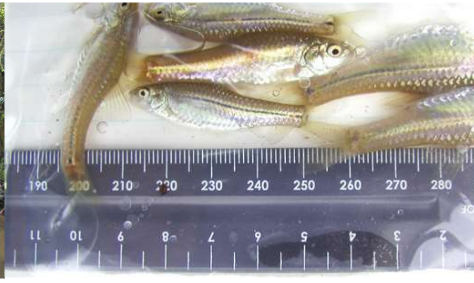





IWMI International Water Management Institute

CGIAR RESEARCH PROGRAM ON Water, Land and Ecosystems

IFPRI INTERNATIONAL FOOD POLICY RESEARCH INSTITUTE

WWF

Towards a Global River Health Assessment Framework

Jeremy Dickens, Chris Dickens, Nishadi Eriyagama, Hua Xie and David Tickner

Ayeyarwady River in Myanmar

Report citation: Dickens, J.; Dickens, C.; Eriyagama, N.; Xie, H.; Tickner, D. 2022. *Towards a global river health assessment framework*. Project report submitted to the CGIAR Research Program on Water, Land and Ecosystems (WLE). Colombo, Sri Lanka: International Water Management Institute (IWMI). 131p. doi: https://doi.org/10.5337/2022.224

The authors: Jeremy Dickens (private) undertook the literature review and drafted much of the text; Chris Dickens (IWMI) was project leader, drafted text and produced the final document; Nishadi Eriyagama (IWMI) drafted text on hydrology; Hua Xie (IFPRI) drafted text on data; and David Tickner (WWF) provided oversight and review comment.

About IWMI

The International Water Management Institute (IWMI) is an international, research-for-development organization that works with governments, civil society and the private sector to solve water problems in developing countries and scale up solutions. Through partnership, IWMI combines research on the sustainable use of water and land resources, knowledge services and products with capacity strengthening, dialogue and policy analysis to support implementation of water management solutions for agriculture, ecosystems, climate change and inclusive economic growth. Headquartered in Colombo, Sri Lanka, IWMI is a CGIAR Research Center with offices in 14 countries and a global network of scientists operating in more than 30 countries.

About WLE

The CGIAR Research Program on Water, Land and Ecosystems (WLE) is a global research-for-development program connecting partners to deliver sustainable agriculture solutions that enhance our natural resources – and the lives of people that rely on them. WLE brings together 11 CGIAR centers, the Food and Agriculture Organization of the United Nations (FAO), the RUAF Global Partnership, and national, regional and international partners to deliver solutions that change agriculture from a driver of environmental degradation to part of the solution. WLE is led by the International Water Management Institute (IWMI) and partners as part of CGIAR, a global research partnership for a food-secure future.

Disclaimer*: This publication has been prepared with care by the authors. Responsibility for editing, proofreading and design/layout of the document, and for any remaining errors and opinions expressed lies with the authors and not the institutions involved. The boundaries and names shown and the designations used on maps do not imply official endorsement or acceptance by IWMI.*

International Water Management Institute (IWMI)

127 Sunil Mawatha, Pelawatta, Battaramulla

Mailing address: **P O Box 2075, Colombo, Sri Lanka**

www. Iwmi.org

Acknowledgements

The authors would like to thank the CGIAR Research Program on Water, Land and Ecosystems (WLE) and World Wide Fund for Nature (WWF) for providing financial support to draft this report. We would also like to thank the 60 experts who joined online workshops to discuss this work and the way forward, in particular those who shared presentations with the meeting, i.e., Stuart Crane, Hazel Thornton, Maria Nuutinen, Paul Franklin and Mike Morris.

Collaborators

International Water Management Institute (IWMI), Colombo, Sri Lanka

International Food Policy Research Institute (IFPRI), Washington, DC, USA

World Wide Fund for Nature (WWF), UK

Donors

CGIAR Research Program on Water, Land and Ecosystems (WLE)

World Wide Fund for Nature (WWF), UK

Benthic macroinvertebrates from a biomonitoring sample showing moderately healthy (left) and poor health (right) (Picture C. Dickens)

SUMMARY

Maintaining the state or health of rivers is a vital part of sustainable development. Healthy rivers are able to support and maintain key ecological processes and thus ecosystem services on which society depends. The monitoring and reporting of river health is well documented at a local scale, but at national, regional and global scales the existing methods are generally unable to provide useful information. This has resulted in a dearth of river health data and information in the likes of the SDGs, the CBD Post-2020 framework, the IPBES etc., reflecting a world-wide lack of appreciation of this vital natural resource.

This report seeks to understand the present global situation of river health (RH) assessment and reporting to develop an understanding of the key attributes of successful approaches, especially in relation to the Emergency Recovery Plan for freshwater biodiversity and GBO5 Sustainable Freshwater Transition (both published in 2019). We restrict our selection of reviewed frameworks to those applicable at either the regional (multi-basin or multi-national) or global scales, as we considered these most likely to possess the traits necessary for global expansion. Where a framework is proposed with the intention to supplant past protocols, the most modern version is used. The report also considers potentially novel ways of moving data to the global scale and ends with the requirements for a global framework for monitoring and reporting on river health. In the process it does not provide detail of site-based methods even where these can be incorporated into global frameworks, as there are already several reviews of such methods.

This report focusses on river health and recognizes this as a part of aquatic ecosystem health. While rivers are by nature integrated with floodplain and some palustrine wetlands, with groundwater and with estuaries, the focus remains on the freshwater lentic (flowing) systems. The principles elucidated will however apply to many of the other aquatic ecosystems.

The key issue that needs to be resolved right at the outset of the design of a RH framework, is just what to include in the definition of river health (RH). While several of the frameworks reported have included human values in their definitions, which has the advantage of promoting the concept of RH into society, it is suggested that this dilutes the urgent need to reflect the state of the very resource itself, which must be based on the biophysical character of the ecosystem alone. If there is no knowledge of the state of the resource itself, then understanding the role of this resource in society is meaningless. It is proposed that these are two different considerations, first the state of the river ecosystem itself, and secondly the relationship to society. For this RH framework, it is proposed that it is the ecosystem alone that is of relevance, and that this indicator can then be used to provide a second layer contribution to other indicators that include the human perspective. This approach is however open to debate and could change for a future framework, but it is an important issue that will need to be resolved before a final framework can be recommended or adopted.

The definition embraced in this report as the most representative of the objectives of RH monitoring and reporting is thus ***"the ability of the river ecosystem to support and maintain key ecological processes and a community of organisms with a species composition, diversity and functional organisation as comparable as possible to that of natural habitats within a region"***.

A number of regional and global RH frameworks were reviewed and have contributed to this report. Regional frameworks included those of the European Water Framework Directive, the USA, Mekong basin, and country based frameworks from South Africa, China, Australia and New Zealand. That from New Zealand in particular has provided substantial material that has potential to guide a global

framework. The value of these regional and country frameworks is that they demonstrate RH monitoring in action, although not always achieving the kind of ecological reporting that we would hope to achieve with this global framework.

The global frameworks included the SDGs, the CBD Post2020 Biodiversity Framework, Group of Earth Observation Biodiversity Observation Network (GEO BON), the Freshwater Health Index, Planetary Boundaries, the Incident Threat Index, the System for Environmental Economic Accounting (SEEA), the Environmental Performance Index (EPI) and the recent Emergency Recovery Plan (ERP). These global frameworks provide useful approaches that could be adopted for a global RH framework even though they themselves did not include any or sufficient RH or river ecosystem data.

Also included in the review are several successful global indices that have a global reach, even though in all cases they did not meet the requirements of RH monitoring that are the objective here. These included the IUCN Red List Index, the Biodiversity Intactness Index, the Living Planet Index, the Water Footprint, the Connectivity Status Index, the Canadian Water Quality Index, the Global Water Quality Index, the SDG 6.6.1 indicator, the Mekong River Commission Water Quality Index and the Aquastat index. While none of these indices were successfully reporting on RH at any level of confidence, they all provide approaches that are of value for a global RH framework.

The lessons learned from the above frameworks are many, however just how these will be sifted out and incorporated into a future framework remains the task of the next step in the process.

The following were recognised as the key attributes of a successful framework:

- **Consistency** - understanding of what constitutes ecosystem health and how to measure it
- **Representativeness** - includes measurement of a full range of the core components of ecosystem health
- **Robustness** - rigorous science with justified selection of components and indicator variables based on empirical evidence
- **Informativeness** - easily understood
- **Flexibility** - can be meaningfully applied across a wide range of waterbodies
- **Scalability** - application remains consistent across spatial scales
- **Feasibility** - not highly demanding on time, labour or money

Adapted from Clapcott *et al.*, 2018

The following are key characteristics and approaches that enabled existing successful frameworks:

- **Policy driven purpose** - A clear purpose is the foundational element of any framework as it influences decisions to all subsequent aspects of the framework development, from the definition of terms, choice of data acquisition methods, to processing and reporting.
- **Clear and consistent definitions** – the essential definition of RH is shown above, however, any new program would need to be clear on which definition to embrace.
- **Using conceptual models to direct the program** – with an overarching model such as the DPSIR, having a conceptual model helps to ensure that the program is fit for purpose and also that the results are not mis-represented.
- **Clear consideration of the key components** of a RH framework – there is clear separation of some frameworks that are based on ecological components and others that include socio-economic aspects. These would have different purposes.
- **RH indicator types** – some frameworks monitor the ecosystem directly while others make use of proxies where there is a risk that interpretation will be misleading.

- **Processing of data** - data processing usually involves three steps. 1) The aggregation of raw data to the appropriate scale for each metric. 2) Data are standardised to a common scale, to ensure consistency and flexibility. This often involves comparison to reference data. 3) The integration (or combination) of data at the indicator, component, or overall ecological condition levels for reporting.
- **Reporting of results** – reporting needs to consider the scale of the report, and also how to best communicate the data that is being used. Many formats for such reports have been produced, with perhaps the most useful being circulate pie-charts that provide an integrated score but also allow component scores to aid interpretation.

This report provides a baseline to formulate a global RH framework and will be followed up with further steps. Supported by IWMI and funded by the WWF, more conceptual thinking will be done to draft a paper that proposes a way forward for a global RH framework. It is hoped that this will provide a springboard for a renewed research effort to develop a global framework that can be adopted at a global level, feeding into global reports such as the SDGs, the CBD and many others.

CONTENTS

LIST OF FIGURES

1 INTRODUCTION

1.1 Rivers as aquatic ecosystems

Rivers are aquatic or freshwater-related ecosystems, a term which may also include wetlands, swamps, bogs, peat, paddies, wadis, streams, canals, lakes, reservoirs, groundwater aquifers, estuaries, and mangroves. The following definition for aquatic ecosystems is adapted from the Millennium Ecosystem Assessment (MEA, 2005) by Dickens and McCartney (2021): *"A water-related ecosystem is a dynamic complex of plant, animal, and microorganism communities and the non-living environment dominated by the presence of flowing (lotic) or still (lentic) water, interacting as a functional unit."*

This report focusses on rivers (lotic ecosystems) and recognises that they are a part of freshwater aquatic ecosystems.

