatened and endangered to recovered and be removed as species of concern, species selected because they are sensitive to changes in freshwater ecosystems (as opposed to their threat status) would continue to serve their role as indicators until such a time as they are extirpated entirely.

3. For as many of these species as data are available, population trends for each of them are calculated for the relevant time period as:

$$\Delta N_{i,j} = ln\left(\frac{N_{i,j}}{N_{i-1,j}}\right)$$

where $\Delta N_{i,j}$ is the change in population size or abundance measure from time $t = i - 1$ to time $t = i$, *j* is the species, $N_{i-1,j}$ is the population size of species *j* at time $t = i - 1$ and $N_{i,j}$ is the population size of species *j* at time $t = i$.

Following the methods of the Living Planet Index (Loh et al. 2005), the average of the population trends across all species for which data are available is calculated as:

$$\overline{\Delta N_i} = \frac{1}{n_i}\sum_{j=1}^{n_i} \Delta N_{i,j}$$

where $\overline{\Delta N_i}$ is the average of population size or abundance changes from time $t = i - 1$ to time $t = i$, n_i is the number of species for which there is population/abundance trend data across the time period, and *j* is the species index. The composite population trend value across all species is calculated as:

$$PT_i = exp(\overline{\Delta N_i})$$

where PT_i is the population trend value across all species from time $t = i - 1$ to time $t = i$. In the absence of population trend data, the default value for PT_i is 1.0.

4. The combination of the proportion of assessed species in the basin that are endangered and threatened, the change in the number of species of concern and the population/abundance trends of those species for which population data are available, i.e., the value of the *species of concern* indicator, is calculated as:

$$ISC_i = min\{ISC_{i-1}(I_{TE,i} \times \Delta SC_i \times PT_i, 100\}.$$

For the very first assessment of the basin at time $= 1$, $ISC_0 = 100$. For cases where no information is available on population/abundance trends, $PT_i = 1$.

4.4.2 Changes in presence and population size of invasive and nuisance species

Invasive and nuisance species are nonindigenous or alien species that “threaten the diversity or abundance of native species, the ecological stability of infested waters, or commercial, agricultural, aqua cultural or recreational activities dependent on fresh waters” (Nonindigenous Aquatic Nuisance Prevention and Control Act of 1990). All known invasive and nuisance species that occur and pose a threat to freshwater health in the basin should be identified. The Global Species Data Base and any other available databases for the region can be used to identify such species. However, local expert knowledge should be used to verify the presence and species identity of invaders in the assessment basin. Monitoring data on population sizes or other measures of abundance, such as biomass, density or area or extent of cover, should be gathered for as many of these species as possible. A change in the indicator value should reflects either a greater intensity of threats posed by the increases in the invader or a lessening of such threats due to beneficial management actions that reduce populations of the invader within the basin.

Steps for calculation of indicator:

The index for *invasive and nuisance species* is calculated in three parts: an index denoting the number (i.e., richness) of invasive and nuisance species, a calculation of the change in the number of invasive and nuisance species, and the average population trend across all invasive and nuisance species for which there is data. These three parameters are then combined to give an overall index for the status and change in *invasive and nuisance species*.

1. An index for the number of invasive and nuisance species is calculated as:

$$I_{IN,i} = \begin{cases} 1 - \frac{n_{IN,i}}{10}, & \text{for } 0 \leq n_{IN,i} \leq 8 \\ 0.1, \text{ for } n_{IN,i} \geq 9 \end{cases}$$

   where $n_{IN,i}$ is the number of invasive and nuisance species in the basin at time *t* = *i*.

2. A change in the number of invasive and nuisance species is calculated as:

$$\Delta n_{IN,i} = \frac{n_{IN,i-1}}{n_{IN,i}}$$

   where $\Delta n_{IN,i}$ denotes the change in the number of invasive species from time $t = i - 1$ to time $t = i$, $n_{IN,i-1}$ is the number of species of concern at time $t = i - 1$ and $n_{IN,i}$ is the number of species of concern at time $t = i$. In cases where $n_{IN,i} = 0$ then $\Delta n_{IN,i} = 1$ as this denotes the case of a decline to 0 invasive or nuisance species.

3. For as many of these species as data are available, population/abundance trends for each are calculated for the relevant time period as:

$$\Delta IN_{i,j} = ln\left(\frac{IN_{i-1,j}}{IN_{i,j}}\right)$$

   where $\Delta IN_{i,j}$ is the change in population size (or biomass, density, area or extent of cover) from time $t = i - 1$ to time $t = i$, *j* is the species, $IN_{i-1,j}$ is the population size (or biomass, density, area or extent of cover) of species *j* at time $t = i - 1$ and $IN_{i,j}$ is the population size of species *j* at time $t = i$.

   Following the methods described above, the average of the population trends across all species for which data are available is then calculated as:

$$\overline{\Delta IN_i} = \frac{1}{n_i}\sum_{j=1}^{n_i} \Delta IN_{i,j}$$

   where $\overline{\Delta IN_i}$ is the average of population size (or biomass or density) changes from time $t = i - 1$ to time $t = i$, n_i is the number of species for which there are population trend data across the time period, and *j* is the species index. The composite population trend value across all species is calculated as:

$$IPT_i = exp(\overline{\Delta IN_i})$$

where IPT_i is the population trend value across the invasive and nuisance species from time $t = i - 1$ to time $t = i$.

4. The combination of change in the number (i.e. richness) of invasive and nuisance species and the population trends of those species for which measures of abundance are available, i.e., the value of the *invasive and nuisance species* indicator, is calculated as follows:

$$INS_i = min\{INS_{i-1}(I_{IN,i} \times \Delta n_{IN,i} \times IPT_i, 100\}.$$

For the very first assessment of the basin at time $= 1,\ INS_0 = 100$. For cases where no information is available on population/abundance trends, $IPT_i = 1$.

5. GUIDELINES FOR EVALUATING ECOSYSTEM SERVICES INDICATORS

We have developed a common framework, analogous to elements of risk assessment frameworks (Covello and Merkhoher, 2013), to derive a systematic process for evaluating ecosystem services (ES) that attempts to describe and quantify the ability of an ecosystem to deliver the services that people demand or expect of it. Under this framework, at least two aspects of provisioning and regulating ES need to be considered: 1) the possibility of demand for the ecosystem service not being met and (2) variability of the occurrence, timing or magnitude of events that leads to unmet demand. The assessment is carried out by dividing the area of interest (river basins for water-related ES) into spatial units in which the delivery of ES can be evaluated, and thus, the objective for non-compliance of demand from the ES can be set. For certain ecosystem services, a univariate or 'crisp' threshold-based objective can be defined that will be directly quantifiable, while others will have multi-variable or 'fuzzy' threshold-based objectives that may require indirect estimates. For example, for the provision of water to various sectors and cities from a basin, the volume demanded is directly quantifiable and hence, non-compliance can be evaluated based on whether the demand is met or not. On the other hand, when considering the damages from flood events to the inhabited areas within a basin, it is harder to assign a threshold to any reduction or lack of capacity of a freshwater ecosystem to regulate floods and evaluation may depend on the evaluation methods used.

To calculate the ecosystem service indicator (ESI) using the spatial units and objectives or thresholds, three dimensions are evaluated: scope (F1), frequency (F2) and amplitude (or excursion) (F3). These dimensions are similar to those used in the CCME Water quality index (Saffran et al. 2001) and mirror the aspects of 'risk source,' 'exposure' and 'consequences' used in many risk calculations (Merkhofer, 2012; Covello and Merkhoher, 2013). These three dimensions are defined as:

- Scope (F1): The number of spatial units in area of interest that are unable to meet the objective or threshold.
- Frequency (F2): The frequency with which the objectives or thresholds are not met.
- Amplitude (or excursion) (F3): The amplitude or magnitude under which the objectives or thresholds are not met.

The final value for each dimension is scaled between (0-100) before combining into a final score. Data quality and availability to determine the three dimensions will vary based on the ecosystem service being evaluated and the area of assessment: In some cases, only 1-2 of the dimensions can be calculated with any confidence from the available data. In some cases only one or two dimensions can be calculated with certainty (Modarres, 2006) in which case the robustness or certainty of the evidence should be reported when calculating the final scores, as follows:

1. If able to only determine F1: $ESI = 100 - F1$ (low evidence)
2. If able to only determine F1 and F2: $ESI = 100 - \sqrt{F1 \times F2}$ (medium evidence)
3. If able to determine all three: $ESI = 100 - \sqrt{F1 \times F3}$ (high evidence)

5.1 INDICATOR SET-UP AND DATA REQUIREMENTS

The tables below define the set-up for the indicators for Provisioning (PRO) and Regulation and Support (REG) under Ecosystem Services. The tables define the granularity of the spatial unit (SU) to be used for the calculation, the type of data that maybe used in calculation process as well as the minimum data that should be used for process. Finally, in line with the theoretical framework (Section 1), the possible definitions for the objective (for determining threshold), scope (F1), frequency (F2) and amplitude (F3) are provided. After considering the data available and definitions below, the user should follow the steps articulated in the sections below to calculate the indicators.

5.2 PROVISIONING (PRO) SERVICES

5.2.1 Water supply reliability relative to demand (WaSD)

Spatial Unit:	Location/demand sites and sectors; aggregate to basin
Type/Class of Input required:	1. Summary of monthly demand from various sectors and the supply actually provided. 2. Environmental flow requirements and actual discharge at monitored points in the rivers.
Suggested source of 'minimum' data to enable calculation:	Local (monthly or seasonal) data for 1-5 years
Objective:	Water demand is met
F1:	Where demand is not met
F2:	Frequency with which monthly demand is not met
F3:	Amplitude/Excursion: The difference between supply and demand when demand is not met

5.2.2 Biomass for consumption (BiCN)

Spatial Unit:	Fishing lots/sub-basins; aggregate to basin
Type/Class of Input required:	Estimates of biomass used/acquired for consumption; can be in the form of catch or production units as available.
Suggested source of 'minimum' data to enable calculation:	Data will be site specific and availability will vary considerably
Objective:	Based on loss of productivity
F1:	Sub-basins where catch/productivity has dropped
F2:	How frequently is the reported catch/productivity below expected levels?
F3:	Amplitude/Excursion based on: the magnitude of loss of catch/productivity

5.3 REGULATION AND SUPPORT SERVICES

5.3.1 Sediment regulation (SeRG)

Spatial Unit:	Reservoirs, deltas, flood plains and/or river reaches; aggregate to basin
Type/Class of Input required:	1. Current reservoir sedimentation rate or loss of capacity rate and design threshold for sediment deposition in the reservoir. 2. River bank erosion 3. Rate of deposition on floodplain and threshold, if any 4. Area and/or rate of deposition/erosion from delta; and expected or average rate based on historical records
Suggested source of 'minimum' data to enable calculation:	Local sources for past 1-5 years
Objective:	Expected deposition or erosion of floodplain/delta based on design threshold
F1:	Number of locations where threshold is exceeded
F2:	Frequency with deposition/erosion incidence exceed threshold (annually)
F3:	Amplitude/Excursion: difference between actual rate and threshold

5.3.2 Deviation of water quality metrics from benchmarks (DvWQ)

Spatial Unit:	River reaches/sub-basin; aggregate to basin
Type/Class of Input required:	Target 'class' WQ targets for the river reach considered and the actual 'class' or actual WQ quality of that reach
Suggested source of 'minimum' data to enable calculation:	Monthly modelled or recorded WQ for 1-5 years and WQ targets of each reach
Objective:	WQ target is met
F1:	Number of locations where WQ target is not met
F2:	Frequency with which WQ targets are not met
F3:	Amplitude/Excursion: gap between WQ targets and actual values

5.3.3 Flood regulation (FlRG)

Spatial Unit:	City or sub-basin; aggregate to basin
Type/Class of Input required:	Flood outlines, interception of flow in upstream areas
Suggested source of 'minimum' data to enable calculation:	Floods frequency/statistics over the past 5 years with flood lines if available or record of intensity/damage.

Objective:	Based on damage severity
F1:	Number of locations where floods occurred
F2:	Frequency of flood incidences
F3:	Amplitude/Excursion: ranking based on extent of damage to flood regulation supply

5.3.4 Exposure to water-associated diseases (ExWD)

Spatial Unit:	Cities/sub-basin; aggregate to basin
Type/Class of Input required:	Identify important water associated diseases for the region and incidence or fatality rate
Suggested source of 'minimum' data to enable calculation:	From local administration for last 5 years
Objective:	Based on incidence ratio or case-fatality ratio
F1:	Number of locations where disease occurs
F2:	Frequency of disease outbreak
F3:	Amplitude/Excursion: based on either incidence ratio or case-fatality ratio

Other notes:

As a broad heading, we are assessing water associated infectious diseases which can be classified into one of five categories (Yang et al. 2012):

- water-borne, enteric microorganisms (e.g., typhoid and cholera) that enter water sources through fecal contamination and cause infection through ingestion of contaminated water. Also includes water-borne pathogens (e.g., *Cryptosporidium, Giardia*) transmitted through ingestion of, or exposure to, contaminated water. They can be described as water-carried diseases and are a subset of water-borne diseases;
- water-based, caused by flukes or nematodes which have an aquatic phase to their life cycle, (e.g., schistosomiasis);
- water-related, transmitted by insect vectors that have an aquatic phase to their life cycle (e.g., malaria and trypanosomiasis);
- water-washed, transmission is due to poor personal and/or domestic hygiene resulting from a lack of appropriate water; and
- water-dispersed, infections of agents that proliferate in fresh water and enter the human body through the respiratory tract (e.g., *Legionella*).

5.4 STEPS FOR CALCULATIONS

1. **Determine the spatial unit (SU)** and produce a GIS layer that shows their location and coverage of the basin**.**
2. **Determine the time-period (or evaluation period) for the indicator calculation.** If the calculation is to go beyond F1, the evaluation period has to be divided into smaller duration of

time (termed here as 'instances'). For example, for an evaluation period of 5 years, each year may be considered as an instance over which to group the events. The test of whether a demand is met or not is conducted within the period represented by the instance.

3. **Determine from the data whether F1 and F2 can be calculated.** If information regarding the number of SUs affected by lack of delivery of ecosystem services is available, then F1 can be calculated using the following formula:

$$F1 = \left(\frac{Number\ of\ SUs\ that\ did\ not\ meet\ demand\ at\ least\ once}{Total\ number\ of\ SUs}\right) \times 100$$

If distribution of events where demand is not met is available over the evaluation period is available, then F2 can be calculated by considering over which instances was demand met or not met:

$$F2 = \left(\frac{Number\ of\ instances\ where\ demand\ was\ not\ met}{Total\ number\ of\ instances\ monitored}\right) \times 100$$

4. **Determine from the data if F3 can be calculated and is 'sharp' or 'fuzzy.'** If any information for the ability to meet demand over the instances and magnitude of departure is available, it may be possible to calculate F3. This evaluation procedure is encapsulated within a measure of 'excursion' for each instance that is deemed to return a non-compliant value. Excursion for each instance *i* (Ex_i) can be calculated as follows:

a) Services where a univariate 'sharp' threshold for non-compliance can be defined:

Here, an objective value (such as target volume to meet water demand) for that particular instance can be defined, and thus, excursion for each case where demand is not met can be evaluated.

When target is to not fall short of this objective, excursion can be defined as:

$$Ex_i = \left(\frac{objective_i}{instance\ value_i}\right) - 1$$

Alternately, when the target is to not exceed the objective, excursion can be defined as:

$$Ex_i = \left(\frac{instance\ value_i}{objective_i}\right) - 1$$

b) Services where a univariate 'sharp' threshold for non-compliance cannot be defined:

Here, a single objective may be hard to define. We recommend that in these, excursion for each instance *i* be ranked on a scale of 1 to 10 to correspond with low to high gap between demand and supply. The values can be defined through stakeholder surveys or through tracking and combining a few metrics relevant to the ecosystem service. For *n* instances among the SUs where objective is not met is then collated and into a normalized sum of excursions (nse) such that:

$$nse = \frac{\sum_{i=0}^{n} Ex_i}{Total\ number\ of\ instances\ monitored}$$

Note that for the normalization process, the total number of instances monitored – whether demand is met or not met – is used. This is done so that the excursions are scaled with respect to all information available about the system and not biased toward instances where demand is not met.

Finally, F3 is now calculated by scaling *nse* to a 0-100 scale using the asymptotic function proposed by Saffran et al. (2001):

$$F3 = \left(\frac{nse}{nse + 1}\right) \times 100$$

5. **Combine F1, F2 and F3.**
    - If able to only determine F1: $ESI = 100 - F1$ (low evidence)
    - If able to only determine F1 and F2: $ESI = 100 - \sqrt{F1 \times F2}$ (medium evidence)
    - If able to determine all three: $ESI = 100 - \sqrt{F1 \times F3}$ (high evidence)

5.5 WORKED EXAMPLES

The following two worked out examples using coarse/synthetic data help demonstrate the application process, the derivation of as well as tracking between the three dimensions <F1, F2, F3>.

5.5.1 Flooding in Bangladesh

Bangladesh has a relatively flat terrain dominated by the Ganges-Brahmaputra delta. Ganges, Brahmaputra and Meghan along with their tributaries crisscross the nation before flowing out to the Bay of Bengal – making it highly susceptible to flooding. In this example, we apply the ESI for flood regulation over the 64 districts covering Bangladesh.

Data source: Spatial database on major floods from Dartmouth Flood Observatory. Includes date and period of event, rough extents on area affected, main cause of flooding and estimates of lives lost as well as people displaced.

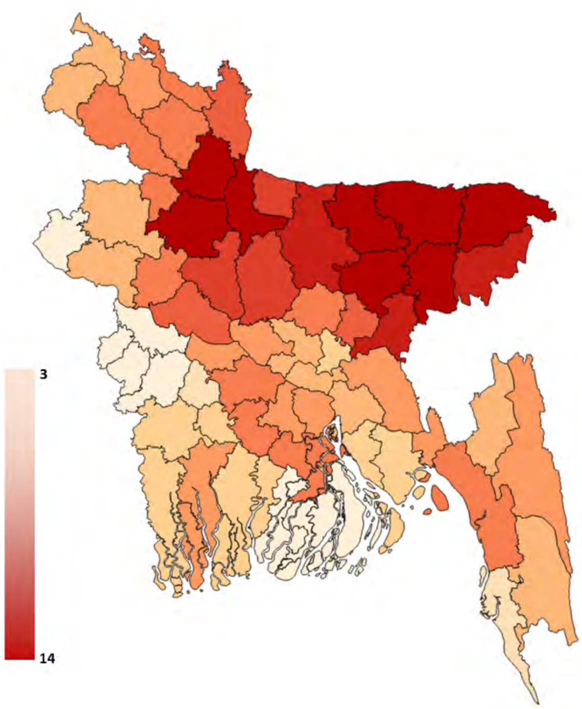


Figure 7: Number of floods affecting each district over the 17-year period ranging from 2000-2016.

Processing: A sub-set of the data was extracted for the period 2000-2016, where flood sources include monsoonal rain and failure of defenses. The spatial extent of each flood was then intersected with the district outlines. The flood frequency table thus derived is shown in Table 8, which tabulates the number of floods affecting each district for each year over the 17-year period. In terms of the indicator calculation, this give 17 years' x 64 districts = 1088 instances. Figure 7 denotes the distribution of the floods obtained from the data. The data available is sufficient for first approximation of scope and frequency. The information associated with number of lives lost and people displaced due to the floods is summarized in the plot of Figure 8. Flood depth-damage relationships are currently not available nor applicable given the coarse nature of the flooding information available. Thus, if we attempt to gauge the magnitude of the floods for amplitude calculation, a 'fuzzy' approach may be appropriate.

Table 8. Flood frequency for 64 districts in Bangladesh from 2000-2016, grouped over 1-year intervals

objectid	divcode	distcode	division	district	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	SUM
1	4	1	Khulna	Bagerhat				1			1	2			1	1						6
2	2	3	Chittagong	Bandarban	2			1				1					1	1		1		7
3	1	4	Barisal	Barguna								2			1			1				4
4	1	6	Barisal	Barisal				1				3			2	1		1			1	9
5	1	9	Barisal	Bhola								2						1				3
6	5	10	Rajshahi	Bogra	1	1	1	1		1		2	2		2				1		1	13
7	2	12	Chittagong	Brahamanbaria	1			1	1	1		2	2		1			1	1		1	12
8	2	13	Chittagong	Chandpur	1			1				3			1			1				7
9	2	15	Chittagong	Chittagong	2			2				2					1	1		1		9
10	4	18	Khulna	Chuadanga				1			1	1									1	4
11	2	19	Chittagong	Comilla	1			1				3			1			1			1	8
12	2	22	Chittagong	Cox's Bazar	1			1				1					1			1		5
13	3	26	Dhaka	Dhaka				1				2			1			1	1		1	7
14	5	27	Rangpur	Dinajpur	1	1	1	1		2		1	2									9
15	3	29	Dhaka	Faridpur				1			1	3			2			1			1	9
16	2	30	Chittagong	Feni	2			1				3						1		1		8
17	5	32	Rangpur	Gaibandha	1	1	1	1		2		2	2		1				1		1	13
18	3	33	Dhaka	Gazipur				1				2	2		1			1	1		1	9
19	3	35	Dhaka	Gopalganj				1			1	2			2	1		1			1	9
20	6	36	Sylhet	Habiganj	1	1	2	1	1	1		2	2					1	1			13
21	3	39	Dhaka	Jamalpur	1	1	1	1		1		2	2		2				1		1	13
22	4	41	Khulna	Jessore				1			1	2				1					1	6
23	1	42	Barisal	Jhalokati								2				1		1				4
24	4	44	Khulna	Jhenaidah				1			1	1									1	4
25	5	38	Rajshahi	Joypurhat	1	1		1		1		1	2						1		1	9
26	2	46	Chittagong	Khagrachhari	2			1				2						1		1		7
27	4	47	Khulna	Khulna		1		1			1	2			1	1					1	8
28	3	48	Dhaka	Kishoreganj	1	1	1	1	1	1		2	2					1	1		1	13
29	5	49	Rangpur	Kurigram	1	1	1	1		2		1	2		1							10
30	4	50	Khulna	Kushtia				1			1	1									1	4
31	2	51	Chittagong	Lakshmipur				1				3			1			1				6
32	5	52	Rangpur	Lalmonirhat	1	1	1	1		2		1	2									9
33	3	54	Dhaka	Madaripur				1				3			2			1			1	8
34	4	55	Khulna	Magura				1			1	1			1						1	5
35	3	56	Dhaka	Manikganj				1				2			2			1	1		1	8
36	6	58	Sylhet	Maulvibazar	1		1	1	1	1		2	2		1			1	1			12
37	4	57	Khulna	Meherpur				1			1	1										3
38	3	59	Dhaka	Munshiganj				1				3			1			1			1	7
39	3	61	Dhaka	Mymensingh	1	1	1	1				2	2		1			1	1		1	12
40	5	64	Rajshahi	Naogaon		1		1		1		1	2								1	7
41	4	65	Khulna	Narail				1			1	2				1					1	6
42	3	67	Dhaka	Narayanganj				1				2			1			1			1	6
43	3	68	Dhaka	Narsingdi	1			1				2	2		1			1	1		1	10
44	5	69	Rajshahi	Natore		1		1		1		1	2		1				1		1	9
45	5	70	Rajshahi	Nawabganj		1		1					2									4
46	3	72	Dhaka	Netrakona	1	1	2	1	1	1		2	2					1	1		1	14
47	5	73	Rangpur	Nilphamari	1	1		1		2		1	2									8
48	2	75	Chittagong	Noakhali	1			1				3						1				6
49	5	76	Rajshahi	Pabna		1		1			1	1	2		2				1		1	10
50	5	77	Rangpur	Panchagarh		1		1		2		1	2									7
51	1	78	Barisal	Patuakhali								2						1	1			4
52	1	79	Barisal	Pirojpur							1	2			1	1		1				6
53	3	82	Dhaka	Rajbari				1			1	2			2				1		1	8
54	5	81	Rajshahi	Rajshahi		1		1		1		1	2								1	7
55	2	84	Chittagong	Rangamati	2			1				2					1	1		1		8
56	5	85	Rangpur	Rangpur	1	1	1	1		2		1	2									9
57	4	87	Khulna	Satkhira		1					1	2			1	1						6
58	3	86	Dhaka	Shariatpur				1				3			2			1			1	8
59	3	89	Dhaka	Sherpur	1	1	1	1				2	2		1				1		1	11
60	5	88	Rajshahi	Sirajganj		1		1		1		2	2		2				1		1	11
61	6	90	Sylhet	Sunamganj	1	1	3	1	1	1		2	2					1	1			14
62	6	91	Sylhet	Sylhet	1		3	1	1	1		2	2		1			1	1			14
63	3	93	Dhaka	Tangail		1	1	1				2	2		1			1	1		1	11
64	5	94	Rangpur	Thakurgaon		1		1		2		1	2									7

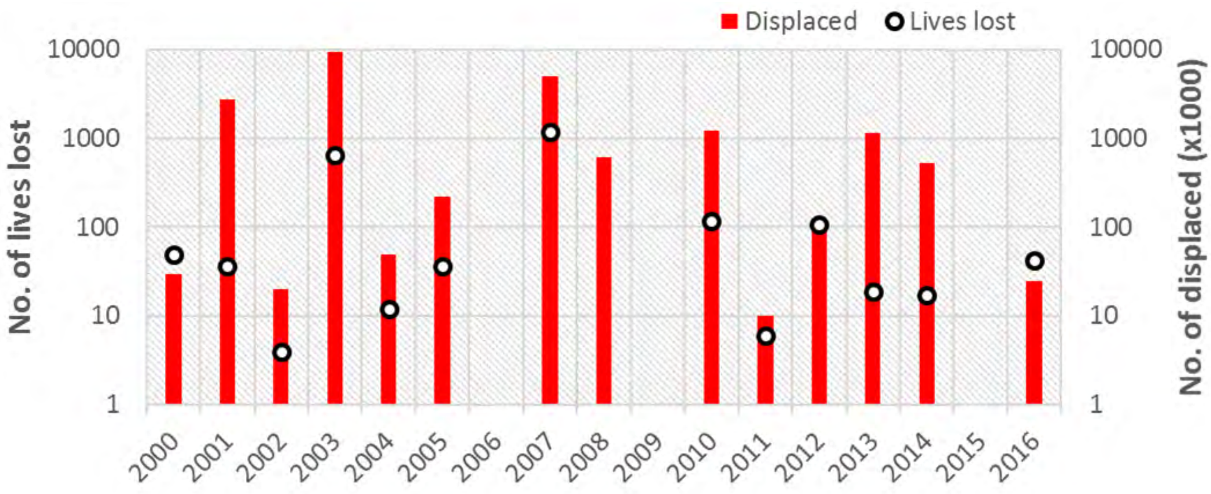


Figure 8. Data on lives lost and people displaced from the flooding.

Scope (F1): Since all districts are affected by flooding over the evaluation period – minimum 3 floods as seen in Figure 7 – F1 is straightforward at 100.

Frequency (F2): From the frequency table, of the 1088 instances (64 districts x 17 years), 409 instances see flooding. Therefore, F2 = (409/1088) * 100 = 37.59

The ecosystem service Indicator (ESI) calculated just using F1 and F2 would therefore follow as:

$$ESI\ (with\ F1\ andF2) = 100 - \sqrt{100 \times 37.59} = \ 38.70$$

Amplitude (F3): Since a clear definition of threshold is not available, here we test the method proposed for 'fuzzy' objectives, where for each instance the 'excursion' needs to be ranked from 1 to 10. To test the sensitivity of this approach, in the first case, we assign all excursions as 10. This is equivalent to saying that since all floods lead to either some lives lost or some displacement of people, they are all unacceptable and, thus, get the highest excursion. Following the formulas in Section 5.4 (Step 4b), F3 becomes 78.99 and, thus, ESI = 11.12.

On the other hand, if we assign as excursions as 1 (equivalent to saying that, although these floods are damaging, still they are part of the system and not a deviation from the natural extents and frequency), F3 becomes 27.32 and, thus, ESI = 47.73.

5.5.2 Water supply reliability in Dongjiang

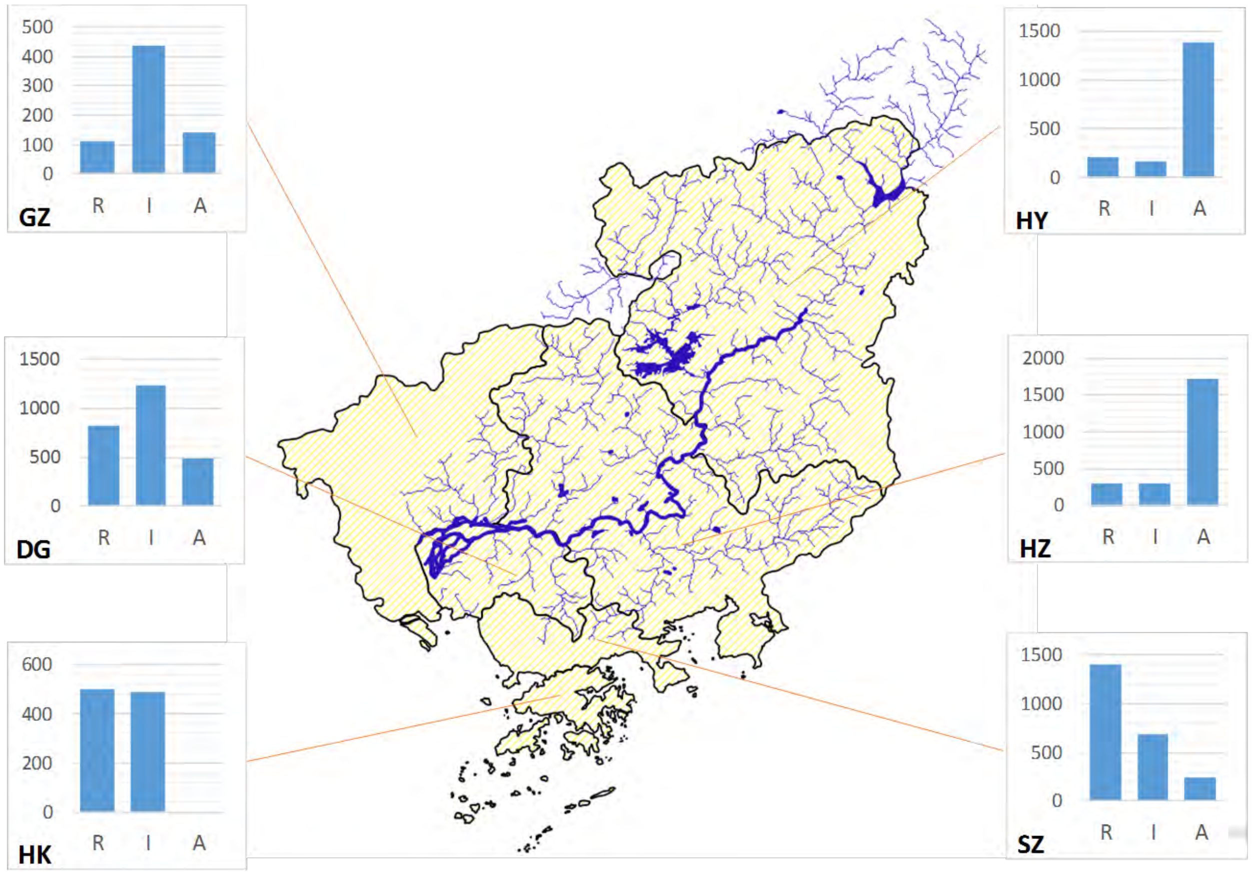


Figure 9. Annual Water Demand, broken down by main city/provinces and sector in million cub.m.