1.2 Ecosystem and river health

The health of anything has an intrinsic understanding to all of us, and when applied to ecosystems the concept is the same. As with the health of a human or any other system, health is a concept that holistically integrates a wide array of issues. Aquatic ecosystem health has been defined as: '*The ability of the aquatic ecosystem to support and maintain key ecological processes and a community of organisms with a species composition, diversity, and functional organisation as comparable as possible to that of undisturbed habitats within the region*' (Schofield and Davies, 1996) after Karr and Dudley 1981:55-68]. Burkhard et al. (2008) noted that ecosystem health is a concept that integrates environmental conditions with the impacts of anthropogenic activities to give information for the sustainable use and management of natural resources. The concept is similar to that of the "state" of an ecosystem i.e. the particular condition of the ecosystem at that time, a concept used to great effect in the DPSIR framework (Smeets and Weterings, 1999) which documents the causal drivers (e.g. economic development) and pressures (e.g. pollution, water abstraction) that result in a particular state of the ecosystem (e.g. the state of the fish, invertebrates etc), and then the resulting impacts (e.g. loss of benefits) and responses (e.g. government policy to redress the problem).

It is the nature of a health assessment that if only one aspect is damaged, then it is likely that the overall health of the system would suffer. It is for this reason that the Water Framework Directive of the EU prescribes a "one-out all-out" policy in determining the state of rivers, meaning that if the river fails in one respect, then it fails overall.

In recent years, rapid economic growth, irrigation, urbanization, population increase, and infrastructure development have had profound impacts on water-related ecosystems and the benefits they provide for people. When considering river discharge (flows of water), withdrawal has decreased the flows in many rivers, exacerbated by dam construction, which means less water, less or changed water-related habitat and, consequently, changes in the portfolios of ecosystem services that rely on the presence of water. Also, the changes in flow amplitude as a result of dams leads to further alterations to ecosystem services, while degradation caused by pollution, loss of assimilative capacity in terms of water quality, higher temperatures, invasion of alien species, fragmentation of river ecosystems by dams, altered flows, unsustainable harvesting of fish and other organisms, and many other anthropogenic stressors also compound the problem and are major threats to the health of water-related ecosystems and their services. The alteration of river flows is often regarded to be the most serious and continuing threat to the ecological sustainability of rivers globally (Lundqvist 1998; Bunn and Arthington 2002).

1.3 Monitoring and reporting on river health at scale

Monitoring the health or state of aquatic ecosystems has had a long history and the practice has become entrenched in many countries around the world. In its original form, monitoring was mostly carried out *in situ*, with assessments of habitat and biota often integrated to provide a holistic picture of the health of the river, lake and/or wetland site. Methods that have been used encompass drivers of the ecosystem health such as hydrology, water quality, geomorphology, sediments; river habitats; riparian vegetation and instream plants including periphyton, diatoms; fish; benthic macroinvertebrates etc. These indicators are either used independently to represent the health of the whole ecosystem, or alternately a suite of these indicators is combined to more holistically represent the health of the ecosystem *(Note: greater detail is given later in the report)*.

Freshwater ecosystems including rivers are in trouble across the globe as their health deteriorates. The global Millennium Ecosystems Assessment (MEA 2005) concluded after a comprehensive survey of data that there has been a "substantial and largely irreversible loss in the diversity of life on Earth" and that ecosystems (including aquatic ecosystems) "are being degraded or used unsustainably". More recently, regional assessments conducted by the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services (IPBES 2018) for the African region, as an example, highlighted that all major ecosystems are threatened, with wetlands being the most threatened of all ecosystems. The lack of progress to resolve the decline in ecosystems is graphically illustrated by the general failure to meet the Aichi Biodiversity Targets of the Convention on Biological Diversity (CBD), where it has been shown, again by way of example, that more than 50% of the countries in Africa would not meet these targets (UNEP-WCMC 2016).

In an attempt to reverse the decline, major global initiatives have now thrust ecosystems onto the global development agenda that includes the following: Millennium Ecosystem Assessment (MEA 2005), The Economics of Ecosystems and Biodiversity (TEEB) for water and wetlands (Russi et al. 2013), the World Water Development Report (WWDR) on nature based solutions (WWAP 2018), the IPBES, the Framework for Freshwater Ecosystem Management approved by the UN Environment Assembly (UNEP 2017), the Ramsar Convention, Agenda 2030 on Sustainable Development (the SDGs), the Post-2020 Framework of the Convention on Biological Diversity and now the UN Decade on Ecosystem Restoration. All of these, together with globally impactful periodic reports such as the WWF Living Planet Report and the Ecological Footprint indicator, and the Water Footprint indicator, are all lacking an adequate global representation of aquatic ecosystem health and they have generally had to rely on proxies. The recent high-level motivation to the CBD of an Emergency Recovery Plan for Freshwater Biodiversity (Tickner, et al. 2020) and GBO5 Sustainable Freshwater Transition of the CBD call for indicators of biodiversity conservation and restoration and would also be supported by measures of aquatic ecosystem health. At a global scale the shortage of biodiversity data has become a pressing issue as it prevents proper reporting on the state of ecosystems. There are however two sources of freshwater biodiversity. http://www.freshwaterplatform.eu/ a part of the BIOFRESH project to document aquatic biodiversity but limited to Europe, and https://www.gbif.org/ a global repository of biodiversity data and information - both have connections to a wide range of biodiversity data and information including aquatic but neither yet provide data to a level that it would be useful for periodic reporting on the state of river health.

Over the past few decades, the need to represent river health at a national and ultimately at a global scale has become pressing. As an approach, site-based data has been often been used to represent the state of aquatic ecosystems even at a basin scale, yet the challenge for global reporting is that the roots of river health monitoring are at an *in situ* scale, with relevant global datasets seemingly scarce and contributing little. In the SDGs reporting process, indicator 6.6.1 (*Change in the extent of water related ecosystems over time*) was simplified to use proxies of ecosystem health (spatial extent, basic water quality, river discharge) to indicate the ecosystem state. Initial attempts to include the direct assessment of ecosystem health were rejected as there were no recognized methods available that could be applied at a global level.

Many countries and regions have existing programmes that monitor aquatic ecosystem health, but these differ around the world to the extent that any attempt to gather harmonized global data is difficult. Moving to the future, river health could be determined using a variety of different approaches ranging from top-down approaches that for example use Earth Observation data, or measures of the status of land-based stressors predicting the impact on river health, to conversely bottom-up approaches that focus on the instream biota and use conventional field monitoring and/or citizen science methods for data gathering. Each of these have their advantages but not all are suited to global scale reporting. Data from new sources, including cutting edge technologies (such as EO, eDNA) could help improve global monitoring of progress towards international goals, such as the SDGs (Hsu *et al.*, 2016).

1.4 This Report

This report focusses on river health and recognizes this as a part of aquatic ecosystem health. While rivers are by nature integrated with floodplain and some palustrine wetlands, with groundwater and with estuaries, the focus remains on the freshwater lentic (flowing) systems. The principles elucidated will however apply to many of the other aquatic ecosystems.

This report seeks to understand the present global situation of river health monitoring to develop an understanding of the key attributes and approaches of successful frameworks, especially in relation to the Emergency Recovery Plan for freshwater biodiversity and GBO5 Sustainable Freshwater Transition (both published in 2020). We restrict our selection of reviewed frameworks to those applicable at either the regional (multi-basin or multi-national) or global scales, as we considered these most likely to possess the traits necessary to inform a global monitoring approach. Where a framework is proposed with the intention to supplant past protocols, the most modern version is used. The report also considers potentially novel ways of moving data to the global scale and ends with the requirements for a global framework for monitoring and reporting on river health. In the process it does not provide detail of site-based methods even where these can become part of integrated indices, as there are already reviews of such methods that are not listed here while the review of modern concepts of river health are scarce (Boulton, 1999); Fairweather, 1999, etc.).

Table 1.1 provides a summary of the regional and global frameworks for some component of RH that have contributed to this review, and also some of the large-scale indices that are available. Not all of these are comprehensive in their scope by documenting RH in its entirety, but they all could contribute something useful to the design of a global framework for RH reporting.

Table 1.1 Summary of frameworks and indices reviewed in this report

Acronym	Framework name	Scale	Reference
	Regional Frameworks		
WFD	Water Framework Directive	European Union	CEC, 2000

Acronym	Framework name	Scale	Reference
NARS	National Aquatic Resource Surveys	USA	USEPA 2006, 2020a
REMP	River EcoStatus Monitoring Programme	South Africa	Dallas et al, 2008; Kleynhans et al. (2008).
RHI	River Health Index	China	Xie et al., 2020
NRHP	National River Health Programme	Australia	FBM and DNRM, 2001; Halse et al., 2002
SRA	Murray-Darling Basin Sustainable Rivers Audit	Australia	Davies, et al. 2010, Whittington et al. (2001)
IECA	Integrated Ecosystem Condition Assessment	Australia	Department of the Environment and Energy, 2017
FBEHF	Freshwater Biophysical Ecosystem Health Framework	New Zealand	Clapcott et al., 2018
MIF	Mekong River Basin Indicator Framework	Mekong basin	MRC, 2019c
Global Frameworks			
SDG 6	Sustainable Development Goals	Global	UN, 2015; UN Water, 2017
CDB post-2020	Convention on Biological Diversity Post-2020 Biodiversity Framework	Global	CBD, no date
GEO BON	Group of Earth Observation Biodiversity Observation Network	Global	Scholes et al. (2012); GEO BON (2015)
FHI	Freshwater Health Index	Global	Vollmer et al., 2018
PB	Planetary Boundaries	Global	Rockström et al., 2009; Steffen et al., 2015
ITI	Incident Threat Index	Global	Vörösmarty et al., 2010
SEEA	System for Environmental Economic Accounting for Water	Global	UNDESA 2012
EPI	Environmental Performance Index	Global	Hsu et al., 2016
ERP	Emergency Recovery Plan	Global	Tickner et al., 2020
Successful indices with global reach			
RLI	IUCN Red List Index		IUCN, 2021
BII	Biodiversity Intactness Index		Scholes and Biggs, 2005
LPI	Living Planet Index		WWF, 2020a
WF	Water Footprint		Hoekstra et al., 2011
CSI	Connectivity Status Index		Grill et al., 2019
WQI	Canadian Water Quality Index		CCME, 2002
WATQI	Global Water Quality Index		Hsu et al., 2016
SDG 6.6.1	Sustainable Development Goal 6.6.1		UN SDG Indicators Repository (no date)
WQI	Mekong River Commission Water Quality Index		MRC, 2019a
Aquasat	Aquasat		Ross et al., 2019

2 REVIEW OF EXISTING METHODS FOR RIVER HEALTH ASSESSMENT

Below we review some of the most widely used or most recently developed frameworks for river health assessment at the regional and global scales. Table 1.1 provides a summary of all of the frameworks and indices that have been included in the review. For each, we summarise its benefits and shortfalls and the lessons that can be learned in relation to its applicability to a global approach.