The Dongjiang or East River is part of the Pearl River System and provides freshwater for five Chinese mainland cities alongside being the main water supply source for Hong Kong. In this worked example, we apply ESI to examine the water supply reliability based on estimated demand and supply of freshwater from the river.

Data source: Study published by Zhang et al. (2008). The water demands obtained from this study are shown in Figure 9. In the study, Conditional Reliability Module (CRM) in the Water Right Analysis Package (WRAP) was used in the study to generate reliability of water supply for the intra-basin cities, such as Heyuan (HY), Huizhou (HZ) and Dongguan (DG) and cities located outside the basin, such as Hong Kong (HK), Shenzhen (SZ) and Guangzhou (GZ). Demand is further divided among the following sectors: residential use (R), industry (I) and agriculture (A).

Processing: We interpolate the results from a simulation of a severe drought for the year 1991 (annual results available from the study) based on available rainfall data to generate a synthetic dataset of monthly reliability of water supply for each demand. This is tabulated in the supplementary material (Table 9) and depicted in the plot of Figure 10.

With five cities having all three sectors, and Hong Kong having only two, we have the equivalent of (5x3 + 1x2 =) 17 spatial units. The evaluation period is one year broken down into 12 months. Therefore, the total number of instances are (17x12 =) 204 instances. Since, in this case, the threshold is univariate; non-compliance and its magnitude can be calculated using formulas in Section 5.4 (Step 4a) with objective being 100% reliability. The three dimensions <F1, F2, F3> thus obtained are <52.94, 17.65, 33.64>. Finally, the indicator value obtained by combining the three dimensions for the synthetic severe drought year in 1991 is 57.80.

Table 9. Water supply reliability for Dongjiang; monthly values interpolated from annual averages

	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Annual
HK[R]	100	100	100	100	100	100	100	100	100	100	100	100	100
HK[I]	100	100	100	100	100	100	100	100	100	100	100	100	100
SZ[R]	100	100	100	100	100	100	100	100	100	100	100	100	100
SZ[I]	100	100	100	100	100	100	100	100	100	100	100	100	100
SZ[A]	100	100	100	100	100	100	100	100	100	100	100	100	100
HY[R]	100	100	100	100	100	100	100	100	100	100	100	100	100
HY[I]	100	100	100	100	100	100	100	100	100	100	100	100	100
HY[A]	100	100	100	55	40	30	20	20	100	100	100	100	72.29
HZ[R]	100	100	100	100	100	100	100	100	100	100	100	100	100
HZ[I]	100	100	100	100	100	100	45	55	100	100	100	100	91.76
HZ[A]	100	100	100	40	20	10	5	5	100	100	100	100	65.09
DG[R]	100	100	100	100	75	85	30	36	100	100	100	100	85.47
DG[I]	100	100	100	100	70	80	20	30	100	100	100	100	83.33
DG[A]	100	100	100	100	75	85	25	36	100	100	100	100	85.09
GZ[R]	100	100	100	100	70	80	20	30	100	100	100	100	83.33
GZ[I]	100	100	100	100	70	80	20	30	100	100	100	100	83.33
GZ[A]	100	100	100	100	75	85	25	36	100	100	100	100	85.9

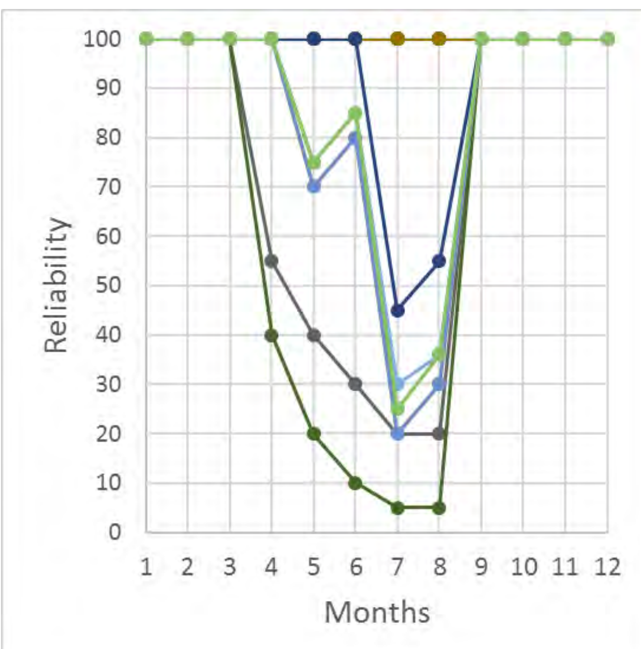


Figure 10. Reliability of supply interpolated to monthly values.

Sensitivity between the multiple levels of information:

Since ESI can be calculated with multiple levels of information (with F1 alone, with F1 and F2, and with all three), it's important to gauge the sensitivity of the result as more or less information is available. In other words, if information that allows only <F1> or only <F1, F2> to be calculated is available, does the final indicator value derived have any skill in depicting the state of the system?

To test for this, we construct a simple experiment using the Monte Carlo method. Under this, the reliability for each 204 instances in Table 9 is generated at random. The number generation is such that each instance in a set of experiments has a fixed probability of failure. For example, if the probability of failure is set at 10%, then any randomly generated reliability of industrial supply for Hong Kong has a 10% probability of getting a reliability value below

100%. We run the experiment for three different values of probability of failure: 10%, 20% and 66.67%, respectively. Under each set of experiments, 5,000 tables of reliability are generated. The results from these simulations are plotted in Figure 11.

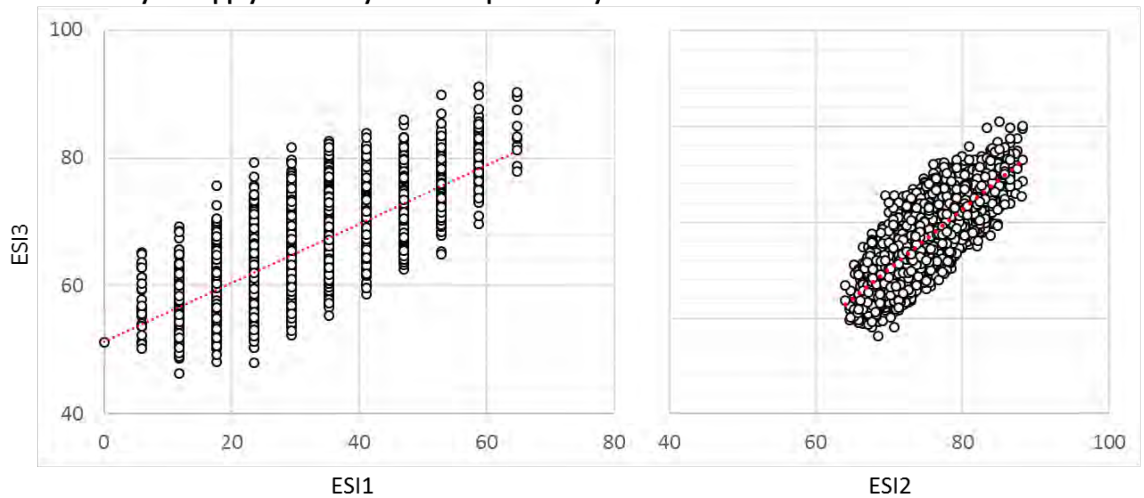
The results for all three simulations show some degree of tracking between the ESIs calculated with different levels of information. With 17 SUs, ESI with F1 can, in principle, take 18 different values ranging from no SUs non-compliant to all SUs non-compliant. However, in the simulations we see a lower range. This is due to the very low probability for ordered compliance with increasing probability of failure at any instance. For example, the probability for compliance at any instance in the reliability table for the 10% probability case is 0.9. For the whole year to be compliant in scope, all 12 instances have to be compliant. Therefore, the probability becomes $0.9^{12} = 0.282$. In Figure 11, we see 12 ESI values for F1 ranging from 0 SUs compliant to 11 SUs compliant. For 12 SUs compliant, the probability will become $0.282^{12} = 2.5E\text{-}7$. As the probability of failure increases, this range further decreases. And in the experiment with probability of failure set to 66.67%, no case is generated by the Monte Carlo simulation where any SU is compliant in scope.

ESI with F1 is seen to give a lower evaluation of the ecosystem service with values shifting significantly as a system becomes more prone to failure. However, some degree of tracking is present; lower ESI values (with F1) would have skill in predicting lower ESI values (with F1, F2 and F3). However, the value itself may differ significantly. ESI (with F1 and F2) and ESI (with F1, F2 and F3) track closer when the system is less prone to failure – with the regression line having nearly 45-degree slope and the values tightly spread around the regression line. As the probability of failure increases, this relationship shifts, although still tracking closer in terms of magnitude than ESI (with F1).

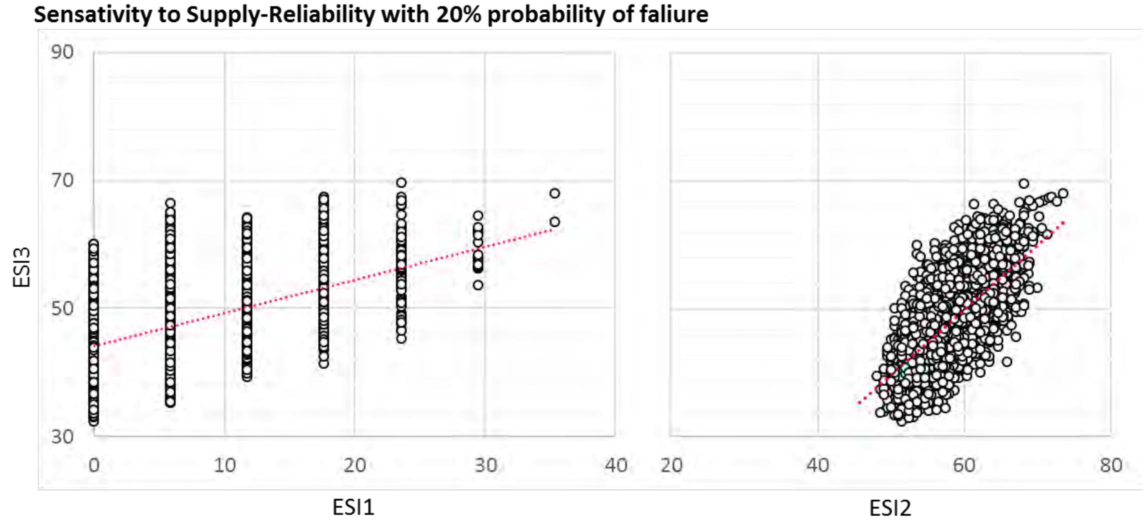
The sensitivity seems to suggest that while calculating ESI with 1 or 2 dimensions only, as in cases similar to this test, the compliance of ecosystem services will be underestimated – and the degree of underestimation cannot be established with high confidence without prior information about amplitude. However, even with this limitation, the ESI calculated with lower number of dimensions clearly demonstrates skill in depicting the state of the system and has value in gauging the gap between supply and demand.

Figure 11: Sensitivity with (a) 10% probability of failure; (b) 20% probability of failure; and (c) 66.67% probability of failure.

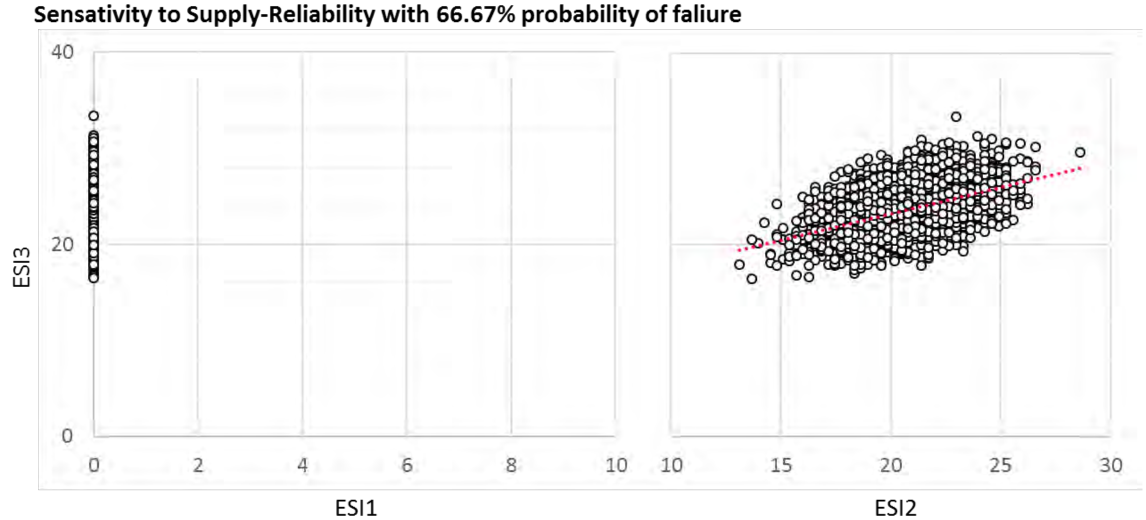
Sensativity to Supply-Reliability with 10% probability of faliure



Sensativity to Supply-Reliability with 20% probability of faliure



Sensativity to Supply-Reliability with 66.67% probability of faliure



5.6 CULTURAL SERVICES

This class of ecosystem services is the most difficult to assess quantitatively. It requires a simultaneous understanding of the ecological and cultural contexts (Daniel et al. 2012) and the changes in their values are not always clearly linked to ecological changes (Chan et al. 2012). Cultural services are also typically "bundled," in the sense that an ecosystem that holds heritage, spiritual and/or inspirational values may very well also provide more tangible recreational benefits (Plieninger et al. 2013). Recreational benefits (e.g. angling and sport fishing; tourism) are the most commonly quantified (Hernandez et al. 2013), but there have been numerous attempts to assess the less tangible cultural services. In monetary valuation exercises, these services should not be assessed separately as they are not independent and evaluating them as distinct services leads to double counting.

We strongly recommend considering a context-appropriate measure for cultural services in the Freshwater Health Index. In basins where you expect that conservation and cultural heritage benefits (5.6.1) will be highly correlated with recreation benefits (5.6.2), it will be suitable to measure just one, though some information on the omitted service should be included in the accompanying narrative. Below, we briefly summarize some of the most common proxies (based on secondary data) used in other assessments as well as methods that can be employed to collect primary data more tailored to the specific issue(s) in your basin.

5.6.1 Conservation/Cultural Heritage sites

Cultural heritage, in this context, refers to biophysical features, historical objects as well as traditional practices that are significant in some way to the present. We include species conservation within this category as representing the existence/bequest benefit of maintaining biological and physical assets for future generations. Examples include Aboriginal fish traps (Bark et al. 2014), Pacific lampreys (Close et al. 2002) or the Balinese *subak* system for water allocation (Lansing and Fox, 2011). One common method to assess these benefits is by measuring protected/conservation areas (Bottrill et al. 2014) as this is a signal that stakeholders have identified values worth preserving in these areas.

1) **Create map with area boundaries:** Begin by estimating the total area within the basin that has a designation as "protected" for their heritage values. Protected areas (PAs) may not be explicitly preserving water-related cultural services and so use judgment to determine whether to exclude these. The World Database on Protected Areas (https://www.protectedplanet.net/) offers a starting point, but this should be supplemented with information on other national/provincial/local recognized sites. If you elect to use area as the final metric, it would have to be scaled against what might be considered the "ideal" amount of protected area, which is a nontrivial task. There is also a range of management effectiveness within these areas, meaning that, in principle, some areas are better able to protect, and thus provide heritage values, than others. Finally, the cultural value of these sites is not necessarily a function of their size. In fact, their scarcity may contribute to their value. For these reasons, it is advisable to complete a second step in the assessment, by weighting by protected area (PA) according to the relative amount of cultural value they provide.

2) **Weighting sites**: There are various ways to assign weights to PAs. Participatory research is usually recommended so that stakeholders inform these differential values (Hernandez et al. 2013). As a starting point, weights could be assigned based on a hierarchy of PA classifications. For example, sites with the UNESCO World Heritage or Ramsar designations might be weighted highest, with

descending weights assigned to nationally, provincially and locally recognized sites. However, consultation with stakeholders help refine this approach as well as provide insight into the specific features (e.g., water quality) that influence their perception of the relative value of a site. A summary of participatory methods for doing so can be found in Chan et al. (2012) and a more specific description of engaging stakeholders to quantify values from sites on a pre-identified map in Plieninger et al. (2013).

5.6.2 Recreation

Water features are frequently an important part of outdoor recreation sites, whether for activities such as fishing, boating and swimming, or as a scenic setting for hiking or birdwatching. Recreation and eco-tourism are sometimes treated separately, with the latter being a sub-class that decisionmakers may want to isolate as an economic activity. For the purposes of the Freshwater Health Index, this distinction is unnecessary.

Fishing has been one of the more common recreation activities quantitatively assessed, because it often requires travel (a proxy for its monetary "value") and many jurisdictions require and keep data on licenses. However, many recreational sites may be multi-purpose, so having a more comprehensive measure of recreation is often desirable. Gathering additional data on the features that make a site attractive for recreation also enable one to estimate changes in the future.

Assessments may begin by estimating the recreation *potential* within a basin, or even the *opportunity,* which also take account of site accessibility (Paracchini et al. 2014). Inputs typically include data on water quality, proximity to other sites, ecological integrity of the surrounding landscapes, roads and trails and population location. These maps provide a baseline that can be confirmed with actual data on visits or used to model how future changes could affect recreation. Their main limitation is that they measure the potential *supply* of the service only.

We recommend using some measure of demand, in the form of person-days of use, or an economic proxy such as the travel cost one incurs to recreate. National parks, for example, typically collect information on visitors, but this is less likely to be the case for less prominent recreational sites. The advent of ubiquitous digital cameras, geo-tagging and online photo sharing makes it possible to harness "revealed preference" data from recreators and not only map demand but also use regression analysis to assess explanatory variables, such as landscape features or proximity to major roads (Tenerelli et al. 2016). The social media site Flickr makes its database of geotagged photographs accessible and offers a promising source of data for low-cost assessments of outdoor recreation demand, though the photographs are skewed towards North America and Europe (Wood et al. 2013).

The alternative, more conventional (and costly) method to collect data on recreation demand is to conduct a survey at the site(s) of interest. This requires skilled surveyors to be posted at recreation sites to collect data from visitors (the distance they travelled, amount of time they are spending, activities they engaged in and features they enjoy about the site). Thus, while it is the most hands-on and time consuming approach to an assessment, it provides the most reliable estimates of demand (which can be converted into an economic valuation) and helps identify less obvious forms of recreation, such as daily strolls in a riverside park, that are nonetheless valuable to stakeholders.

6. GUIDELINES FOR EVALUATING GOVERNANCE & STAKEHOLDERS INDICATORS

6.1 WHO ARE STAKEHOLDERS?

Stakeholders are the actors—from individual citizens to community groups to local and international organizations—that depend on freshwater services in a basin and/or are involved in the decisions that affect the basin. Stakeholders can be *internal* or *external* to the basin, with the latter group not being directly dependent on the basin's resources but nonetheless interested in the outcomes and influential through policy dialogue and financing. These stakeholders operate within the governance sub-system and its rules, but at the same time, they may play a role in changing (or subverting) the rules.

In large or complex basins, it may be necessary to begin with a stakeholder mapping exercise, to identify the key groups (or individuals) and their interest in and influence over decisions in the basin. Typical categories include governments and their line agencies (national, state/provincial and local), river basin organizations, businesses (including private service providers), farmers, fishermen, researchers and civil society groups. Additional stakeholders may include inter-governmental organizations, international lending institutions and international non-profit organizations.

6.2 SURVEY

Of the 12 proposed sub-indicators within the Governance & Stakeholders category, the majority involve some amount of subjectivity reflecting stakeholders' perceptions. While it is desirable to use objective, empirical data, such as counting the number of multi-stakeholder meetings a river basin organization holds in a year, such data are imperfect proxies for the actual principles or processes of interest. When measuring governance, perception data (what an individual believes to be occurring) is particularly valuable, because decision makers base their actions on their perceptions, and there is also frequently a divergence between *de facto* and *de jure* governance (Kaufman et al. 2010). Put another way, in a perception-based survey, individuals are asked to apply a subjective rating scale rather than answer only objective questions (yes/no or numeric responses). To develop comprehensive, systematic and comparable data for our Governance & Stakeholder indicators, we recommend deploying a survey to a cross section of stakeholders and then subsequently repeated for the next round of assessments.

6.2.1 Implementing the survey

A survey instrument has been developed to be administered to a group of regional experts familiar with water management issues in the basin. The instrument has been designed to correspond to 11 of the Governance and Stakeholder sub-indicators – one "module" per sub-indicator, containing between 3-7 questions each (see Appendix B). The master file is in English but has also been translated into Chinese and other languages. The original version has been screened and approved by Conservation International's Institutional Review Committee, but before administering the survey, it is advisable to ensure that it is also compliant with your own institution's research ethics policy as well as any other ethical policies that might be in place in the jurisdictions where you wish to survey.

The survey was designed to take approximately one hour for respondents to complete and can be administered in person or remotely (e.g., mailed or sent electronically), in one-to-one settings or in groups. An in-person group meeting may be most efficient for the surveyor(s) and can be built into the agenda of a broader meeting. If administered in person, the surveyor(s) must take care to not suggest answers or otherwise bias respondents; information should be limited to the survey instructions and clarifying terms that may be unclear.

In principle, any stakeholder in the basin can take the survey, but participants should only complete the questions that they feel qualified to answer. Many questions require familiarity with details about water resource management and related topics, and so we recommend using a non-probability sampling technique referred to as expert sampling. *This is not meant to be a representative sample of the population*, which requires probability sampling and hundreds or thousands of respondents to infer that results represent the perception of the general population in the basin. Instead, respondents should be invited based on their experience with water governance issues in the basin. We recommend referring to your stakeholder mapping first and identifying stakeholder groups with high levels of interest and engagement in the basin, as they should be the most familiar with the current governance dynamics. With perception-based surveys, there are no “wrong” answers, but respondents should be able to explain their responses and to provide insight into areas of poor (or strong) performance as well as areas of disagreement among respondents.

Your sampling should cover all tiers of government, along with industry representatives, non-profit organizations and academic researchers. There is no minimum number of respondents per stakeholder group, and in some cases, it may be desirable for stakeholder groups to be given an opportunity to formulate an “official” response that incorporates multiple inputs and allows them time for consultation before answering. Survey responses should be kept anonymous, but we recommend recording respondents’ sector affiliation (e.g., provincial government) as well as their location within the basin, which can be as simple as “upstream,” “mid-stream,” and “down-stream.” With this information, and with enough respondents, it is possible to analyze the data for differences or commonalities. A larger number of respondents reduces the influence any one individual respondent has on results and can even offer some statistical significance to analyzing sectoral or geographical differences. It should be made clear to respondents and in subsequent communications about the results that the results reflect the perceptions of an expert panel, and the number of respondents should be noted.

It may also be necessary to allow respondents to reconsider and adjust their responses after learning about the mean values for the group. Respondents would receive the averaged results (and standard deviations) calculated for the group and then be asked to reconsider their initial responses that diverge from the mean. However, the means do not represent a “true value” since the topics are subjective, and so respondents can have legitimate reasons for deviating substantially from the mean. For this reason, standard deviations should be calculated and recorded as a measure of the uncertainty of the final indicator values.

While the priority of the survey is to elicit information that translates into quantitative sub-indicators, it will likely be necessary and useful to follow up and elicit qualitative information, particularly in areas of poor performance. The timing and amount of effort on this will relate to expectations for the narrative report that accompanies the indicators. This is one advantage of administering the survey on a one-to-one basis, though further interviewing should take place once the survey has been completed. If qualitative information is collected through a group discussion format, there is no guarantee that all perspectives will be heard.

6.2.2 Survey questions

See Appendix B for full survey.

6.2.3 Analyzing responses

A spreadsheet tool has been developed to help record and analyze responses (reference to where it can be downloaded). Prior to analyzing responses, you should determine whether to weight individual survey questions. Since each survey module contains multiple questions, the individual questions can be weighted before being aggregated into a numeric score for the sub-indicator. For example, for a question on “Resource Rights,” respondents may want to place greater weight on water abstraction (A) and emissions/pollution rights (B) than rights related to land use (C) and fisheries (D). The same principles apply as outlined in Section 2.4 if weights are to be assigned at this stage, but the default is to leave each question unweighted.

We recommend against assigning weights to individuals but this is also a possibility. In practice, it is used to compensate for under- or over-representation within a sample, but *a priori*, there is no sound way to determine what the demographics of a multi-stakeholder group should look like.

Begin recording responses by assigning an alpha-numeric code to each respondent. We suggest this so that you can note the sector (e.g., “G” for government) and location of the respondent, but the important point is to have a unique identifier for each respondent. Each respondent occupies one column in the spreadsheet.

The survey template uses a 0-10 rating scale for each question, so these scores should be entered accordingly. For unanswered questions, leave the cells blank. We have adopted a 0-10 scale for three reasons. First, it is a better approximation of an interval-type scale than, for example, a 5-point scale, which is ordinal and not intuitively linked to equidistant intervals. The scores on the survey are multiplied by 10 to correspond to values in the Governance and Stakeholders component of the Index. Second, having a wider range for responses should help avoid responses clustering too tightly to the (neutral) midpoint. Third, and related, having the wider range should help reveal smaller changes in between assessment periods.

As respondents’ data are entered, the number (N), average (Ave) and standard deviation (SD) are automatically calculated for each question and then aggregated into sub-indicator, major indicator and a final score. Like the overall Index, these indicators are aggregated using a weighted geometric mean. By default, weights are set to be equal for each question, for each sub-indicator and for each major indicator. As discussed in Section 2.4, there may be valid reasons to have users provide weights for these steps, which will affect aggregated scores. Alternatively, the spreadsheet can be used as an exploratory tool, to determine how different weighting scenarios would affect aggregated scores. The magnitude of difference is unlikely to be great, but the weighting exercise itself can be useful to gain further insight into respondents' underlying preferences, and using the weighted values should be a better approximation of the "true" score.

With perception data, though, there are bound to be disagreements about the true score for a question relating to an indicator. For this reason, the SD provides an initial screening test to highlight areas with highly variable scores amongst respondents. If using the 0-10 scale, we recommend flagging questions with a SD > 2 for further investigation amongst the group of respondents. While there are valid reasons for respondents to perceive the same issue quite differently, it also could be a matter of individuals interpreting the *scale* differently. Once initial results have been analyzed, having a follow up discussion on items of disagreement and then subjecting respondents to the same survey to allow them to adjust responses is recommended.

With a sufficient sample size (>20, and representatives from different sectors and locations), it may be worthwhile to examine respondents' characteristics as explanatory factors. It is unlikely that statistically significant differences will be found among groups, but the identifying information is collected to explore that possibility. It also helps reveal potential biases attributable to a respondent's location within the basin or sectoral affiliation. This information does not invalidate responses, but it can provide useful context when interpreting index values and developing policy and water management decisions.

7. STRESSORS AND SCENARIOS

One of the purposes of the Freshwater Health Index, beyond assessing the health of a basin, is projecting the effect of stressors and scenarios on the values of the indices in the face of risk and uncertainty. This can help to evaluate the effect of potential planning efforts on freshwater health, rank alternative proposed management actions with respect to freshwater health, and explore the projected effects of new environmental conditions or changes in the basin.

Scenario planning, or scenario analysis, is a framework for exploring options and developing more robust plans in the face of irreducible uncertainty (Peterson et al. 2003). Scenarios are highly uncertain yet plausible futures which can include novel stressors. They could represent plausible future states of a system under different climate projections, different proposed management or development plans such as the placement of a dam or expansion of irrigation, or different uncertain effects of a management or development plan on the system such as low, medium or high deviations from natural flow regime under a dam siting proposal or diffuse changes, such as forest loss or increases in fertilizer use.

While the overriding objective of the Freshwater Health Index is to assess freshwater sustainability, its strengths lie in highlighting trade-offs between ecosystem services and evaluation of proposed management or development plans to provide decision support at the basin scale where management decisions have greatest relevance. For instance, evaluations of the indices under different climate change or management scenarios can be performed in the Freshwater Health Index framework (see Table 10) to project a range of outcomes which can then be subject to decision rules (Regan et al. 2005; Polasky et al. 2011). In cases where data is lacking and directional changes in indicator values can be ascertained, they can be useful in judging the relative effects of different scenarios or rankings of management or development proposals. Ranking of options can be sufficient to inform decisions under uncertainty in many circumstances. Moreover, the framework is intended to be embedded within an adaptive management framework, updating information as we learn more about the effects of uncertain scenarios, such as climate change and its impacts.

Table 10. Major types of environmental change and indicators most likely to be directly affected

Climate change	**Land-use change***	**Water allocation change**
Deviation from natural flow	Deviation from natural flow	Deviation from natural flow
Groundwater storage	Water quality	Groundwater storage
Biodiversity	Drainage basin condition	Biodiversity
Water stress/reliability	Biodiversity	Water stress/reliability
Biomass for consumption	Water stress/reliability	
Flood regulation	Sediment regulation	
	Flood regulation	
	Conservation/cultural heritage	

*Here we include decisions about the sites of for dam construction as a special case of land-use change though, in practice, dam siting scenarios should be evaluated independently.

7.1 SCENARIOS

The aim of scenario analysis using the Freshwater Health Index approach is to enable stakeholders to consider the changes encompassing multiple dimensions of water-related ecosystems and ecosystem services. Instead of focusing on defining a definitive outcome on the state of the system in response to future changes and developments, the process intends to help stakeholders identify potential trade-offs they are likely to face while managing the freshwater system and assess what steps they might carry out to mitigate risk to the overall health of the system. Figure 12 below illustrates the evolution of the freshwater social-ecological system with changes in forcing over time. Working through scenarios may help explore the impacts of such forcing on freshwater health.

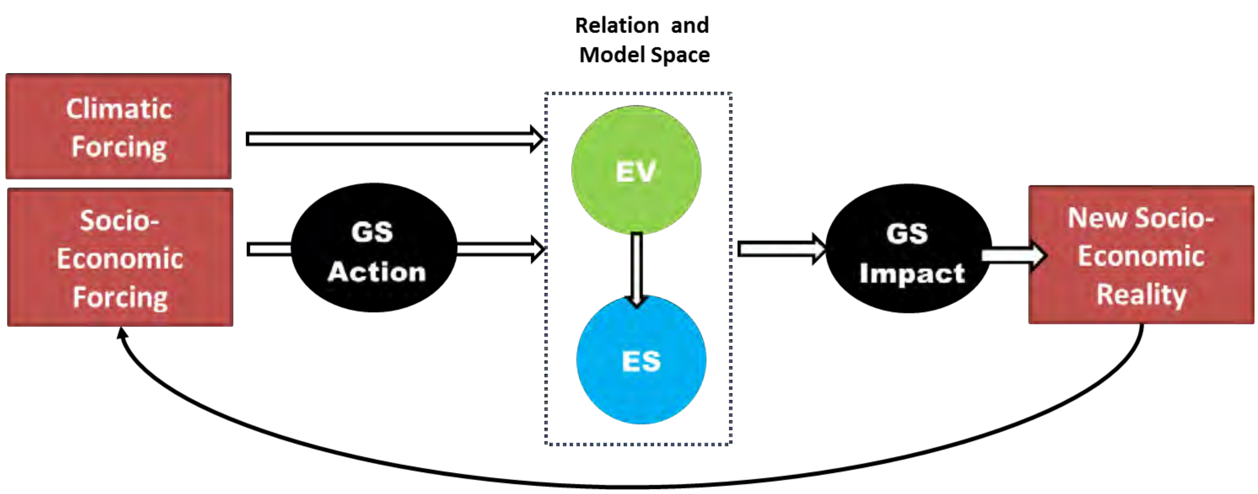


Figure 12. Climate and socio-economic forcing on the Freshwater Health Index components.