2.1 Review of regional or national frameworks

2.1.1.1 The Water Framework Directive (WFD), EU

The European Union's Water Framework Directive (CEC, 2000) was one of the world's most ambitious pieces of environmental legislation at the time it was written and set the precedent for subsequent frameworks. It was established to streamline legislation to assist with water pricing across the EU and outlines water policy and objectives for EU member states. The purpose of the Directive is to establish a framework to prevent further deterioration and enhance protection and improvement of aquatic ecosystems (surface, ground, transitional and coastal waters), whilst promoting sustainable water use and maintaining socioeconomic systems (Article 1). For all inland surface waters (rivers & lakes), both natural and artificial, the objective for member states was thus to achieve at least 'good' ecological status (or good ecological potential for heavily modified water bodies) by 2015 and at the latest 2027 at the scale of individual river basins (Article 3). The Directive then provides a framework for the characterisation of river basins and monitoring and classification of ecological status, but it is the choice of Member States as to the exact methods and measures employed. Firstly, ecological quality is considered to comprise three components, namely:

- Biological elements, including the composition and abundance of aquatic flora, benthic invertebrates and fish, and phytoplankton for lakes.
- Hydro-morphological elements supporting biological elements, including water flow quantity & dynamics, residence time (lakes), groundwater connection, river continuity, depth, width variation (rivers), bed structure and substrate, bed quantity (lakes), and riparian zone/ lakeshore structure].
- Physico-chemical elements supporting biological elements, including thermal, oxygenation, salinity, acidification, & nutrient conditions, transparency (lakes), and specific pollutants relevant to the locality].

For each of these elements, member states are required to establish reference conditions by river basin type (differentiated by ecoregion, then type (altitude, size and geology) or a similar degree of differentiation). Reference conditions are also established for heavily modified or artificial waterbodies. Within each river basin, they are required to identify pressures and assess any impacts on the ecological status of water bodies (Article 5) and set up monitoring programmes to establish coherent and comprehensive overviews of the status within each (Article 8). The framework provides guidance on what indicator variables must be included for comprehensive monitoring of each element of ecological quality. Ecological status is then graded as an ecological quality ratio, based on the degree of change from the reference conditions, and divided into five classes: 'high', 'good', 'moderate', 'poor' or 'bad' with standard definitions (see Annex V) and colour coding (see below). The integration of the quality ratio for each variable into elements, and elements into a single value of ecological status has subsequently mostly used the average or 'one out, all out' approaches (Hering *et al.*, 2010). Given that implementation of the WFD varies by country, and given that many European water bodies cross national boundaries, the European Commission has paid a lot of attention to inter-calibration of ecological status categories between member states, particularly standardisation of reference conditions.

The greatest advantage of the WFD is that it provides a means of integrating information from existing monitoring schemes that differ in the types of data considered, sampling designs and geographical context, by providing standardised protocols for data acquisition and analysis. This has enabled transboundary cooperation in the development of river-basin management plans and resulted in comprehensive basin-wide pictures of ecological status. The strength of the standardised approaches for data collection (i.e., what elements must be measured, not how) and means of data analysis (standardisation) are breakthroughs in this regard, removing the need for complicated calibration between countries. However, the greatest criticisms of the WFD are that the timeframe proposed for countries to achieve good ecological status were unrealistic, more time is needed to collect representative data, that it lacks functional indicators, and that the 'one out, all out' approach to integrate indicators is too restrictive (Hering *et al.*, 2010).

The greatest lesson learned for a global RH assessment is that a standardised protocol is extremely useful for integrating data from various sources, especially where monitoring systems already exist, as it is highly flexible and robust, and would be highly beneficial to emulate. A global framework must be adaptive to allow the inclusion of new indicator variables and changes in methods for integrating data, which is provided for by a standardised protocol. The need for long-term goals is also clear, as it takes time to collect representative data, especially involving new approaches requiring validation.

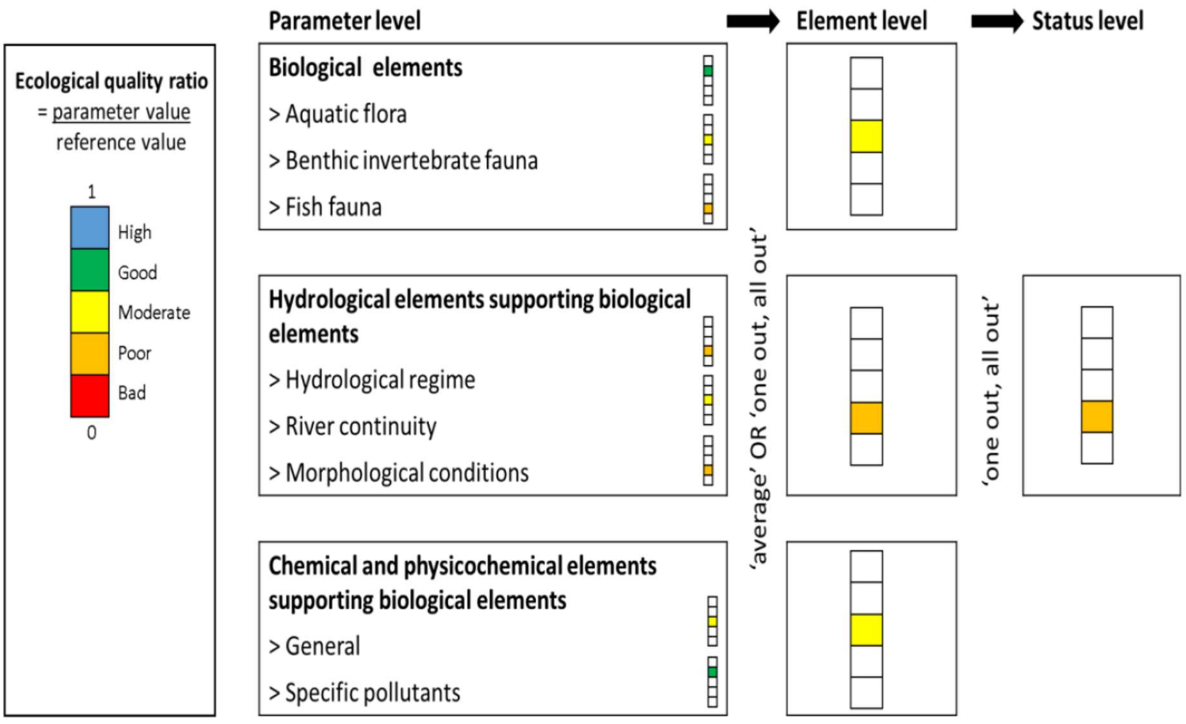


Figure 2.1: Example of how indicator 'ecological quality ratios' are calculated and combined to estimate quality elements in the Water Framework Directive (Van de Bund and Solimini, 2007; Clapcott et al.*, 2018).*

2.1.1.2 National Aquatic Resource Surveys (NARS), USA

The National Aquatic Resource Surveys (NARS) are collaborative programmes between the Environmental Protection Agency (EPA), states and tribes to assess the health of aquatic resources in the USA (USEPA, 2006, 2020a). Of interest to us are the 1) National Rivers and Streams Assessment (NRSA), and National Lakes Assessments (NLA), but protocols also exist for wetlands and coastal waters. The objective is to measure the extent to which rivers and streams and lakes support healthy biological conditions and recreation (for lakes), the extent of major environmental stressors and if there are trends in the conditions present, at the national and ecoregion levels. The survey uses a probability-based sampling design to randomly select sampling sites across the country that reflect the full range in character and variation of rivers and streams and lakes in the US so include weighting by size (as stream order) and spatial distribution (ecoregions) to ensure representative results. Surveys are repeated every 5 years to monitor trends. As with the WFD, surveys involve standardised methods of data collection and analysis for comparability with data collected, which varies between states and at the local level in some place. The biological condition is based on a direct measure of aquatic life and is considered representative of the ecosystem conditions. Biological indicators include fish, benthic macro-invertebrates & periphyton for rivers and streams, although only macro-invertebrates are considered in the first assessment (USEPA, 2006), in addition to zooplankton and Chlorophyll *a* (as a measure of eutrophication) for lakes. Environmental stressors are chemical and physical factors that can negatively influence biological conditions. Chemical indicators include, phosphorous & nitrogen concentrations and acidification, for rivers, streams and lakes, in addition to salinity for rivers and streams, and atrazine, dissolved oxygen, and sediment mercury for lakes. Physical indicators include streambed sediments, instream fish habitat, riparian vegetative cover, and riparian disturbance for rivers and streams and drawdown, human disturbance, lakeshore habitat, physical habitat complexity, and shallow water habitat for lakes. Recreation conditions are determined by factors related to human health, including *Enterococci* (a faecal indicator), mercury in fish tissue, and algal toxins (microcystin), in addition to cyanobacteria in lakes). For each indicator, the condition is defined as “Good”, “Fair”, or “Poor” based on its value relative to the percentile thresholds of the distribution of values for the least-disturbed reference sites. Sites with missing data are included as “Not Assessed”. For stressors, values are given for their extent (percentage of river length) and relative risk (i.e. risk of “poor” biological condition when a stressor is poor).

The greatest **advantages** of the NARS framework are that it uses widely used indicators to integrate monitoring methods across various regions (states etc), providing a comprehensive view of freshwater health across the USA at the basin-scale, whilst the standardisation of variables to a common scale enables comparability across different ecoregions. The randomised sites selection also provides statistically robust results. **Weaknesses** include the long time-intervals (5 years) between assessments, which is a result of the high sampling effort and cost required, meaning that the results become less useful to direct management activities. The design is also highly rigid and not scale-independent (not allowing application at smaller scales) because of the probability-based sampling design. The selection of widely used indicators may also compromise the representativeness of the data by overlooking variables relevant to certain areas. The classification of conditions relative to percentile thresholds in the distribution of the reference condition is also rather arbitrary.

The key **lessons** to be learned from the NARS framework for a global RH assessment are that, like the WFD, standardisation of variables is key for widespread applicability and a standardised protocol is essential to enable consolidation of data from various existing surveys. However, the probability-based sampling design is much too rigid for applicability at multiple scales, especially below the river basin scale (although they could still foreseeably be used within individual regions under a global assessment), and is also expensive because of access limitations to randomly selected sites. The definition of 'ecological conditions' relative only to biological conditions is also more restrictive than most the other frameworks reviewed and does not provide a single score integrative of all aspects of the ecosystem. However, the reporting of results for indicators separately has its value in enabling the interpretation of biological conditions relative to chemical or physical conditions.