* NOTE: GS refers to Governance & Stakeholders, EV refers to Ecosystem Vitality and ES refers to Ecosystem Services.

In the sections below, we identify four major classes of modifications that may be relevant to the development of a freshwater basin. For each class, in general, the stakeholders can work through the following steps:

1. Define and quantify as best possible the change in parameters for freshwater systems. For example, in a certain climate scenario, the expected change might be that precipitation will drop by x%. The aim is to be comprehensive, covering the main changes that may be introduced into the system, yet brief.
2. With a diverse set of stakeholders, work through the indicators for ecosystem vitality and ecosystem services that have been calculated in the current assessment and, for each indicator, identify (in response to step 1): the direction of change in the indicator and expected shift in terms of percentage drop or rise in the indicator value.

3. Where models are available that can be used to project value for all or a subset of indicators based on the parameters modified in step 1, this exercise should be carried out to build confidence in the values obtained in step 2.
4. If the values obtained in step 2 and 3 are significantly different, consultation between stakeholders and experts should be carried out to explore the underlying causes.
5. Based on the outcomes of steps 2, 3 and 4, stakeholders should identify which dimensions of freshwater health (captured through the respective indicator) are being stressed, if at all, due to forcings and consider the responses they may carry out to mitigate negative impacts.
6. Depending on the scenario and the action considered, stakeholders may want to iterate over the steps or explore the impact of alternative management actions.

7.1.1 Global Climate Change

Three major areas of uncertainty are at play in decision making in the face of climate change: 1) uncertainty in the climate projections due to the different general circulation models (GCMs) employed and the gas emission scenarios used as input for these models (IPCC 2014), 2) uncertainty in the impacts to freshwater ecosystems and hydrology under these projections, and 3) uncertainty in the effects of specific management and planning decisions on freshwater systems under climate and associated changes (Lawler et al. 2010).

A warmer climate accompanied by changes in precipitation patterns will affect hydrologic regimes, biogeochemical cycling, community composition and productivity, and wetland ecosystem structure and function (Arnell and Gosling 2013, Grimm et al. 2013, Pyne and Poff 2016, van Dijk et al. 2015). Sea level rise will likely inundate many areas, increasing the salinity of freshwater wetlands, triggering salt water intrusion in aquifers and altering biotic communities and water quality (Craft et al. 2009; Weston et al. 2006). Alterations to natural disturbance regimes, such as fire or intense hurricanes, could also have significant effects on freshwater health (Michener et al. 1997). Shifts in the frequency, timing and intensity of rainfall events can affect the transport of sediments, nutrients and other constituents from wetlands as well as precipitate larger flooding events (Arnell and Gosling 2016). Perturbations in hydroperiod and hydrologic flows can significantly affect aquatic communities and associated biogeochemical processes that ultimately have effects on water quality (Delpla et al. 2009). The temporal sensitivity of freshwater resources to climate change ranges from within-year to annual to multi-year to centuries (Ford and Thornton 2012).

The Ecosystem Vitality and Ecosystem Services components will be the most relevant for examining the effects of climate change on the indicators. Evaluation of the indicators under climate change scenarios should use the underlying hydrologic model as the foundation to which climate-related data is linked (see section 3.2). Projected changes in precipitation will be of most relevance in assessing the effects of climate change on hydrological systems, and this will directly affect evaluations of deviation from natural flow regime and change in groundwater storage. Ideally, projections of precipitation should be based on established general circulation models most suitable for the basin; they should span a range of low to high emissions scenarios; and they should be downscaled to a meaningful or readily available spatial scale. In the absence of such data, hypothetical scenarios can be used to explore the effects of changes in precipitation, e.g., baseline scenario plus or minus 10% change in rainfall. The outputs of the hydrologic model with projected precipitation then feed into other models in the model chain (Figure 3) to enable calculation of other indicators (e.g., water quality, flood damage models).

Changes in temperature may also affect the biotic components of freshwater ecosystems, such as the population sizes of species of concern and invasive species. This will be difficult to estimate and will rely on information on physiological tolerances. However, changes in the distribution of species habitat due to changes in precipitation and temperature can be modeled using Species Distribution Models (SDMs), which indicate relative changes in biodiversity.

To fully evaluate effects of sea level rise on coastal basins, a fully three-dimensional (3-D) circulation model of the basin, such as that described by Zheng and Weisberg (2012), is recommended (NRC 2014). Ideally, this should be coupled with a regional hydrologic model and a regional atmospheric model (e.g., Maxwell et al. 2011). Understanding salinity intrusion in coastal wetlands and aquifers used for water supply requires a surface water flow model coupled with a variable-density groundwater flow model. Since it is unlikely that this is possible for most coastal basins, it will be necessary to include the effects of sea level rise in a piecemeal and incomplete fashion without the benefit of a systems model that couples all the relevant components. To this end, plausible hypothetical scenarios (e.g., 1.5-meter sea level rise) may be the only option in examining the effects of sea level rise on freshwater systems. In such cases, it will also be necessary to consider the range of effects this has on the indicators most likely to be affected. For instance, under water quality, it will be necessary to consider a measure of salinity in the climate change scenario analysis even if that was not considered in the baseline assessment.

We recommend the following steps in applying the Freshwater Health Index to evaluate the outcomes of climate change:

1. Assessors are encouraged to first think systematically through the potential mechanisms of the impact of climate change on the basin. The identification of likely mechanisms of impact will help with defining key indicators used in Freshwater Health Index assessments in the context of climate change. This diagnostic process may be aided by development of diagrammatic models.

2. Assessors should identify the indicators relevant to the mechanisms of change in the indices under climate change identified in Step 1. These indicators will most likely include Water Quantity, Water Quality and Biodiversity under Ecosystem Vitality, and Provisioning, Water Quality and Regulation and Support under the Ecosystem Services component.

3. To incorporate future climate impacts on the indicators more explicitly, assessors are encouraged to identify available data and models that can be used to estimate indicators under climate change. When models exist but relevant climate change data is not available to parameterize models, assessors should select plausible scenarios for temperature, precipitation and sea level rise. Climate scenarios should span a range of high to low emissions scenarios or high to low changes in temperature, precipitation and sea levels. In cases where models are not available to estimate indicators (e.g., water quality), inferences informed by data, results in peer-reviewed literature or expert opinion will need to be made on the effects of these changes in the relevant indicators. If this is not possible, then such indicators should not be subject to evaluation under climate change, and the baseline values for these indicators should be used in the assessment. A meaningful time horizon should be selected that is relevant to the management and planning goals in the basin and can reflect anticipated changes but also aligns with the climate projections available for the basin. For example, a projection to 2040 is in the range of typical spatial planning parameters, whereas a projection for 2100, in addition to being even more uncertain, far exceeds most planning and management scenarios.

4. Finally, hydrological models coupled with other systems models (see Figure 3) should be run with the climate data or scenarios to provide values for the indicators under the climate change scenarios. These can then be aggregated (as described above) to provide a scenario-relevant index for each component that can then be compared to the baseline indices to gauge projected changes in freshwater health.

7.1.2 Land-use Change

Land-use change includes the conversion of land from one state to another for human use (e.g., natural areas to croplands) or, less frequently, as restoration to a more natural condition. Land-use change remains the most serious threat to biodiversity, and it undermines the capacity of ecosystems to supply fresh water (Foley et al. 2005). Several indicators in the Ecosystem Vitality and Ecosystem Services components could change with land-use change, most notably and directly drainage-basin condition under Ecosystem Vitality. Depending on the type of land-use change, other indicators may be affected as well. For instance, if natural land is converted to cropland, ancillary Ecosystem Vitality effects are likely, such as declines in biodiversity due to reduction in habitat, declines in water quality from fertilizer use and surface run-off, and greater deviations from the natural flow regime or changes in groundwater storage due to increased irrigation pressure. The Ecosystem Services indicators may also be affected with some declines and some increases in indicator value. Using natural-to-cropland land-use change as an example, greater stress on the water supply will likely occur (resulting in a decrease in the average annual water stress indicator value) but total amount of biomass for consumption will increase. Regulation and support indicators may also decline, such as water quality metrics and changes in sedimentation and nutrient retention. If the natural lands served as recreational areas, the cultural/aesthetic indicator may also decline or increase depending on the recreational use (e.g., bird watching opportunities may increase in certain types of croplands).

The Freshwater Health Index can also be used as a tool to explore the effects of infrastructure planning. Freshwater infrastructure includes dams for hydropower or water consumption, and dredging, channelization or straightening of rivers for navigation. All of these can have effects on Ecosystem Vitality and Ecosystem Services indicators. With the Ecosystem Services component, infrastructure can have detrimental effects on flood regulation, water quality and sedimentation, but it can also improve intra-annual variability of supply relative to demand. It can affect water quantity and quality, drainage basin condition and biodiversity indicators under Ecosystem Vitality. The suite of indicators that might change under an infrastructure planning scenario will be highly context dependent. However, the steps outlined below provide guidance on the strategy for proceeding with a land-use change or infrastructure planning scenario.

We recommend the following steps in calculating the Freshwater Health Index under land-use change scenarios:

1. Assessors should ascertain the exact type and extent of land-use change at the outset of the scenario assessment and the timeframe for conversion. This should include the location and area of conversion, the type of conversion (from state A to state B – see section 4.3.2) and the specific types of activities that will occur upon conversion.

2. Assessors are encouraged to think systematically through the potential mechanisms of the impact of the land-use change on the basin. The identification of likely mechanisms of impact will help with defining key indicators used in Freshwater Health Index assessments in the context of land-use change. This diagnostic process may be aided by development of diagrammatic models.

3. Assessors should identify the indicators relevant to the mechanisms of change in the indices under land-use change identified in Step 2. These indicators could include any under the Ecosystem Vitality and Ecosystem Services components.

4. To incorporate future land-use change impacts on indicator values more explicitly, assessors are encouraged to identify available data and models that can be used to estimate indicators under the type of land-use change imposed. For example, a map of the location, area and type of proposed land-use change can directly inform calculation of the land cover naturalness indicator, which can in turn be used to inform changes in the number and population size of species of concern. In cases where models are not available to estimate indicators (e.g., recreation use), inferences informed by data, peer-reviewed literature or expert opinion will need to be made on the effects of these changes in the relevant indicators. If this is not possible, then such indicators should not be subject to evaluation under land-use change, and the baseline values for these indicators should be used in the assessment.

5. Finally, available models (see Figures 3 and 4) should be run with the data relevant to the type, location and amount of proposed land-use change to provide values for the indicators under this scenario. These can then be aggregated (as described above) to provide a scenario-relevant index for each component that can then be compared to the baseline indices to gauge projected changes in freshwater health.

7.1.3 Water Allocation and Trade-offs

Since fresh water is often limited, and numerous needs exist within a basin, water of sufficient quality must be allocated to different uses such as municipal human consumption, agriculture, industry, energy, environmental flows, etc. Water needs to be allocated in a manner that achieves economic, social and environmental goals. Water allocation, therefore, makes trade-offs between the priorities of stakeholders, reliability of water supply, equity, economic growth and maintenance of ecosystems (http://www.sswm.info/content/water-allocation). Water allocation can change for a variety of economic, social or environmental reasons, including greater need for municipal water due to population growth, industrial or agricultural expansion, or drought, as well as water diversion or transfer projects to deliver water into or out from a basin.

We recommend that the Freshwater Health Index be used as additional decision support to supplement existing water allocation models to explore the effect of water allocation decisions in the basin. For instance, the REALM (REsource ALlocation Model) (Perera et al. 2005) or the Water Evaluation and Planning (WEAP) water-allocation model (Yates et al. 2005) can be used to structure the water allocation scenarios and provide input on water quantity and quality for the Ecosystem Vitality component and provisioning and regulation/support for Ecosystem Services component for the Freshwater Health Index. Other indicators also may be relevant for different water allocation scenarios, e.g. Conservation/Cultural/Heritage sites. The values of the Freshwater Health Indices can then be used to rank alternative water allocation scenarios as additional information in the decision-making process.

8. UPDATING ASSESSMENTS

It is recommended that assessments are updated every five years unless rapid change occurs, in which case the assessment should be conducted in response to the change. However, scenario analysis can be undertaken at any time and will be prompted by the need for input into decision making.

9. REFERENCES

Alessa, L., Kliskey, A., Lammers, R., Arp, C., White, D., Hinzman, L., and Busey, R. 2008. The arctic water resource vulnerability index: an integrated assessment tool for community resilience and vulnerability with respect to freshwater. Environmental Management, 42(3):523-541.

Angermeier PL. 2000. The natural imperative for biological conservation. Conservation Biology 14(2):373-381.

Arnell, N. W., and Gosling, S. N. 2013. The impacts of climate change on river flow regimes at the global scale. Journal of Hydrology, 486:351-364.

Arnell, N. W., and Gosling, S. N. 2016. The impacts of climate change on river flood risk at the global scale. Climatic Change, 134(3):387-401.

Bark, R. H., Barber, M., Jackson, S., Maclean, K., Pollino, C., & Moggridge, B. 2015. Operationalising the ecosystem services approach in water planning: a case study of indigenous cultural values from the Murray–Darling Basin, Australia. International Journal of Biodiversity Science, Ecosystem Services & Management, 11(3):239-249.

Binder, C.R., Hinkel, J., Bots, P.W.G. and Pahl-Wostl, C. 2013. Comparison of Frameworks for Analyzing Social-ecological Systems. Ecology and Society, 18(4):26.

Bottrill, M., Cheng, S., Garside, R., Wongbusarakum, S., Roe, D., Holland, M. B., ... & Turner, W. R. 2014. What are the impacts of nature conservation interventions on human well-being: a systematic map protocol. Environmental Evidence, 3(1):16.

Brown B.B. 1968. Delphi Process: A Methodology Used for the Elicitation of Opinions of Experts. The RAND Corporation, Santa Monica, CA.

Brown C and Lall U. 2006. Water and economic development: The role of variability and a framework for resilience. Natural Resources Forum **30**(4):306-317.

Burroughs R. 1999. When stakeholders choose: process, knowledge, and motivation in water quality decisions. Society and Natural Resources **12**(8):797-809.

Cai, M., McKinney, D.C. and Lasdon, L.S. 2002. A framework for sustainability analysis in water resources management and application to the Syr Darya Basin. Water Resources Research, 38(6):21-2 – 21-14.

Caro, T. 2010. Conservation by proxy: Indicator, umbrella, keystone, flagship and other surrogate species. Island Press, Washington, D.C.

CCME (Canadian Council of Ministers of the Environment). 2001. Canadian water quality guidelines for the protection of aquatic life: CCME Water Quality Index 1.0, User's Manual. In: Canadian environmental quality guidelines, 1999, Canadian Council of Ministers of the Environment, Winnipeg. http://www.ccme.ca/files/Resources/calculators/WQI%20User's%20Manual%20(en).pdf

Clarkin K., Connor A., Furniss M.J., Gubernick B., Love M., Moynan K. et al. 2005. National inventory and assessment procedure—for identifying barriers to aquatic organism passage at road-stream crossings. US Dep Agric for Serv. http://www.fs.fed.us/biology/nsaec/fishxing/publications/PDFs/NIAP.pdf

Close, D. A., Fitzpatrick, M. S., & Li, H. W. 2002. The ecological and cultural importance of a species at risk of extinction, Pacific lamprey. Fisheries, 27(7):19-25.