2.1.1.3 River EcoStatus Monitoring Programme (REMP), South Africa

The REMP (previously the River Health Programme RHP) embraces a tiered approach to assessment based on input effort. The most basic is the Present Ecological State and Ecological Importance and Sensitivity (PESEIS) (Department of Water and Sanitation, 2014) assessment that is a first broad assessment that has been applied for the entire country at a fine scale, where the Present Ecological State (PES) represents the broadest desktop evaluation of river reaches. This is followed by the EcoStatus approach (that in turn has a tiered approach) summarised in a River Health Programme Manual (Dallas et al, 2008) that references several Ecological Classification manuals that include the *Hydrological Driver Assessment Index* (HAI), *Geomorphological Driver Assessment Index* (GAI) (Rowntree, 2013), *Physico-Chemical Driver Assessment Index* (PAI) (Department of Water Affairs and Forestry, 2008), *Fish Response Assessment Index* (FRAI) (Kleynhans, 2008), *Macro-Invertebrate Response Assessment Index* (MIRAI) (Thirion, 2007), *Riparian Vegetation Response Assessment Index* (VEGRAI) (Kleynhans, Mackenzie and Louw, 2008), and *Index of Habitat Integrity* (IHI) (Kleynhans, 1996; (Kleynhans, Louw and Graham, 2009a, 2009b). These evaluate the present ecological conditions (PES) of various biophysical attributes (or components) of riverine ecosystems in relation to their natural conditions.

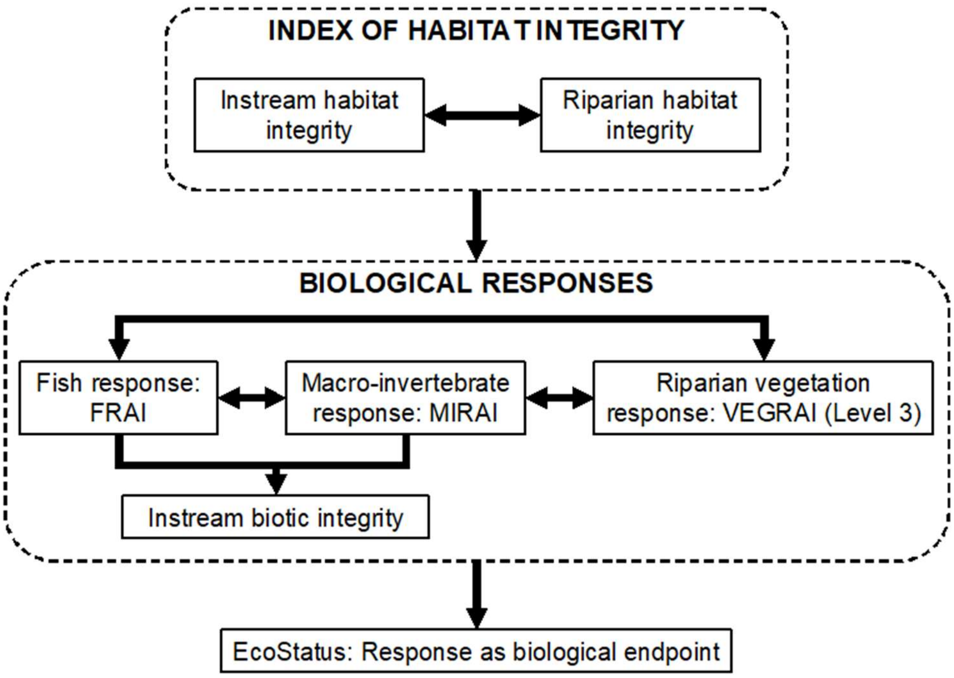


Figure 2.2 The relationship between the indices used to determine river ecological health according to the EcoStatus procedure (from Dallas et al 2008)

The EcoStatus is the ecologically integrated state representing both driver attributes (hydrology, geomorphology and physico-chemistry) and response attributes (fish, invertebrates and habitat) of river ecosystems. Ecological Categories (ECs) of attributes and EcoStatus are represented on a scale from A (Natural) to F (critically modified) and represented visually on the colour spectrum from blue (A) to red or black (F). For each component, indices are calculated using rule-based modelling approaches. These are based on expert ratings of the degree of change from natural, from 0 (no change) to 5 (maximum relative change), for measures relevant to each. Measures are then weighted in terms of their importance to determining the EC under natural conditions for the specific river reach. Depending on the level of detail required and time and funding constraints, they also provide rules for determining EcoStatus at six different levels of assessment varying from rapid desktop to comprehensive. All require the assessment of biotic responses (i.e., fish, invertebrates and riparian vegetation, although the detail of analysis is varied), whilst the more rapid assessments replace more comprehensive driver indices (HAI, GAI & PAI) in favour of simpler habitat indices such as the IHI (which occurs in two levels), as they generally integrate the effects of changes to ecological drivers.

Advantages of the EcoStatus approach include its comprehensivity, taking into account the relationships between ecological drivers and biotic attributes, providing one of the most comprehensive approaches of all the frameworks considered. The comprehensive weighting systems used to calculate indices of each attribute and the EcoStatus status also make the final estimates highly representative of actual ecological conditions. The ability to adapt the level of assessment according to the level of detail required is also highly advantageous by making it more accessible to projects with restricted resources. However, major shortfalls also include the weighting system, which is complicated (requiring expert knowledge that uses the Likert approach in the PESEIS but MCD and hierarchical decision analysis in the EcoStatus methods), and involves a large degree of subjectivity that may lack consistency and an empirical basis.

Lessons to be learned for a global assessment include that ecological status can be determined by focusing on biotic response attributes, whilst ecological drivers, determined in parallel, are used for interpretation of changes in biotic conditions, as in NARS.

2.1.1.4 The River Health Index (RHI), China

The framework for the evaluation of freshwater health in China (Xie *et al.*, 2020) is designed to work in collaboration with the River Chief System (RCS) of water governance in China (Xie *et al.*, 2020). The political system in China is one of centralised authority and a hierarchical style of governance. In water governance, this is achieved using the River Chief System (RCS), where 'River Chiefs' are responsible for the management of water bodies at various levels (provincial, municipal, county and township) and their performance evaluated using a quantitative assessment method, outlined in (Development Research Center, 2019). This includes an evaluation of the outcomes of water protection and management. This made it necessary to establish a standard quantitative approach applicable nationwide to assess both freshwater health and the effectiveness of the RCS. The authors therefore conducted an extensive review of freshwater heath assessments from around the world and the regional standards in China to determine the core components and best practices of effective monitoring systems before proposing a new nationwide framework. The newly proposed framework thus involves two main categories a) ecosystem integrity (physical habitat, water quantity, water quality, aquatic life and ecological processes) and non-ecological performance (social services and water governance) (Xie *et al.*, 2020). 'Ecological integrity', they suggest, cannot be indicated by a single measure so must necessarily integrate measures of all the relevant components, namely 'physical habitat', 'water quality', 'water quantity', 'aquatic life', and 'ecological processes'. 'Non-ecological performance' is unique to this index and is included purely in accordance with the RCS standards but is based on the recognition that ecological and non-ecological factors are intrinsically linked and interdependent (Xie *et al.*, 2020). The 'social services' sub-category includes provisioning (water supply, aquaculture), regulating (e.g., flood protection) and cultural services (e.g., recreation, aesthetic beauty, & spiritual significance). 'Water governance' includes managerial protocols that facilitate the implementation of water protection activities and efficient water governance at different administrative levels. For each subcategory, they list a mandatory and optional indicators for inclusion, although the source of data, means of data collection or measured value involved can be varied as they are all standardised to a scale of 0 -100. To calculate the RHI, indicators are first weighted before averaging to calculate the score for each subcategory. Subcategory weights, however, are fixed in the calculation of 'Ecological Integrity' as physical habitat (0.2), water quantity (0.15), water quality (0.15), and aquatic life (0.2), and 'Non-ecological Performance' as Social Services (0.2) and Governance (0.1)), which are thus fixed at a ratio of 7:3 in the calculation of the overall RHI score on a scale of 0 -100. Although evaluation is mostly carried out at the reach (or lake section) level, scores for larger scales (e.g., the entire river/ lake) are calculated by averaging the scores of each monitored section, weighted according to their length. The score is further categorized as "very healthy" (RHI ≥ 90), "healthy" (75 ≤ RHI < 90), "subhealthy" (60 ≤ RHI < 75), "unhealthy" (40 ≤ RHI < 60), or "hazardous" (RHI < 40).

The main advantages of the RHI framework are that it is highly flexible to regional conditions and inclusive of existing monitoring programs thanks to the standardised protocol that requires the measurement of key indicators but does not define the source or way in which they are measured. It is also highly scalable, applicable from the reach to national scales. The inclusion of non-ecological factors is also beneficial to the centralised style of government in China and the incorporation of whether good ecological health is being met in the assessment of the river chief's governance incentivises good governance.

In terms of the lessons learned for a global RH assessment, the RHI, once again, shows the strength of a standardised protocol for application across various contexts and spatial scales, namely through the indication of key indicators that must be included a means for standardising and integrating the measurements obtained. The inclusion of non-ecological factors, particularly governance, in the quantification of ecological health, although interesting, is not suitable for global application and would require adaption to the various forms of government around the world.

2.1.1.5 The National River Health Program (NRHP), Australia

The National River Health Program (NRHP) was established as a collaboration by the Australian Commonwealth government and states in 1992 as a consistent and standardised method of assessing river health to provide the information needed to reverse the degradation of inland waters (eWater, DSEWPC and LWA, no date; Davies, 2000; FBM and DNRM, 2001; Halse *et al.*, 2002). The aims of the NRHP were to: 1) monitor and assess the ecological condition of Australia's rivers, 2) assess the effectiveness of current management practices, and 3) provide better ecological and hydrological data on which to base management decisions. The first phase of the project, the Monitoring River Health Initiative (MRHI), from 1993 – 2001, aimed to develop a means to assess ecological conditions using aquatic macro-invertebrates, leading to the creation of the AUSRIVAS (**Au**stralian **Ri**ver **A**ssessment **S**cheme) models, which is modelled on the British bioassessment system RIVPACS (River Invertebrate Prediction and Classification System). These models predict the macroinvertebrate families likely to occur (with a probability >0.5) at sites according to the environmental variables present in the absence of anthropogenic stress (e.g. position, habitat etc.), against which the number of families observed 'O' can be expressed as a ratio (Davies, 2000). The O/E values are then divided into the following 'biological condition' bands: X (>1.12, more biologically diverse than reference), A (0.88 – 1.12 'similar to reference', B (0.64 – 0.87, 'significantly impaired'), C (0.40 – 0.63, 'severely impaired') and D (0 – 0.39, 'extremely impaired') (Nichols and Dyer, 2013). AUSRIVAS models have thus been developed for all the main habitat types found in Australian rivers (riffles, edge, pools and bed habitats) and for all states and territories (e.g. (FBM and DNRM, 2001; Halse *et al.*, 2002). The second phase of the NRHP was then to conduct the First National Assessment of River Health (FNARH) using the AUSRIVAS models and by 1999 over 5000 sites had been assessed. Since then it has been used to survey freshwater ecosystems across Australia and forms the scientific basis of subsequent RH monitoring frameworks both in Australia (e.g. Tasmanian River Health Monitoring Program (Hardie, Bobbi and Uytendaal, 2018)) and abroad. Although river health is based on the bioassessment of macroinvertebrates, the AUSRIVAS also included a physical assessment. This measures 44 geomorphic, physical and chemical variables, including site variables (classification, position, dimensions), physico-chemical variables, habitat variables (substrate surface areas), and habitat characteristics (riparian, aquatic veg, algae, detritus, flows). These are used for interpretation of results and diagnosis of site conditions but are separate to the results of the bioassessment.

Advantages of the NRHP are that the AUSRIVAS method is easy to undertake and provides rapid results. The use of predictive models also eliminates the need for reference sites within any particular survey. This has resulted in their widespread adoption in Australia. However, a recent review of the performance of the AUSRVIAS models (Bruce C. Chessman, 2021) found that 48 % of studies utilising AUSRIVAS rated its performance, in terms of its repeatability and capacity to discriminate among sites with different degrees of human impacts, as 'poor' and only 28 % as 'good'. The AUSRIVAS O/E indices are thus "weak or inconsistent indicators of anthropogenic stress" (Chessman, 2021). Reasons for this include: (1) variable reference-site status (i.e. subjective site selection and inclusion of sites with human influence (e.g. regulated rivers), (2) inappropriate model predictors (i.e. predictor variables used for the predictive models (e.g. salinity) may be subject to anthropogenic influence), (3) limitations of O/E indices (i.e. no account of abundance, sensitivity to chance detection/non-detection of taxa (particularly in environments with low taxonomic richness), which also depends on the choice of threshold), (4) inconstant sampling methods, and (5) neglect of non-seasonal temporal variability (i.e. aseasonal or supraseasonal regimes that govern much of Australia). The Ephemeroptera–Plecoptera–Trichoptera (EPT) and stream invertebrate grade number – average level (of SIGNAL) have been shown to out-perform the AUSRIVAS O/E indices (e.g. (Walsh, 2006; Cox, et al. 2019; Chessman, 2021), offering better alternatives. They also do not rely on build-in reference data. Further drawbacks that eventually led to the cessation of the NRHP are the huge financial and labour-intensive requirements to carry out large-scale assessments.

Lessons learned for a global RH assessment include that the use of a single metric to define RH limits its representativeness, suggesting that the integration of multiple indicators offers a more representative means of assessment. It also emphasizes the importance of appropriate reference site selection, as inappropriate and inconsistent selection can seriously bias the results. Nevertheless, the NRHP is once again an example of a framework that infers ecological health from the biological conditions, suggesting that the focus of a global assessment should be biological, but perhaps one that incorporates multiple components.

2.1.1.6 The Murray-Darling Basin Sustainable Rivers Audit, Australia

The sustainable rivers audit (SRA) is the assessment of river ecosystem health in the Murray-Darling Basin (MDB), Australia, involving the Australian Government, state jurisdictions and Murray-Darling Basin Authority. Although similar to previously mentioned frameworks (i.e. WFD, NARS, EcoStatus Reports), it is designed specifically to represent functional and structural links between ecosystem components, biophysical condition and human interventions (Davies *et al.*, 2010). Using measurements obtained *in situ* and from models, indicators of condition are combined for five themes: hydrology, fish, macroinvertebrates, vegetation and physical form, although for the first audit, only the hydrology, fish and macroinvertebrates were considered. Similar to the NARS, sampling sites were selected randomly by valley (i.e., sub-basins) and altitudinal zones to ensure representativeness. Indicator conditions are estimated as ratios to the reference conditions (estimated as the condition had there been no human intervention) and scaled 0 -100 with the reference condition = 100. For each theme, indicators are combined into sub-indices and sub-indices into a single 'Ecosystem Health Index' using expert-system rules (weighting) based on underlying conceptual models. Conditions are then classified as five categories according to their difference from the reference condition: Good (80 – 100, near reference condition), Moderate (60 – 79, moderate difference), Poor (40 – 59, large difference), Very Poor (20 – 39, very large difference) and Extremely Poor (0 – 19, Extreme Difference).

For the **hydrological** theme, metrics were modelled for 'current' and 'reference' scenarios for the same time-periods, the latter excluding human influences. The five hydrological indicators include (1) High-Flow Events Indicator (magnitude of high flows), (2) Low- and Zero-Flow Events Indicator (magnitude of low flows; proportion of time without flow), (3) Flow Variability Indicator (coefficient of variation of monthly flows), (4) Seasonality Indicator (timing of min and max flows). For **fish**, the two indicators include (1) Expectedness and (2) Nativeness. Expectedness provides information on species richness relative to the reference condition based on the metrics a) Observed to Expected Ratio and b) Observed to Predicted Ratio. Both compare the number of native species predicted to occur under reference conditions and those collected, although the first corrects the number of species 'expected' to occur downward by taking into account rare species. Nativeness provides information on the proportions of native versus alien species in the a) biomass, b) abundance, and c) species richness. For **Macro-Invertebrates**, samples are collected by standardised kick-sampling and weep-netting using the AUSRIVAS protocol (Davies, 2000) twice per year, including edge and instream habitat, and condition determined using two indicators. The first is an O/E metric comparing the 'observed' and 'expected' families. The reference condition being developed through application of filters, based on traits determining family distributional limits for temperature, hydrology, geomorphology and biogeography. Filter variables were estimated applied using EO data. The second is a SINGAL O/E Metric. This also compares the 'observed' vs. 'expected' taxa' but after the application of tolerance scores using SINGAL (Stream Invertebrate Grade Number Average Level), which reflects the sensitivity of macroinvertebrates to disturbances (0 = high tolerance, 10 = high sensitivity). Using expert-rules, the two indicators were combined to provide a single metric of macroinvertebrate conditions. For the **physical form**, condition is determined from changes in physical form (channel form, bank dynamics, bed dynamics, floodplain features), and sediment dynamics at the river-reach scale. The form is assessed *in situ*, whilst sediment dynamics are assessed using the SedNet model (Wilkinson et al. 2004). Data are aggregated for larger (basin) scale reporting. Reference conditions are determined from least-disturbed sites, historical maps, aerial photographs and model outputs. For the **vegetation**, although not included in past assessments, factors will be assessed for riparian and floodplain vegetation at a) catchment and b) reach/ floodplain unit scales and combined into a single indicator value. Indicators include taxonomic composition and disturbance, nativeness/ weediness, function, and structure.

Advantages of the SRA include the inclusion of indicators obtained via modelling, namely for hydrology and sediment dynamics, as they are scale-independent and include a determination of the reference conditions. There is also strong potential for use of EO data for large-scale riparian vegetation assessment in future applications. The sampling design is also highly representative, covering a range of relevant ecological components and follows standardised protocols, whilst the conceptual models describe the links between ecological themes and drivers so are useful to support adaptive management and communication of results. A major concern raised with the SRA is how to treat reference conditions under a climate change scenario as the ecosystem adapts to long-term changes in temperature and runoff.

Lessons to be learned from the SRA for a global assessment include the usefulness of using modelled data for widespread scale-independent assessments, including for determining reference conditions (e.g., flows) but that this requires accurate empirical basis. It also opens the debate on how climate change should be factored into the determination of reference conditions. This seems to have two possible solutions. One could either 1) include the changes in condition because of CC within the condition assessment and therefore relative to a reference for pre-CC conditions or 2) exclude changes in the condition because of CC and therefore have a dynamic reference, based on the reference conditions that take into account CC but no other human impacts.

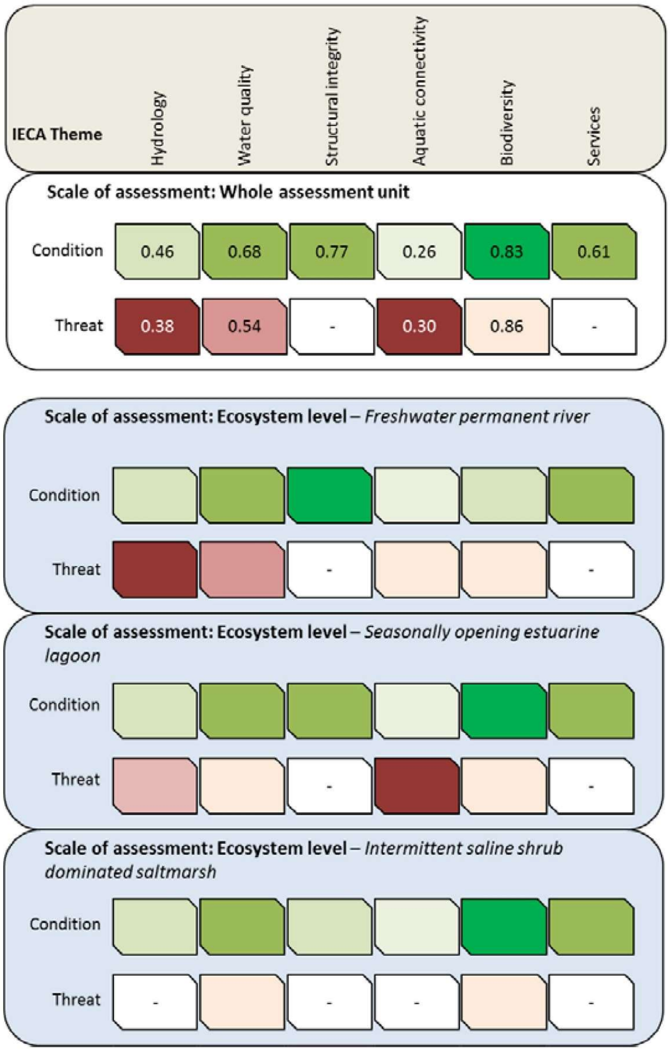
2.1.1.7 Integrated Ecosystem Condition Assessment Framework (IECA), Australia

The Integrated Ecosystem Condition Assessment (IECA) is a manual providing a flexible method for undertaking integrated ecological conditions assessments of aquatic ecosystems, including rivers, wetlands and estuaries (Department of the Environment and Energy, 2017). It is Module 5 of the Aquatic Ecosystems Toolkit; a series of five manuals to guide the classification and condition assessment of aquatic ecosystems and provide guidance on how to identify high ecological value aquatic ecosystems, developed by the Aquatic Ecosystems Task Group and later, the Wetlands and Aquatic Ecosystem Sub Committee (WAESC) (Australian Government and State and Territory representatives). It is designed for application at various scales and is flexible in its application, so is meant to build on existing methods and programs being implemented in Australia for intended use by national, state and regional agencies tasked with assessing aquatic ecosystem conditions. The IECA framework's objectives are thus to assess and report on 1) the status and trends in ecological conditions and threats (relating to predetermined baseline or reference points for priority ecological values of aquatic ecosystems) and 2) the effectiveness of management interventions at maintaining/ improving aquatic ecosystem conditions. It is developed for application to all inland and estuarine ecosystems and can operate at multiple spatial scales. Central to the framework is the ability to build on and incorporate existing methods and programs adopted by Australian jurisdictions. The nested or hierarchical structure of the framework is thus a key feature and it is critical for condition assessments to include both biotic and abiotic components of the ecosystem and where appropriate ecosystem services. Ideally data should come from different organisational levels (species, communities, biotopes) (Borja *et al.*, 2016). They thus delimit eight steps preceded by a planning phase necessary to develop an ecosystem health framework for any particular region.