Cote D, Kehler DG, Bourne C, and Wiersma YF. 2009. A new measure of longitudinal connectivity for stream networks. Landscape Ecology **24**(1):101-113.

Covello, V. T., and Merkhoher, M. W. (2013). Risk assessment methods: approaches for assessing health and environmental risks. Springer Science and Business Media.

Craft, C., Clough, J., Ehman, J., Joye, S., Park, R., Pennings, S., ... and Machmuller, M. 2009. Forecasting the effects of accelerated sea-level rise on tidal marsh ecosystem services. Frontiers in Ecology and the Environment, 7(2):73-78.

de Groot RS, Wilson MA, and Boumans RM. 2002. A typology for the classification, description and valuation of ecosystem functions, goods and services. Ecological Economics **41**(3):393-408.

Daniel TC, Muhar A, Arnberger A, et al. 2012. Contributions of cultural services to the ecosystem services agenda. Proceedings of the National Academy of Sciences **109**(23):8812-8819.

Dickin SK, Schuster-Wallace CJ, Elliott SJ. 2013. Developing a vulnerability mapping methodology: Applying the Water-Associated Disease Index to Dengue in Malaysia. PLoS ONE 8(5): e63584. doi:10.1371/journal.pone.0063584

Dudgeon D, Arthington AH, Gessner MO, et al. 2006. Freshwater biodiversity: importance, threats, status and conservation challenges. Biological Reviews **81**:163-182.

Foley, J. A., DeFries, R., Asner, G. P., Barford, C., Bonan, G., Carpenter, S. R., ... and Helkowski, J. H. 2005. Global consequences of land use. Science, 309(5734):570-574.

Ford, D.E. and Thornton, K.W. 2012. Water resources in changing climate. In Firth, P. and Fisher, S.G. (eds.) Global Climate Change and Freshwater Ecosystems. Springer Verlag, NY. Pp 26-47.

Gehrke P, Brown P, Schiller CB, Moffatt DB, and Bruce A. 1995. River regulation and fish communities in the Murray–Darling river system, Australia. Regulated Rivers: Research and Management 15:181–198.

Gippel, C.J., Zhang, Y., Qu, X., Kong, W., Bond, N.R., Jiang, X. and Liu, W. 2011. River health assessment in China: comparison and development of indicators of hydrological health. ACEDP Australia-China Environment Development Partnership, River Health and Environmental Flow in China. The Chinese Research Academy of Environmental Sciences, the Pearl River Water Resources Commission and the International WaterCentre, Brisbane, September.

Gleick PH. 1998. Water in crisis: paths to sustainable water use. Ecological Applications 8:571-579.

Gregory KJ. 2006. The human role in changing river channels. Geomorphology 79:172-191.

Grey D and Sadoff CW. 2007. Sink or swim? Water security for growth and development. Water Policy 9(6):545-571.

Grimm, N. B., Chapin, F. S., Bierwagen, B., Gonzalez, P., Groffman, P. M., Luo, Y., ... and Schimel, J. 2013. The impacts of climate change on ecosystem structure and function. Frontiers in Ecology and the Environment, 11(9):474-482.

GWP (Global Water Partnership). 2009. Global Water Partnership Strategy 2009-2013. Stockholm, Sweden. http://www.gwp.org/Global/GWP-CEE_Files/About/GWP_Strategy_2009-2013_final.pdf

Haines-Young, R., and Potschin, M. 2010. The links between biodiversity, ecosystem services and human well-being. In Raffaelli, D., and C. Frid (eds.), Ecosystem Ecology: A New Synthesis. BES Ecological Reviews Series. Cambridge: Cambridge University Press, pp.110–139.

Haines-Young, R. and Potschin, M. 2013. Common International Classification of Ecosystem Services (CICES): Consultation on Version 4, August-December 2012. EEA Framework Contract No EEA/IEA/09/003 (Download at www.cices.eu or www.nottingham.ac.uk/cem)

Hart, B. T., Maher, B., and Lawrence, I. 1999. New generation water quality guidelines for ecosystem protection. Freshwater Biology 41(2):347-359.

Hooper B. 2010. River basin organization performance indicators: application to the Delaware River basin commission. Water Policy 12:461-478.

IPCC, 2014: Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Core Writing Team, R.K. Pachauri and L.A. Meyer (eds.)]. IPCC, Geneva, Switzerland, 151 pp.

IUCN. 2012. IUCN Red List Categories and Criteria: Version 3.1. Second edition. Gland, Switzerland and Cambridge, UK: IUCN. iv + 32pp.

Ivey JL, Smithers J, de Loë RC and Kreutzwiser RD. 2004. Community capacity for adaptation to climate-induced water shortages: linking institutional complexity and local actors. Environmental Management 33:36-47.

Konikow LF and Kendy E. 2005. Groundwater depletion: A global problem. Hydrogeology Journal 13:317-320.

Ladson AR, White LJ, Doola, JA, Finlayson BL, Hart BT, Lake PS, and Tilleard JW. 1999. Development and testing of an Index of Stream Condition for waterway management in Australia. Freshwater Biology 41(2):453-468.

Lansing, J. S., & Fox, K. M. 2011. Niche construction on Bali: the gods of the countryside. Philosophical Transactions of the Royal Society of London B: Biological Sciences, 366(1566):927-934.

Lawler, J.J., Tear, T.H., Pyke, C., Shaw, M.R., Gonzalez, P., Kareiva, P., Hansen, L., Hannah, L., Klausmeyer, K., Aldous, A., Bienz, C., and Pearsall, S. 2010. Resource management in a changing and uncertain climate. Frontiers in Ecology and the Environment 8(1):35–43.

Le TV., Nguyen HN, Wolanski E, Tran TC, and Haruyama S. 2007. The combined impact on the flooding in Vietnam's Mekong River delta of local man-made structures, sea level rise, and dams upstream in the river catchment. Estuarine, Coastal and Shelf Science 71(1):110-116.

Lehner, B., Grill G. 2013. Global river hydrography and network routing: baseline data and new approaches to study the world's large river systems. Hydrological Processes, 27(15):2171–2186. Data is available at www.hydrosheds.org.

Lemos MC and Agrawal A. 2006. Environmental governance. Annual Review of Environment and Resources 31:297-325.

Loh, J., Green, R. E., Ricketts, T., Lamoreux, J., Jenkins, M., Kapos, V., and Randers, J. 2005. The Living Planet Index: using species population time series to track trends in biodiversity. Philosophical Transactions of the Royal Society B: Biological Sciences, 360(1454):289-295.

Machado, A. 2004. An index of naturalness. Journal for nature conservation, 12(2):95-110.

Maxwell, R. M., J. K. Lundquist, J. D. Mirocha, S. G. Smith, C. S. Woodward, and A. F. B. Tompson. 2011. Development of a coupled groundwater–atmosphere model. Monthly Weather Review 139:96-116.

McGinnis, M. D., and E. Ostrom. 2014. Social-ecological system framework: initial changes and continuing challenges, Ecology and Society, 19(2):30, 10.5751/es-06387-190230.

MEA (Millennium Ecosystem Assessment). 2005. Millennium Ecosystem Assessment Ecosystems and Human Well-Being: Synthesis, 2005. Island Press, Washington, DC.

Merkhofer, M. W. (2012). Decision Science and Social Risk Management: A Comparative Evaluation of cost-benefit analysis, decision analysis, and other formal decision-aiding approaches (Vol. 2). Springer Science & Business Media.

Michener, W. K., Blood, E. R., Bildstein, K. L., Brinson, M. M., and Gardner, L. R. 1997. Climate change, hurricanes and tropical storms, and rising sea level in coastal wetlands. Ecological Applications, 7(3):770-801.

Modarres, M. (2006). Risk analysis in engineering: techniques, tools, and trends. CRC press.

Moglia M, Alexander KS, and Sharma A. 2011. Discussion of the enabling environments for decentralised water systems. Water Science & Technology 63:2331-2339.

Moore, R. E. 1966. Interval analysis (Vol. 4). Englewood Cliffs: Prentice-Hall.

Naiman J, and Dudgeon D. 2011. Global alteration of freshwaters: Influences on human and environmental well-being. Ecological Resources 26:865–873.

Nazemi A, and Wheater HS. 2015. On inclusion of water resource management in Earth system models–Part 2: Representation of water supply and allocation and opportunities for improved modeling.

Hydrology and Earth System Sciences 19(1):63-90.

NRC (National Research Council). 2014. Progress Toward Restoring the Everglades: The Fifth Biennial Reivew, 2014. National Academies Press, Washington, DC. 240 pp.

OECD. 2008. Handbook on Constructing Composite Indicators: Methodology and User Guide. Available for download at http://www.oecd.org/std/42495745.pdf

OECD. 2015. Stakeholder Engagement for Inclusive Water Governance. OECD Publishing, Paris. DOI: http://dx.doi.org/10.1787/9789264231122-en

Ostrom E. 2009. A general framework for analyzing sustainability of social-ecological systems. Science 325:419-422.

Pahl-Wostl C, Conca K, Kramer A, Maestu J, and Schmidt F. 2013. Missing links in global water governance: a processes-oriented analysis. Ecology and Society, 18(2):33. http://dx.doi.org/10.5751/ES-05554-180233

Pahl-Wostl C. 2015. Water Governance in the Face of Global Change: From Understanding to Transformation. Springer International Publishing, Switzerland.

Paracchini, M. L., and Capitani, C. 2011. Implementation of a EU wide indicator for the ruralagrarian landscape. JRC Scientific and Technical reports, European Commission, Joint Research Centre. Institute for Environment and Sustainability.

Perera, B. J. C., James, B., and Kularathna, M. D. U. 2005. Computer software tool REALM for sustainable water allocation and management. Journal of Environmental Management, 77(4):291-300.

Peterson, G.D., Cumming, G.S. and Carpenter, S.R. 2003. Scenario planning: a tool for conservation in an uncertain world. Conservation Biology 17(2):358–366.

Poff NL, and Zimmerman JKH. 2010. Ecological responses to altered flow regimes: a literature review to inform the science and management of environmental flows. Freshwater Biology 55:194–205.

Polasky, S., Carpenter, S.R., Folke, C. and Keeler, B. 2011. Decision-making under great uncertainty: environmental management in an era of global change. Trends in Ecology and Evolution 26(8):398-404.

Prüss-Üstün A, Bos R, Gore F, Bartram J. 2008 Safer water, better health: Costs, benefits and sustainability of interventions to protect and promote health. Geneva: World Health Organization.

Pyne, M. I., and Poff, N. L. 2016. Vulnerability of stream community composition and function to projected thermal warming and hydrologic change across ecoregions in the western United States. Global Change Biology. doi: 10.1111/gcb.13437

Regan HM, Colyvan M, and Burgman MA. 2002. A taxonomy and treatment of uncertainty for ecology and conservation biology. Ecological Applications, 12(2):618-628.

Regan, H.M., Ben-Haim, Y., Langford, B., Wilson, W.G., Lundberg, P., Andelman, S.J. and Burgman, M.A.

2005. Robust decision making under severe uncertainty for conservation management. Ecological Applications 15(4):1471-1477.

Rodriguez, J.P., Beard, T.D.J., Bennett, E.M., Cumming, G.S., Cork, S.J., Agard, J., Dobson, A.P. and Peterson, G.D. 2006. Trade-offs across space, time, and ecosystem services. Ecology and Society, 11:28-41.

Rogers P, and Hall AW. 2003. Effective water governance (Vol. 7). Global water partnership. http://www.gwp.org/global/toolbox/publications/background%20papers/07%20effective%20water%20governance%20(2003)%20english.pdf

Saaty, T. L. 1990. How to make a decision: the analytic hierarchy process. European journal of operational research, 48(1):9-26.

Saffran, K., Cash, K., and Hallard, K. (2001). Canadian water quality guidelines for the protection of aquatic life. CCME Water Quality Index, 1(0):34-1.

Schultz, L., Folke, C., Österblom, H., and Olsson, P. (2015). Adaptive governance, ecosystem management, and natural capital. Proceedings of the National Academy of Sciences, 112(24):7369-7374.

Shields, P. and Rangarjan, N. 2013. A Playbook for Research Methods: Integrating Conceptual Frameworks and Project Management. Stillwater, OK: New Forums Press. p. 24.

Smakhtin, V., Revenga, C., Döll, P., Tharme, R., Nackoney, J., and Kura, Y. 2004. Taking into account environmental water requirements in global-scale water resources assessments (Vol. 2). IWMI.

Smith, E.T. and Zhang, H.X. 2004. Developing key water quality indicators for sustainable water resources management. Water Environment Federation. Available for download at http://acwi.gov/swrr/Rpt_Pubs/WEFTEC04_SWRR_Zhang.pdf

Sukopp, H. 2004. Human-caused impact on preserved vegetation. Landscape and urban planning, 68(4):347-355.

SWRR. 2005. Sustainable Water Resources Roundtable Preliminary Report. Advisory Committee on Water Information, Water Information Coordination Program, US Geological Survey, US Dept of Interior. Available for download at http://acwi.gov/swrr/Rpt_Pubs/prelim_rpt/index.html

TEEB (The Economics of Ecosystems and Biodiversity). 2011. The Economics of Ecosystems and Biodiversity in National and International Policy Making. Earthscan, London & Washington.

Tengberg A, Fredholm S, Eliasson I, Knez I, Saltzman K, and Wetterberg O. 2012. Cultural ecosystem services provided by landscapes: assessment of heritage values and identity. Ecosystem Services 2:14-26.

Tropp, H. 2007. Water governance: trends and needs for new capacity development. Water Policy, 9(S2):19-30.

Turkelboom, F. et al. 2014. Chapter 18 - CICES Going Local: Ecosystem Services Classification Adapted for a Highly Populated Country. Part IV Ecosystem Services: Tools and Practices. Ecosystem Services pp. 223-247.

UNEP. 2007. Global Drinking Water Quality Index Development and Sensitivity Analysis Report. Prepared and published by the United Nations Environment Programme, Global Environment Monitoring System (GEMS)/Water Programme.

UNEP. 2008a. Water Quality for Ecosystems and Human Health. United Nations Environment Programme, Global Environment Monitoring System/Water Programme, Ontario, Canada.

UNEP. 2008b. Water Quality: Development of an index to assess country performance. United Nations Environment Programme, Global Environment Monitoring System/Water Programme, Ontario, Canada.

UN-Water. 2013. Water Security and the Global Water Agenda. A UN-Water Analytical Brief. United Nations University, Canada. http://www.unwater.org/downloads/watersecurity_analyticalbrief.pdf

UN-Water. 2015. Water for a Sustainable World. The United Nations World Water Development Report 2015. United Nations Educational, Scientific and Cultural Organization, France. http://unesdoc.unesco.org/images/0023/002318/231823E.pdf

van Dijk, G., Smolders, A. J., Loeb, R., Bout, A., Roelofs, J. G., and Lamers, L. P. (2015). Salinization of coastal freshwater wetlands; effects of constant versus fluctuating salinity on sediment biogeochemistry. Biogeochemistry, 126(1-2):71-84.

Vogel RM, Lall U, Cai X, Rajagopalan B, Weiskel PK, Hooper RP and Matalas NC. 2015. Hydrology: The interdisciplinary science of water. Water Resources Research, 51(6):4409–4430.

Vollmer, D., Regan, H.M. and Andelman, S.J. In review2016. Assessing the sustainability of freshwater systems: A critical review of composite indicators. Ambio. 45(7):765-780.