The planning phase involves determining the current context and state of understanding of the assessment unit. This includes first articulating the purpose of the IECA and framing the question, which involves clarifying one's objectives, identifying existing target, trigger or threshold values, stakeholder identification and engagement, and establishing the spatial boundaries of the assessment unit. It also involves groundwork to establish an oversight body, collate existing information, define the scale of assessment, identify ecosystem types, identify any existing conceptual models of the assessment unit, and identify any externalities likely to affect the assessment unit (e.g., floods etc.). The framework itself then consist of eight steps, these include Step 1: **Identify and prioritise values,** Step 2: **Identify and prioritise threats,** Step 3: **Develop Key Evaluation Questions,** Step 4: **Identify and prioritise indicators,** Step 5: **Design assessment and implementation,** Step 6: **Analyse and Aggregate,** Step 7: **Harmonise and Integrate**, and Step 8: **Develop Report Card**. This provides a consistent approach for application across jurisdictions by providing a means to define, assess and report on aquatic ecosystems at various scales and aims to encourage cross-jurisdictional collaboration, particularly for multi-jurisdictional assessment units. Core to this consistency is their definition of ecological conditions as "t*he state or health of individual animals or plants, communities or ecosystems as they relate to values and ecosystem services with reference to specific management goals or objectives and assessment against a defined baseline. Condition indicators can be physical-chemical or biological and represent the condition of the ecosystem. They may also be surrogates for pressures and stressors acting within the ecosystem."* Also important is their adoption of six common themes that summarise the nature of aquatic ecosystems for assessment and reporting, namely hydrology, water quality, structural integrity, aquatic ecosystem connectivity, biodiversity, and ecosystem services. For these, they further define a set of desired and optional indicators and recommend the inclusion of several indicators per theme for greater representativeness. Many existing regional frameworks in Australia will be compatible with the IECA requiring only minor alterations. They thus recommend following the manual as a checklist of sorts, determining if the information from the framework in question is' fit for purpose'.

They describe in detail the considerations and process involved in each step, including for the identification and prioritisation of indicators, sampling design for recording indicators that may be lacking and methods for the aggregation, harmonisation, integration, and reporting of data. They recommend aggregation of metrics across various scales before integration as this allows for the independent sampling of each metric, reducing uncertainty and increasing efficiency. Integration-before-aggregation they recommend against as it requires gathering data for all metrics at the same scale, which is not adequate as the most appropriate scale for the gathering of data varies by metric. It also aggregates errors and is massively affected by the presence of missing data. They recommend the harmonisation of data to a common scale of 'condition scores' (0-100) by comparing the observed changes to reference conditions, which they define as the pre-European conditions. They provide detail on the optimal methods for integration of results for metrics into indicator scores, indicator scores into themes and themes into overall condition scores (i.e., upscaling) to the scale of reporting using averaging, modelling or summing approaches as appropriate to the indicator types. Generally, they recommend weighted averaging for indicators gathered at the site scale and modelling to predict values of indicators that have an assessment unit scale reference value (i.e., those measured per area or length) at larger scales, or summing when values for each subunit are known. Scores for themes (or components) are then divided into five categories (or bands) for condition (state indicators) or threat (pressure indicators) and represented on a colour scale for effective communication of results (see Figure 2.3). These they recommend are most effectively communicated using a dashboard approach, where the boxes indicating the condition and threat scores for each theme are represented separately. They suggest this be carried out at two levels of integration, including for indicators by theme per ecosystem, and for themes at the ecosystem level, as in Figure 2.3. They also find it essential to document assumptions in the methodology and knowledge gaps to ensure transparency.

The main advantage of the IECA is that it is highly flexible and can be applied to any aquatic ecosystem and for different management needs at multiple scales. This is due mostly to its hierarchical nature that can build on existing frameworks allowing for the tailored selection of metrics to use for indicators. It is also highly consistent enabling comparability across jurisdictions due to the clear definition of terms and setting of common themes and standard methods for assessment and reporting of conditions. This, it is hoped, will foster cooperation between jurisdictions. In turn, enabling more effective management of cross-boundary assessment units. In terms of negatives, the novelty of the framework means that potential shortfalls emerging from its application are not yet apparent. However, the major lessons from the IECA for a global RH assessment include detailed guidance on how to carry out the relevant groundwork, select appropriate indicators, and aggregate, harmonise and integrate scores for reporting at varying scales. It thus provides a useful blueprint for the formulation of a standard protocol for assessment at the global scale.



IECA Score	ICEA Band description	Condition	Threat
1 – 0.8	Largely unmodified	Excellent	Negligible
0.79 – 0.6	Slightly modified	Very good	Minor
0.59 – 0.4	Moderately modified	Good	Moderate
0.39 – 0.2	Substantially modified	Fair	Major
0.19 – 0	Severely modified	Poor	Extreme
-	Not assessed		

This hypothetical example shows an output at the Whole of Assessment Unit and three Ecosystem level assessments, Level 2 Integration. No scores are presented for the themes at the Ecosystem level as this is a purely hypothetical/illustrative of the nested nature which report cards could be presented. In a completed assessment the scores for the themes in each Ecosystem would be included in the report card.

Report cards could also be presented for each of the indicators x ecosystem type x theme – i.e. Level 1 Integration.

Figure 2.3: Hypothetical example of how a report card would appear using the IECA framework. Note the dashboard approach indicating the condition and threats to each theme per ecosystem type using a colour scale indicated in the top right-hand corner (sourced from (Department of the Environment and Energy, 2017)

2.1.1.8 Freshwater Biophysical Ecosystem Health Framework, (FBEHF) New Zealand

The Cawthron Report (Clapcott *et al.*, 2018) describes a framework for assessment of biophysical ecosystem health of fresh waters in Aotearoa New Zealand, commissioned by the Ministry of the Environment to help managers meet the requirements of the National Policy Statement for Freshwater Management 2017 and the Environmental Reporting Act 2015. It is based on a series of workshops and literature review (namely the other frameworks outlined here) to identify the key requirements for developing and implementing a framework. The framework's purpose was to "provide a consistent approach for assessing freshwater biophysical ecosystem health" to enable governments, communities and individuals to gauge the maintenance and improvement of ecosystem health. Ecological integrity was defined as *the ecosystem's ability to maintain its evolving structure and function over time in the face of external stress as compared to a reference benchmark*.

The proposed assessment of ecological integrity included aquatic life, physical habitat, water quality, water quantity and ecological processes. The key performance attributes necessary for the practical application of the framework include consistency (broadly applicable across fresh waters), representativeness (integration of multiple components), robustness (informed by science), informativeness (easily understood), flexibility (suits varied application across ecosystem types) and scalability (can be modified for reach- to national-scale assessments). Application of the framework further required knowledge of the suitability of component indicators and their appropriate benchmarks, as well as methods for data aggregation, harmonisation, integration, and reporting. The report thus provides an example of how component indicators can be identified and further recommend the development of conceptual models to illustrate core components and indicator links to management options and the development of best practice guidelines for data analysis and reporting (including analysis of existing data at multiple scales). A useful feature of the approach is that reporting changes with scale, with less detail and more synthesis at larger scales.

This report is highly beneficial to the purposes of this study due to its similar objective, to provide a widely applicable RH assessment framework. The key performance attributes identified by the authors and the methods proposed for each are thus equally applicable to a global framework. The core components of an integrated assessment approach and definition of ecological integrity are also useful to a global RH assessment. Therefore, the report provides as a good template to use for our development of a global RH assessment framework development.

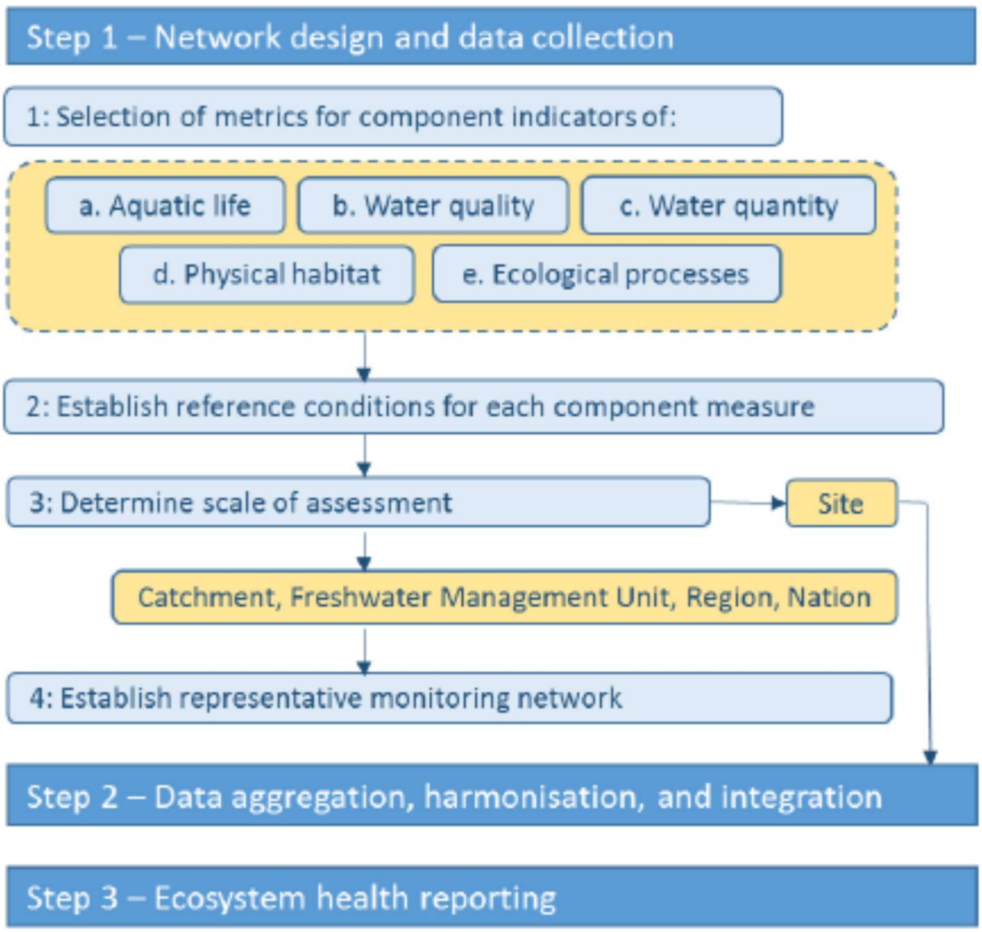


Figure 2.4: Steps in the application of the framework for freshwater ecosystem health in New Zealand (source: (Clapcott et al.*, 2018))*

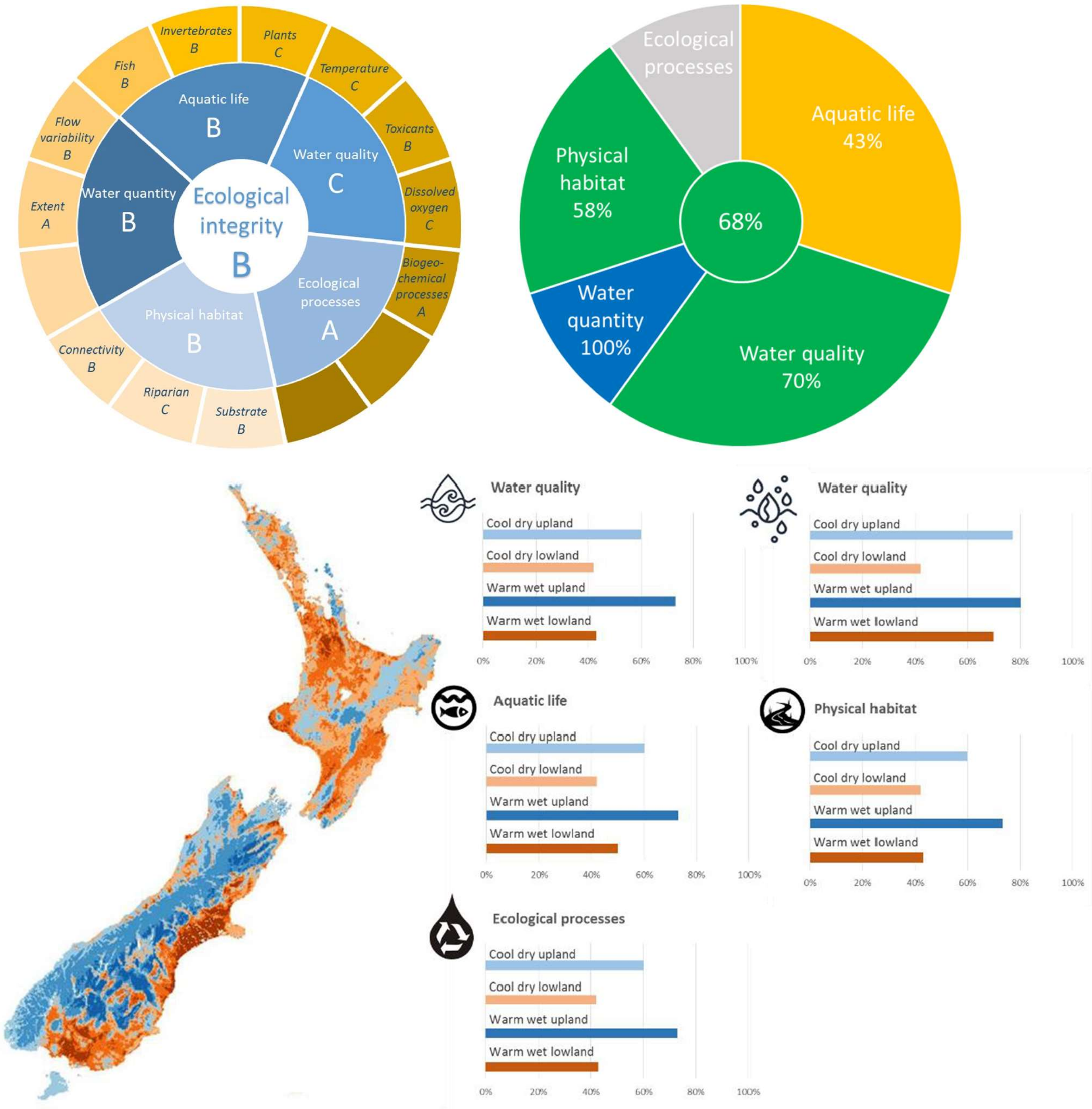


Figure 2.5: Hypothetical report cards proposed for the Freshwater Biophysical Ecosystem Health Framework for New Zealand. Top left, circular diagram showing how integrity can be simultaneously reported at various levels of assessment from the overall score at (the centre), per each of the five core ecological components (middle layer), to individual indicators (outer layer). Condition is indicated by colour and scores categories shown. Top right, a variation of this, showing the integrity of ecosystem components and overall, as a percentage value with corresponding colour scheme, as well as the weighted importance of each component to the contribution of the overall score. Below, conditions of freshwater ecosystems components indicated separately by stream classes.

2.1.1.9 Mekong River Basin Indicator Framework (MIF)

The Mekong River Commission (MRC) is an inter-governmental organisation for regional dialogue and cooperation in the Lower Mekong River Basin, based on the Mekong Agreement between

Cambodia, Lao PDR, Thailand and Viet Nam established in 1995, although the Upper Mekong States of China and Myanmar have been dialogue partners since 1996. The organisation serves as a regional platform for water diplomacy and a knowledge hub of water resources management for the sustainable development of the region. Its mission is to “promote and coordinate sustainable management and development of water and related resources for the mutual benefits of the lower Mekong countries and the people's well-being”. According to the Basin Development Strategy (MRC, 2021) regular basin monitoring and 5 yearly State of the Basin Reports (SOBRs) are integral to the strategic planning cycle. Since 2018, the SOBRs have been (and continue to be) structured around a set of indicators at three hierarchical levels known as the *Mekong River Basin Indicator Framework* (MIF) for monitoring, assessment and reporting on the state of the basin (MRC, 2019c). At the highest level, itis consists of five dimensions*:* ***environment***, ***social***, ***economic***, ***climate change***, and ***cooperation***, indicated by 15 strategic indicators. For the ‘Environment’ dimension, these include *water flow conditions in mainstream*, *water quality and sediment conditions*, *status of environmental assets*, and *overall environmental condition*. At the second level 55 assessment indicators (11 pertaining to the ‘Environment’) were selected aimed at providing more detailed information and supporting the quantification of strategic indicators. These strategic and assessment indicators included:

1) **Water flow conditions in the mainstream**
    a. Dry season flows - compliance with PMFM
    b. Flood season peak flows - compliance with PMFM
    c. Tonle Sap reversal flows - compliance with PMFM
    d. Timing of onset of wet season flow.
2) **Water quality and sediment conditions**
    a. Water quality and ecological health - compliance with PWQ & TGWQ
    b. Suspended Sediment concentrations
    c. Salinity intrusions in the delta
3) **Status of environmental assets**
    a. Wetland area
    b. Condition of riverine habitats
    c. Condition and Status of Fisheries and other aquatic resources
    d. Condition and status of ecologically significant areas

Strategic and assessment indicators are reported using a traffic light colour scheme according to the level of concern and urgency of actions.

No immediate concerns

Some significant concerns to address

Considerable concern, urgent action needed

Insufficient data to form a view, requires action to address knowledge gaps

Finally, at the third and lowest level, 160 monitoring parameters were identified to support the quantification of the above assessment and strategic indicators. These are supported by six monitoring programmes designed around the major areas of concern in the basin, namely flow modifications, sediment reductions, loss of wetlands, deterioration of riverine habitats, and over-exploitation of fisheries. These included *hydrometeorology, sediment*, *water quality*, *aquatic* *ecology*, *and fisheries* monitoring to understand and assess the availability and condition of water resources and environmental & social conditions to understand their linkages with water resources. For hydrometeorological monitoring, the Procedure for the Maintenance of Flows on the Mainstream (PMFM) uses rainfall and water level data gathered at 49 automated hydro-met stations throughout the basin to determine thresholds of flow volumes in flood and dry seasons, as well as the timing of the start of wet season flows, as this is crucial in triggering fish migrations. For sediment monitoring, the Discharge and Sediment Monitoring Program (DSMP) measure discharge and suspended sediment concentrations (SSC). Water quality monitoring is carried out according to the Procedures for Water Quality (PWQ) and Technical Guidelines (TGWQ) and takes into consideration the requirements for human health, aquatic life and agricultural use. The calculation of water quality conditions for the protection of human health follow the equations set out in the Canadian WQI for aquatic life (see Section 3.4.1.6 below) and for agricultural use by comparison of electrical conductivity against the recommended thresholds for irrigation and paddy rice (MRC, 2019a). The water quality scores for human health on a scale from 0 -100 and aquatic life on a sale of 0 -10 can thus both be expressed in five colour-coded categories from A (High Quality, blue) to E (Very Poor Quality, red). For Aquatic ecology, the *Ecological Health Monitoring* (EHM) programme is designed to monitor the river's long-term ecological health by using regular (annual) biological monitoring of four major organism groups (Benthic Diatoms, Zooplanktons, Littoral Macro-invertebrates and Benthic Macro-invertebrates). Phytoplankton were added in 2017 to monitor impacts from Hydropower projects. A healthy ecosystem is indicated by high abundance, high average richness and low Average Tolerance Score Per Taxon (ATSPT). During the sampling period, a set of baseline conditions describing the ecological health of the LMB were established through the MRC Biomonitoring or Ecological Heath Monitoring Report (MRC, 2010). Fisheries monitoring involves collecting data on 1) fish abundance and diversity (specifically changes in fish species diversity, catch composition and abundance to improve understanding of environmental factors to inform fisheries planning and management), 2) fish larvae using drift monitoring and 3) Bagnet (dai) fishery, particularly catches of small mud carp but also total fish abundance, diversity and prices.

For the 3rd strategic indicator, 'status of environmental assets' changes in the wetland area is easily determined from satellite imagery, as are 'ecologically significant areas', which consists of rivers, wetlands, forests and grasslands, although it appears the status and trends for these features are only assessed qualitatively at the strategic indicator level. Riverine habitats include exposed sandy and rocky areas, deep pools and backwaters, however, their status is also only rated qualitatively at the strategic indicator level.