Vrba, J. and Lipponen, A. (eds), 2007. Groundwater Resources Sustainability Indicators. IHP-VI Series in Groundwater No. 14. United Nations Educational, Scientific and Cultural Organization, Paris. http://www.hydrologie.org/BIB/Publ_UNESCO/SOG14.pdf

Walz, U., and Stein, C. 2014. Indicators of hemeroby for the monitoring of landscapes in Germany. Journal for Nature Conservation, 22(3):279-289.

Ward JV and Stanford JA. 1995. Ecological connectivity in alluvial river ecosystems and its disruption by flow regulation. Regulated Rivers: Research & Management 11(1):105-119.

Ward PJ, Jongman B, Salamon P, Simpson A, Bates P, De Groeve T, Muis S, de Perez EC, Rudari R, Trigg MA, and Winsemius HC. 2015. Usefulness and limitations of global flood risk models. Nature Climate Change 5(8):712-715

Weston, N. B., Dixon, R. E., and Joye, S. B. 2006. Ramifications of increased salinity in tidal freshwater sediments: Geochemistry and microbial pathways of organic matter mineralization. Journal of Geophysical Research: Biogeosciences, 111(G1).

WWAP (United Nations World Water Assessment Programme). 2015. The United Nations World Water Development Report 2015: Water for a Sustainable World. Paris, UNESCO. http://unesdoc.unesco.org/images/0023/002318/231823E.pdf

Yang K, LeJeune J, Alsdorf D, Lu B, Shum CK, and Liang S. 2012. Global distribution of outbreaks of water-associated infectious diseases. PLoS Neglected Tropical Diseases 6(2):e1483.

Yates, D., Sieber, J., Purkey, D., andHuber-Lee, A. 2005. WEAP21—A demand-, priority-, and preference-driven water planning model: part 1: model characteristics. Water International, 30(4):487-500.

Zheng, L., and Weisberg, R. H. 2012. Modeling the west Florida coastal ocean by downscaling from the deep ocean, across the continental shelf and into the estuaries. Ocean Modelling, 48:10-29.

APPENDIX A: SUMMARY OF CHANGES TO THE GUIDELINES

Changes from version 1.0 to 1.1: in Section 4.4.1 the Species of Concern sub-indicator of Biodiversity has been changed to reflect the proportion of threatened and endangered species, of the total number of species assessed, in the river basin. Weights have been applied to the number of CR, EN, VU and nationally/provincially listed species.

APPENDIX B: GOVERNANCE AND STAKEHOLDERS SURVEY

Framework for Basin Management (1 of 12)

Integrated water resources management is a guiding framework for coordinating both development and management of all resources within a basin, to maximize welfare without compromising ecological sustainability. In some cases a single agency, such as a river basin authority, is responsible for coordinating and overseeing these functions; the questions below focus on the specific functions as managed within your jurisdiction (e.g. transnational, national or provincial) regardless of whether they are all carried out by the same agency.

Based on your own knowledge of the current situation, please evaluate the degree to which the following functions are being fulfilled throughout the basin. Provide a rating between 1 and 5 following the criteria below. Please skip any items which you do not feel qualified to answer.

Rating	Criteria
1	Function is almost never satisfactory (without conflicts among stakeholder groups)
2	Function is rarely satisfactory
3	Function is sometimes (~50%) satisfactory
4	Function is often satisfactory
5	Function is almost always satisfactory

A) **Policies and actions to advance water resource development and management are coordinated.** For example, but not limited to, if there is river basin organization or commission, how effective is it in coordinating the different agencies, levels of government (e.g., national, provincial, local), and private interests when establishing integrated development plans for the basin?

☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5

B) Infrastructure such as dams, reservoirs, and treatment plants are centrally managed or coordinated. Examples include, but are not limited to: dam operators communicating the timing and volume of reservoir releases, or assessing cumulative impacts of dams.

☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5

C) Financial resources are mobilized to support water resource development and management needs. Examples include, but are not limited to: cost-sharing for common projects, or collecting user fees/taxes.

☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5

D) **Ecosystems conservation priorities are developed and actions implemented.** Examples include, but are not limited to: protecting forested watersheds, maintaining wetland/river connectivity, or developing an aquatic species biodiversity action plan.

☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5

Rules for resource use (2 of 12)

Clear and enforceable rules are recognized as a requirement for the efficient use of scarce resources, and as a means of resolving conflicts. These rules encompass various uses and users of water, and can be both formal (i.e., legislated by a government body) or informal rules administered by communities.

Based on your own knowledge of the current situation, please evaluate the quality and stakeholders' understanding of rules concerning the use of various resources. Provide a rating between 1 and 5 following the criteria below. Please skip any items which you do not feel qualified to answer.

Rating	Criteria
1	Rules are very poorly articulated and/or understood or do not exist
2	Rules are poorly articulated and/or understood
3	Rules are acceptably articulated and/or understood
4	Rules are well articulated and/or understood
5	Rules are very well articulated and/or understood

A) **Quality and clarity of rules for allocating water among different sectors (e.g., municipal, industrial, agricultural)**
Examples include, but are not limited to: prioritizing water according to use, or limits on the timing and amount of water that can be withdrawn.

☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5

B) **Quality and clarity of rules for allocating water among administrative jurisdictions (e.g., cities, provinces, countries)**
Examples include, but are not limited to: determining withdrawals between provinces, or setting minimum flow requirements for rivers that cross administrative boundaries.

☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5

C) **Quality and clarity of rules for groundwater abstraction**
Examples include, but are not limited to: guidelines regarding the depth of wells, or amount of water that can be withdrawn within a certain time period.

☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5

D) **Quality and clarity of rules for wastewater handling and water pollution**
Examples include, but are not limited to: guidelines regarding the discharge of wastewater (e.g. pollutant concentrations, volume, temperature, time of release) into water bodies.

☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5

E) **Quality and clarity of rules for managing land use (including aquaculture) to safeguard water resources**
Examples include, but are not limited to: guidelines regarding soil management practices, the amount of forested land in watersheds, or the volume of runoff allowed for a given plot of land.

☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5

F) **Quality and clarity of rules for freshwater fisheries**
Examples include, but are not limited to: guidelines on catch limits, protected species, or fishing methods.

☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5

Incentives and regulations (3 of 12)

Various management tools, from conventional regulations to market-based instruments can be applied within a governance system. Having a variety of tools offers opportunities to increase the efficiency of interventions (e.g., cost per unit outcome) or lead to a more equitable distribution of benefits.

Based on your own knowledge of the current situation, please evaluate the development of the following management tools. Provide a rating between 1and 5 following the criteria below. Please skip any items which you do not feel qualified to answer.

Rating	Criteria
1	Instrument does not exist or is in earliest stage of discussion
2	Instrument is under development, e.g. guidelines have been circulated
3	Instrument has been developed and is being piloted, but guidelines are subject to refinement
4	Instrument is fully developed, but use is not yet standardized
5	Instrument is fully developed and a standard practice

A) **Environmental and social impact assessments for all major water projects, regardless of funding source, are carried out prior to decisions being taken**
Examples include, but are not limited to: environmental impact assessment (EIA) that is submitted to a government body for evaluation.

☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5

B) **Existence of financial incentives for environmental stewardship**
Examples include, but are not limited to: mechanisms for providing payments for watershed services provided by upstream stakeholders (e.g., farmers, forest managers, local governments).

☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5

C) **Existence of market-based exchange schemes**
Examples include, but are not limited to: tradeable water rights, wetland mitigation banking, or pollutant trading.

☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5

D) **Existence of honorary recognition programs**
Examples include, but are not limited to: publishing lists of industries with good environmental performance, or awards for local governments practicing good water stewardship.

☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5

E) **Existence of land use zoning policy**
Examples include, but are not limited to: requirements for riparian buffers, floodplain development, or forested catchment zones.

☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5

Technical capacity (4 of 12)

Lack of local capacity is often cited as an impediment to a variety of issues in resource management. Here we are referring to people employed in areas of water resource management, service delivery, monitoring and enforcement, and related research, but excluding international consultants.

Based on your own knowledge of the current situation, please evaluate the quality of human resources in water resource development and management in the basin. Provide a rating between 1 and 5 following the criteria below. Please skip any items which you do not feel qualified to answer.

Rating	Criteria
1	Level is very unsatisfactory
2	Level is unsatisfactory
3	Level is satisfactory
4	Level is very satisfactory
5	Level is extremely satisfactory

A) Number of staff (including local consultants) to fulfill necessary functions
Examples include, but are not limited to: backlogs (work waiting to be done) in a particular agency, or open positions remaining vacant due to lack of candidates.

☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5

B) Staff have sufficient expertise to fulfill necessary functions
Examples include, but are not limited to: hydrologists to evaluate a proposed dam, or fisheries ecologists to assess fish stocks.

☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5

C) Opportunities for professional training and certification
Examples include, but are not limited to: financial support or time allocated for continuing education courses related to improving technical skills.

☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5

Financial capacity (5 of 12)

Water resource development and management is often under-financed, particularly for services that do not generate revenue, such as ecosystem protection. Although financial capacity can be measured directly as a function of existing allocations relative to estimated budget needs, qualitative information is also useful in providing insights and identifying priorities.

Based on your own knowledge of the current situation, please evaluate the quality of human resources in water resource development and management in the basin. Provide a rating between 1 and 5 following the criteria below. Please skip any items which you do not feel qualified to answer.

Rating	Criteria
1	Level is very unsatisfactory
2	Level is unsatisfactory
3	Level is satisfactory
4	Level is very satisfactory
5	Level is extremely satisfactory

A) Level of investment in water supply development
Examples include, but are not limited to: financial resources for building and maintaining reservoirs or irrigation systems.

☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5

B) Level of investment in service delivery systems
Examples include, but are not limited to: financial resources for building and maintaining water distribution networks (i.e. piped supply) or household wells.

☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5

C) Level of investment in wastewater handling and treatment
Examples include, but are not limited to: financial resources for building and maintaining community toilets, or treatment systems to process waste water.

☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5

D) Level of investment in ecosystem conservation and rehabilitation
Examples include, but are not limited to: financial resources for protecting wetlands to mitigate flood risk, remediating impaired streams, or rehabilitating fish stocks.

☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5

E) Level of investment in monitoring and enforcement
Examples include, but are not limited to: financial resources for evaluating EIAs, collecting environmental data, inspecting facilities, and enforcing regulations.

☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5

Information and knowledge (6 of 12)

Sound water governance requires information on a range of topics and from many sources. Even in cases where data and information are abundant, if they are not made accessible (across agencies, with citizens, etc.) then they are less likely to aid in wise decision making.

Based on your own knowledge of the current situation, please evaluate the accessibility of information (including data on water quantity and quality, planning documents, and financial information), along with its quality of coverage and transparency (ability to be traced to the source). Provide a rating between 1 and 5 following the criteria below. Please skip any items which you do not feel qualified to answer.

Rating	Criteria
1	Almost never satisfactory
2	Rarely satisfactory
3	Sometimes (~50%) satisfactory
4	Often satisfactory
5	Almost always satisfactory

A) **Information is accessible to interested stakeholders**
Examples include, but are not limited to: reports made freely available through a website, or data available upon request to the agency with the information.

☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5

B) **Information meets expected quality standards, in terms of frequency, level of detail, and subjects of interest to stakeholders**
Examples include, but are not limited to: time series data on streamflow, water levels, or water quality for specific locations within the basin.

☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5

C) **Information is transparently sourced**
Examples include, but are not limited to: methods used to collect data are documented, or authors (source) of these data are clearly identified.

☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5

D) All available, sound and relevant information is routinely applied in decision-making
Examples include, but are not limited to: modifying an infrastructure project based on EIA results, or adjusting fisheries management guidelines based on fish catch data.

☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5

Engagement in decision-making processes (7 of 12)

Stakeholder engagement encompasses the process by which any person or group with an interest in a water-related topic can be involved in decision-making and implementation. It is associated with improved information transfer, better targeted and more equitable plans and policies, improved transparency and accountability, and reduced conflict.

Based on your own knowledge of the current situation, please evaluate the degree to which all stakeholders have a voice within the cycle of policy and planning for water resources development and management. Provide a rating between 1 and 5 following the criteria below. <u>Please skip any items which you do not feel qualified to answer</u>.

Rating	**Criteria**
1	Process as described almost never, or never occurs
2	Process as described rarely occurs
3	Process as described sometimes (~50%) occurs
4	Process as described often occurs
5	Process as described almost always, or always occurs

A) **All relevant stakeholders have been identified and notified when considering major decisions**
Examples include, but are not limited to: mapping and notifying stakeholders affected by a proposed water supply infrastructure project (e.g. construction of a water supply dam).

☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5

B) **Stakeholders are able to provide comments prior to major decisions being taken**
Examples include, but are not limited to: consultation meetings or an information gathering period where stakeholders may provide input regarding a policy or project.

☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5

C) Decisions are responsive to stakeholders' participation
Examples include, but are not limited to: processes for reaching joint agreements among a group of stakeholders prior to approval of a major policy or project, or projects being revised subsequent to stakeholder feedback.

☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5

Enforcement and compliance (8 of 12)

In many societies, there is a gap between laws and their actual enforcement, reflecting either insufficient capacity or a lack of accountability. Enforcement and compliance can be ensured through fines, incentives, or social pressure, but weak enforcement leads to poor management and a lack of confidence in the system.

Based on your own knowledge of the current situation, please evaluate how well existing regulations and agreements are enforced for the following areas throughout the basin. Provide a rating between 1 and 5 following the criteria below. Please skip any items which you do not feel qualified to answer.

Rating	Criteria
1	Enforcement is very poor or no guidelines (formal or informal) exist
2	Enforcement is poor
3	Enforcement is acceptable
4	Enforcement is good
5	Enforcement is very good

A) Surface water abstraction guidelines are enforced
Examples include, but are not limited to: industries restricted from withdrawing more than a specified amount of surface water, or farmers sanctioned for withdrawals during the dry season.

☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5

B) Groundwater abstraction guidelines are enforced
Examples include, but are not limited to: farmers or industries restricted from pumping more than a specified amount of groundwater.

☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5

C) Flow requirement guidelines are enforced
Examples include, but are not limited to: dam operators meeting the expectations of downstream water users, to meet environmental flows, human water needs, and/or flood protection.

☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5

D) Water quality guidelines are enforced
Examples include, but are not limited to: industries and communities complying with requirements related to pollutant discharges, or non-negotiable fines are levied on violators.

☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5

E) Land use guidelines are enforced
Examples include, but are not limited to: environmentally sensitive zones (e.g., catchment forests and wetlands) being protected from development or degradation.

☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5

Distribution of benefits from ecosystem services (9 of 12)

*Equity is an important issue in water resource management, most closely associated with access to safe water and sanitation. Here we extend the concept to include all benefits from ecosystem services in the basin (**water and sanitation, fisheries, flood mitigation, water quality maintenance, disease regulation, and cultural services**).*

Based on your own knowledge of the current situation, please evaluate quality of outcomes, in terms of their share of benefits from water resources, for the following stakeholder groups (groupings may overlap). Provide a rating between 1 and 5 following the criteria below. Please skip any items which you do not feel qualified to answer.

Rating	Criteria
1	Their share of benefits is almost never adequate
2	Their share of benefits is rarely adequate
3	Their share of benefits is sometimes (~50%) adequate
4	Their share of benefits is often adequate
5	Their share of benefits is almost always adequate

A) **Economically vulnerable populations benefit from ecosystem services**
Examples include, but are not limited to: poor households' access to improved water supply sources at a reasonable cost, protection from inland flood risks, or rural compared to urban populations' benefits.

☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5

B) Indigenous people benefit from ecosystem services
Examples include, but are not limited to: exercising customary rights related to water, including for consumptive as well as cultural uses, or maintaining traditional fisheries.

☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5

C) Women and girls benefit from ecosystem services
Examples include, but are not limited to: amount of time collecting water for households, or provision of toilets for females.

☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5

D) **Resource-dependent communities benefit from ecosystem services**
Examples include, but are not limited to: fishermen and smallholder farmers' incomes compared to other economic sectors.

☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5

Water-related conflict (10 of 12)

Tensions among stakeholders are expected when there is competition for scarce resources such as water. An effective governance system should prevent tensions from escalating into conflicts, here defined as a difference that prevents agreement, and therefore delays or undermines a decision taken with the basin.

Based on your own knowledge of the current situation, please evaluate the frequency of conflicts occurring over the past three years regarding water-related issues. Provide a rating between 1 and 5 following the criteria below. Please skip any items which you do not feel qualified to answer.

Rating	Criteria
1	Conflicts almost always occur
2	Conflicts often occur
3	Conflicts sometimes occur
4	Conflicts rarely occur
5	Conflicts almost never occur

A) **Frequency of conflict due to overlapping jurisdictions (e.g., between national governments in transboundary systems, provincial and national government, or between agencies)**
Examples include, but are not limited to: disputes between the local environmental bureau and a national ministry about authority within a floodplain, or between agencies in managing agricultural pollution.

☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5

B) **Frequency of conflict about water rights allocation**
Examples include, but are not limited to: disputes about how water is allocated between two municipalities, or between agricultural and industrial users.

☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5

C) **Frequency of conflict about access**
Examples include, but are not limited to: disputes about having access to safe water and sanitation, or the costs of such access.

☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5

D) Frequency of conflict regarding the siting of infrastructure
Examples include, but are not limited to: disputes about reservoir development and resettlement plans for residents and land owners, or downstream impacts to fisheries or water users.

☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5

E) Frequency of conflict over water quality and other downstream negative impacts
Examples include, but are not limited to: disputes between upstream and downstream stakeholders about dry season flows or pollution concentrations.

☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5

Monitoring mechanisms (11 of 12)

Policy and planning decisions about water resources management are ideally based on sound data and information, which must be collected on a regular basis. Monitoring entails costs and so data collection should be based on needs and assessed relative to resource constraints, where a comparatively wealthy basin might invest in higher spatial and temporal coverage of information.

Based on your own knowledge of the current situation, please evaluate the degree to which different types of data are being collected, analyzed, and used to inform decisions in the basin. Provide a rating between 1 and 5 following the criteria below. Please skip any items which you do not feel qualified to answer.

Rating	Criteria
1	Data are very poorly monitored, or not monitored at all
2	Data are poorly monitored
3	Data are acceptably monitored
4	Data are well monitored
5	Data are very well monitored

A) Overall standard of water quantity monitoring
Examples include, but are not limited to: streamflow being regularly measured, estimated, or modeled in the basin

☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5

B) Overall standard of water quality monitoring
Examples include, but are not limited to: water quality samples taken from water bodies and measured, or water quality being modeled based on data related to discharge of pollutants.

☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5

C) **Overall standard of biological and ecological monitoring**
Examples include, but are not limited to: surveillance undertaken to assess aquatic species (e.g., harvested, threatened, invasive) populations or communities (e.g. macroinvertebrates).

☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5

D) **Overall standard of monitoring access to, and use of, water**
Examples include, but are not limited to: household surveys administered to estimate the coverage of access to improved water and sanitation sources, or estimates of farmers' groundwater extraction.

☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5

Comprehensive planning and adaptive management (12 of 12)

Comprehensive planning is the process of developing goals and objectives concerning water quantity and quality, surface and groundwater use, land use change, river basin ecology, and multiple stakeholders' needs. Adaptive management refers to the ability to handle changes, unintended consequences, or surprises to the water resource system through updating planning and processes using new information

*Based on your own knowledge of the current situation, please evaluate the degree to which **comprehensive** planning at the basin (or sub-basin) scale is taking place. Provide a rating between 1 and 5 following the criteria below. Please skip any items which you do not feel qualified to answer.*

Rating	**Criteria**
1	Process is almost never comprehensive, or does not occur at all
2	Process is rarely comprehensive
3	Process is sometimes (~50%) comprehensive
4	Process is often comprehensive
5	Process is almost always comprehensive

A) **A shared vision is established and used to set objectives and guide future development**
Examples include, but are not limited to: goals for improvement are jointly established by multiple stakeholders, or a process is in place for developing local water plans that inform higher-level (provincial or national) plans.

☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5

B) **The existence and use of strategic planning mechanisms**
Examples include, but are not limited to: basin-specific spatial plans or management plans that guide investments and policy, or climate change adaptation plans.

☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5

C) **The existence and use of an adaptive management framework**
Examples include, but are not limited to: updating plans to reflect new knowledge or changing economic development priorities, or to address issues such as climate change.

☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5

APPENDIX C: FRESHWATER HEALTH INDEX SCIENTIFIC WORKING GROUP BIOSKETCHES

Dr. Sandy Andelman is executive director of Vital Signs, and chief scientist and senior vice president of the Betty and Gordon Moore Center for Science at Conservation International. She previously served as Deputy Director of the U.S. National Center for Ecological Analysis and Synthesis (NCEAS), one of the world's top ecological research institutes. Her scientific expertise includes tropical ecosystems, biodiversity, climate change and interactions between the environment and human well-being. Sandy has pioneered the creation of global monitoring and forecasting systems for climate change and ecological change — early-warning systems — to recognize and predict thresholds of environmental degradation in time to prevent them and to promote resilient human societies. She has a Ph.D. in behavioral ecology from the University of Washington.

Dr. Chusit Apirumanekul is a research fellow, Stockholm Environment Institute. He is a hydrologist with professional experiences in the field of hydro-meteorology, integrated water resource management, flood disaster risk management, flood modeling, early warning system and capacity building, especially in Lower Mekong Region including Cambodia, Lao PDR, Myanmar, Thailand and Vietnam. He also has experiences in flood and water resources modelling.

Dr. Tim Capon is an agricultural and natural resources economist with CSIRO Land and Water based in Canberra. Tim's research interests include the application of behavioural and experimental economics to understanding the factors that shape decisions and market outcomes. Applications include the design of markets for greenhouse gas emissions and soil carbon sequestration, the design of market-based instruments for natural resource management, and climate change adaptation decision-making. A recent focus of his research is on understanding how a real options decision framework can be used to investigate how uncertainties about future climate affect the adaptation and transformation of agricultural and natural systems.

Dr. Naresh Devineni is an assistant professor in the Department of Civil Engineering and NOAA Cooperative Remote Sensing Science and Technology Center, City University of New York (City College). His areas of expertise are hydro-climate modeling, water sustainability and risk assessment, water systems analysis and extremes analysis, statistical methods for water resources. His work addresses the impacts of climate variability and change on water resources, exploring both floods and droughts, their climate determinants over multiple centuries, and how these may affect interlinked human activities at multiple scales of cities, river basins and nations.

Dr. David Dudgeon is chair professor of Ecology and Biodiversity and Director of the School of Biological Sciences at the University of Hong Kong. He has more than 30 years' experience as a teacher and researcher, and is the author of more than 200 scientific papers, numerous book chapters, and several books on freshwater ecology and biodiversity conservation - mainly dealing with tropical Asia. These include Tropical Asian Streams (1999), The Ecology and Biodiversity of Hong Kong (2005 and 2011; published in English and Chinese), and an edited collection, Tropical Stream Ecology (2008). In 2000, Dudgeon was awarded the Biwako Prize in Ecology by the Japanese government. He is a member of a variety of international advisory boards and scientific committees, and is editor-in-chief of the peer-reviewed journal *Freshwater Biology*.

Dr. Tracy A. Farrell is the regional director for the Greater Mekong Program for Conservation International Cambodia. She has spent the past five years developing and leading cross-cutting initiatives in the areas of fresh water, ecosystem services, and wildlife trade. In this role, she creates research agendas, strategic directions, and business plans to refine CI's niche and partnership approach to address these as well as other emerging institutional priorities. She has published broadly across these and other areas in both peer reviewed and popular publications, and has 10 years' experience aligning research and field activities to ensure solid program delivery, largely taking place in North, Central and South America. Before joining CI, she served as dean for the School for Field Studies and was also a visiting professor/instructor for Virginia Tech's Department of Forestry.

Dr. Isabelle Fauconnier is the water policy and sustainability advisor for the Global Water Programme at the International Union for the Conservation of Nature (IUCN). She has worked on institutional reform and governance of water services provision and water resources management for over 15 years with both multilateral and non-governmental organizations. She has conducted field research, project and policy work in Latin America, Africa and North America. Through her work in both urban slums and rural watersheds, Isabelle has focused on linking poverty, social equity and economic development concerns with the improved management of water services and resources. Before joining IUCN, Isabelle worked on water policy research, project design and evaluation with organizations such as the World Bank, WHO and the World Wildlife Fund, providing technical assistance to the governments of Haiti, Venezuela, Argentina, Ghana, Burundi, Cameroon and Morocco on the design and implementation of institutional changes in the water sector. Isabelle holds a PhD in City and Regional Planning from the University of California at Berkeley.

Dr. Glen MacDonald is a Distinguished Professor and the John Muir Memorial Chair of Geography at UCLA. He works on issues of climate change and its impacts, particularly in terms of water resources and wetland systems. He is a Member of the National Academy of Sciences, a Fellow of the American Association for the Advancement of Science, a Fellow of the American Geophysical Union and a Guggenheim Fellow.

Dr. Matthew McCartney is Theme Leader on Ecosystem Services for the International Water Management Institute, Vientiane, Lao PDR, and he specializes in water resources and wetland and hydro-ecological studies. He has participated in a wide range of research and applied projects, primarily in Africa and Asia, often as part of a multi-disciplinary team. Most recently he worked on a number of projects including: a water resource assessment of the Dry Zone of Myanmar, an evaluation of the flow regulating functions of natural ecosystems in the Mekong and a study on integrating built and natural infrastructure in water resource planning in the Tana and Volta River basins. He was a steering committee member on the UNEP Dams Development Project (2002-2004) and a member of the Ramsar Science and Technical Review Panel (2007-2015).

Dr. Amy McNally is an assistant research scientist at the Hydrological Science Laboratory, NASA Goddard and UMD ESSIC studying water resource availability in sub-Saharan Africa and Yemen for the Famine Early Warning Systems Network. Using remotely sensed data and land surface models, her research focuses on improving estimates of soil moisture and evapotranspiration for agricultural drought and water resources monitoring. She received a B.S. in Environmental Biology at SUNY-ESF, and an M.S. in Water Policy and Management from Oregon State University where her research focused on climate change and water sharing agreements in the Middle East. She went on to earn a Ph.D. in Geography at the University of California Santa Barbara. Other research highlights have included studies on malaria and climate change in Africa, the socio-economic impacts of dams in Southwest China, and the impact of aerosols on precipitation in South Korea.

Dr. Cho Nam Ng is an associate professor in the department of geography at The University of Hong Kong. His expertise lies in environmental policy and planning, environmental impact assessment and strategic environmental assessment, nature and heritage conservation, sustainable development, and climate change and energy policy. He received his Ph.D. from Lancaster University.

Dr. Alison (Sunny) Power is a professor in the Department of Ecology and Evolutionary Biology and the Department of Science and Technology Studies. Her research focuses on biodiversity conservation in managed ecosystems, interactions between agricultural and natural ecosystems, agroecology, the ecology and evolution of plant pathogens, invasive species, and tropical ecology. She has led a working group on the roles of natural enemies and mutualists in plant invasions at the National Center for Ecological Analysis and Synthesis. She served as vice-president for public affairs for the Ecological Society of America and as the presidential university fellow of The Nature Conservancy. She served on the Committee on California Agricultural Research Priorities of the National Research Council and the Oversight Committee of the Collaborative Crop Research Program of the McKnight Foundation.

Dr. Helen Regan is a Professor in the Biology Department at the University of California Riverside. She received Bachelor of Science and PhD degrees in applied mathematics. Her research interests are diverse, interdisciplinary and highly collaborative, spanning risk analysis, ecological modeling for global change, decision making and uncertainty analysis particularly in the realm of conservation. She serves on the IUCN Red List Standards and Petitions Subcommittee and on the editorial boards of the journals Ecology Letters and Diversity and Distributions. She was a member of the National Research Council's Committee on the Independent Scientific Review of the Everglades Restoration Progress for four years.

Dr. Kashif Shaad is a postdoc, based at Conservation International Singapore, helped develop the data resources and modeling approaches for the Freshwater Health Index. He recently completed his PhD in Environmental Engineering from ETH Zurich and holds a Masters in Hydroinformatics. His research interests include developing mathematical models and informatics tools for improving water management, and he is keenly following the growing integration of ecology with hydrodynamics.

Dr. Rebecca Shaw is Chief Scientist and Senior Vice President at the World Wildlife Foundation. Previously she worked at the Environmental Defense Fund, where she was responsible for developing and implementing the vision and strategy of the Land, Water and Wildlife program. Prior to joining EDF in 2011, she served first as Director of Conservation Science and then as Associate State Director at the Nature Conservancy's California Chapter. She's also researched the impact of climate change at the Carnegie Institution for Science's Department of Global Ecology at Stanford University. She is a lead author of the section of the 2014 Intergovernmental Panel on Climate Change's Fifth Assessment Report that focuses on impacts, adaptation, and vulnerability, and serves as a member of the California Climate Adaptation Advisory Panel. Rebecca holds an M.A. in environmental policy and a Ph.D. in energy and resources from the University of California, Berkeley.

Dr. Nicholas Souter is the Mekong case Study Manager at Conservation International Cambodia. His areas of expertise are in conservation biology and natural resource management. He has worked extensively on determining the impacts of river regulation on floodplain vegetation dynamics and processes in the South Australian lower River Murray. He represented South Australia on the Murray-Darling Basin Authorities Sustainable Rivers Audit Implementation Working group and was a technical member of the Vegetation assessment group. He has spent the last three years in Cambodia managing Fauna and Flora International's University Capacity Building Project in partnership with the Royal University of Phnom Penh.

Dr. Caroline Sullivan is a Professor of Environmental Economics and Policy at the Southern Cross University, NSW Australia. Her areas of expertise span: Water Management, International Development, Ecological and Environmental Economics, Index development. She has been involved in water and forestry research for over 20 years in Asia, Africa, Europe, the Caribbean, Latin America, Australia and the Pacific. She conceived of and led the work on the development of the Water Poverty Index and has worked on the development of a variety of indices with several organizations including the FAO, the governments of Canada, the UK, Fiji and others, and the African Development Bank.

Dr. Derek Vollmer is a Postdoctoral Researcher in the Betty and Gordon Moore Center for Science, where he is helping develop and apply the Freshwater Health Index in select river basins around the world. Prior to joining CI, he worked as a Doctoral Researcher within the Future Cities Laboratory at the Singapore-ETH Centre for Global Environmental Sustainability. His research there focused on ecosystem services and spatial planning, with a focus on the Ciliwung River catchment in metropolitan Jakarta. Dr. Vollmer has also been a Program Officer in the Science and Technology for Sustainability unit of the U.S. National Academy of Sciences in Washington, DC, where he directed two bilateral studies on U.S.-Chinese cooperation on clean energy, along with studies of multi-stakeholder partnerships, product certification schemes and urban sustainability issues. He holds a B.A. in Government and International Studies from the University of Notre Dame, a M.S. in Environmental Science and Policy from Johns Hopkins University, and a PhD in Spatial Planning from ETH Zurich, Switzerland.

Dr. Raymond Yu Wang is an Associate Professor in School of Government, Sun Yat-sen University. He holds a Ph.D. in Geography from The University of Hong Kong, where he continued post-doctoral research in Faculty of Social Sciences. His main areas of expertise include water governance, environmental policy and environmental politics in China.