Advantages of the MIF and SOBRs are that they provide a comprehensive overview of the basin's conditions and highlight issues that must be addressed for the objectives of the 1995 Mekong Agreement to be met. The results also align strongly with the SDGs, particularly 6.6.1: *change in the extent of water-related ecosystems*. The means of reporting using the traffic-light colour scheme is also effective at communicating results effectively for the integrated management of water resources. Challenges include the lack of data verification for the Chinese portion of the basin by the Chinese government. Shortfalls of the framework include that the basin wide conditions of the Indicator Framework are too broad to be used to describe actual freshwater ecological health. For the 'Status of environmental assets', in particular, many of the assessment indicators are based on very rough assessments. On the other hand, the failure to integrate the results of each of the six sophisticated monitoring systems (including water quality, hydrometeorological, and aquatic life) into a single RH index detracts massively from the representativeness of the results at indicating RH.

Lessons learned include the utility of a hierarchical framework to provide varying levels of detail according to the objectives (decision making (strategic), assessment or monitoring) and the importance of properly integrating indicators within the given framework for representativeness.

2.2 Review of global frameworks

2.2.1.1 UN Sustainable Development Goals (SDG 6)

The Sustainable Development Goals (SDGs) of the UN 2030 Agenda (United Nations, 2015) consist of 17 interconnected global goals aimed at ending poverty, protect the planet and ensuring all people enjoy peace and prosperity. SDG 6 is to “ensure availability and sustainable management of water and sanitation for all”. An integrated monitoring initiative supports countries in monitoring water- and sanitation-related issues. Global indicators for SDG 6 include drinking water, sanitation and hygiene, wastewater treatment, water quality, water-use efficiency, water stress, water resource management, transboundary cooperation, and water-related ecosystems (UN Water, 2017). For water-related ecosystems, the target is to “*protect and restore water-related ecosystems, including mountains, forests, wetlands, rivers, aquifers and lakes by 2030*”. The state of water-related ecosystems is indicated using proxies of ecosystem health (spatial extent, basic water quality, river discharge) as the initial attempts to include the direct assessment of ecosystem health were rejected as there were no recognized methods available that could be applied at a global level, hence this project. For the ‘water quality’ indicator, it is considered good if it does not damage ecosystem function or human health, according to national target levels of selected parameters, namely dissolved oxygen (DO), electrical conductivity, nitrogen, phosphorous, and pH.

A key benefit of SDG 6 is that it provides the potential to consider the inter-relatedness between ecosystem health, water quality, sanitation, water stress, and water governance, all of which are dependent on each other.

2.2.1.2 Post 2020 Framework – Convention on Biological Diversity

The Convention of Biological Diversity of the UN in October 2021 has documented a new approach, the Post-2020 Global Biodiversity Framework, to global efforts to protect biodiversity in all its forms (CBD, no date). At the time of writing only preliminary information is available so what is presented here may change, up until the final adoption of the framework that will take place in 2022.

The Post-2020 Framework includes detail of a number of indicators, divided into:

(a) Headline indicators: A minimum set of high-level indicators which capture the overall scope of the goals and targets of the post-2020 global biodiversity framework which can be used for tracking national progress, as well as for tracking regional and global progress. These indicators, or a subset of them, can also be used for communication purposes;

(**b) Component indicators:** A set of indicators for monitoring each component of each goal and target of the post-2020 global biodiversity framework at the global, national and regional levels;

(c) Complementary indicators: A set of indicators for thematic or in-depth analysis of each goal and target. These indicators would be primarily for use at the global level. It is expected that this list will be dynamically updated, in consultation with the Biodiversity Indicator Partnership, to reflect new scientific and indicator development.

The criteria for selection of these indicators includes:

(a) The data and metadata related to the indicator are (or will be) publicly available;

(b) The methodology for the data product is either published in a peer reviewed academic journal or has gone through a scientific peer review process;

(c) There is evidence that the indicators will be regularly updated with a time lag of less than five years between updates;

(d) There is an existing mechanism for maintaining the indicators, including, for example, by a member of the Biodiversity Indicators Partnership, an intergovernmental organization or a well-established scientific or research institution.

Useful indicators for global river health monitoring are shown in Table 2.1:

Table 2.1 Relevant indicators from the Post-2020 Framework as at November 2021. Only included are river related indicators that serve to quantify the state of the ecosystem in some way.

Headline indicator	**Component indicator**
A.0.2 Species Habitat Index	A.2.1 CMS connectivity indicator (CMS)
	A.3.1 Ecosystem Integrity Index
	A.4.1 Species status information index (GEO BON)
A.0.3 Red list index	
B.0.1 National environmental economic accounts of ecosystem services	B.3.1 Nature’s material contributions including food, water and others (from the environmental economic accounts)
	B.4.1 Nature’s non-material contributions including cultural (from the environmental economic accounts)
2.0.1 Percentage of degraded or converted ecosystems that are under restoration	
	2.2.1 Maintenance and restoration of connectivity of natural ecosystems
3.0.1 Coverage of Protected areas and OECMS (by effectiveness)	3.4.1 Species Protection Index (GEO BON)
5.0.2 Proportion of fish stocks within biologically sustainable levels	
6.0.1 Rate of invasive alien species spread	6.3.1 Rate of invasive alien species impact (GEO BON)
11.0.1 National environmental-economic accounts of regulation of air quality, quality and quantity of water, and protection from hazards and extreme events for all people, from ecosystems	11.2.1 Proportion of bodies of water with good ambient water quality (SDG 6.3.2)
	11.2.3 Level of water stress (SDG 6.4.2)
15.4 Move towards the full sustainability of extraction and production practices, sourcing and supply chains, and use and disposal	15.4.1 Ecological footprint
20.0.1 Indicator on biodiversity information and monitoring, including traditional knowledge, for management*	

The primary value of this framework is that it has similar objectives to those in this project. Although the emphasis is on biodiversity, by definition this would include the state of the ecosystem. The CBD process is ongoing and should be resolved in 2022 at which time should provide a useful contribution.

2.2.1.3 GEO BON – Group on Earth Observations Biodiversity Observation Network

GEO BON has many biodiversity orientated indicators that make use of global datasets where possible. As noted on their website *GEO BON with its scientific partners introduces a set of global indicators integrating biodiversity observations, remote sensing data, and models to address important gaps in our understanding of biodiversity change across local, national and global spatial scales (GEO BON 2015). These indicators are based on biodiversity observations, harmonized across multiple data sources and standardized, to allow for a consistent monitoring basis.*

Included in Table 2.2 are some of the potentially relevant indicators although there is little mention of aquatic ecosystems. Note that many have been excluded as they focus purely on terrestrial ecosystems. Some of these may become relevant however, if terrestrial data is used in a model of river impact.

Table 2.2 GEO BON indicators relevant to river health. Note that despite some promising indicator titles, aquatic ecosystems were often excluded and thus they are not listed.

Indicator	**Description**	**Scale**	**Comment**
Species Habitat Index (SHI)	Measures changes in the estimated size and quality of ecologically intact areas supporting species populations	Locally collected observations and remote sensed habitat	Does NOT include aquatic ecosystems
Biodiversity Habitat Index	Biologically-scaled environmental mapping and modelling to estimate impacts of habitat loss, degradation and fragmentation on retention of terrestrial biodiversity globally, from remotely-sensed forest change and land-cover change datasets.	Remotely-sensed forest change and land-cover change datasets to recent advances in biodiversity informatics, ecological meta-analysis, and macro-ecological modelling	Does NOT include aquatic ecosystems
Global ecosystem restoration index (GERI)	Composite index that integrates structural and functional aspects of the ecosystem restoration process	Remote sensed	Three key and complementary elements of ecosystem restoration: (1) change in ecosystem productivity (2) change in the ecosystem energy balance and (3) changes in land cover
Rate of Invasive Alien	Measures the change in impact risk from invasive alien species (IAS) that are expected to have entered a new region	Country collection of data	Aquatic ecosystem not mentioned

Species Spread	given general observation trends and available impact data		

2.2.1.4 Freshwater Health Index (FHI)

The freshwater health index (Vollmer *et al.*, 2018) integrates information on the sustainability of freshwater ecosystems, human water uses, stakeholders and governance aspects of integrated water resource management (IWRM) as three indicators of ecological vitality (EV), ecosystem services (ES) and governance and stakeholders (GS) at the basin or sub-basin scale. Each component is made up of several major indicators that are in turn comprised of multiple sub-indicators. These include, for ecosystem vitality: **water quantity** (*deviation from natural flow regime, groundwater storage depletion*), **water quality** (*suspended solids in surface water, total N in surface and groundwater, total P in surface and groundwater,* and *other indicators of major concern*), **drainage-basin condition** (*bank modification* (percent of channel modification)*, flow connectivity* (dendritic connectivity index)*, land cover naturalness*), and **biodiversity** (*changes in number and population size of i) species of concern and ii) invasive/ nuisance species*). For ecosystem services: **provisioning** (water supply reliability relative to demand, *biomass for consumption*.), **regulation & support** (*sediment regulation, deviation of water quality metrics from benchmarks*, f*lood regulation, exposure to water-associated diseases*), and **cultural/ aesthetic** (*Conservation/ Cultural Sites*, *Recreation*). For governance and stakeholders: **enabling environment** (*Water resource management*, *right 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*), and **effectiveness** (*enforcement and learning mechanisms*, *enforcement and compliance, distribution of benefits from ecosystem services*, *water-related conflict*).

Indicators were selected based on their relevance and whether empirical data are likely to exist, can be modelled, or can otherwise be collected efficiently and cost-effectively. For ecosystem vitality and ecosystem services, indicators are thus based on monitored or modelled spatial data, whilst governance and stakeholders' preferences are gauged through surveys involving 50 questions on a Likert-type 5-point scale. In both case studies, these were obtained at stakeholder workshops (Vollmer *et al.*, 2018; Bezerra *et al.*, 2021), including members familiar with the governance in the region. To enable comparison between sub-indicators, they are normalised to a common non-dimensional scale of 0-100, indicating negative to positive connotation. The data sources considered for sub-indicators can thus be varied according to what is available for the region in question. The geometric mean of sub-indicators is then calculated for indicator values, and of indicators for components, using weightings that denote the important of each in the determination of the aggregated value. Although there are various weighting methods, the authors chose the Analytical Hierarchy Process (AHP) (Saaty, 2005) as it is well suited to the hierarchical nature of the indicators and allows input from a large number of stakeholders. It is thus subject to stakeholder opinion. For ecosystem vitality, however, sub-indicators and indicators are not weighted (i.e., weighted equally) as their relative importance to freshwater ecosystems is objective and should be informed through empirical, rather than subjective, means. However, the authors do not aggregate component values into a single FHI value as separately they provide important information on the linkages between them. The resulting scores and weightings may then be depicted as a circular diagram (Figure 2.6) for each component with wedges indicating the indicators and sub-indicators, coloured according to the score value (red (0) – blue (100)), and whose thickness corresponds to their relative weighting.

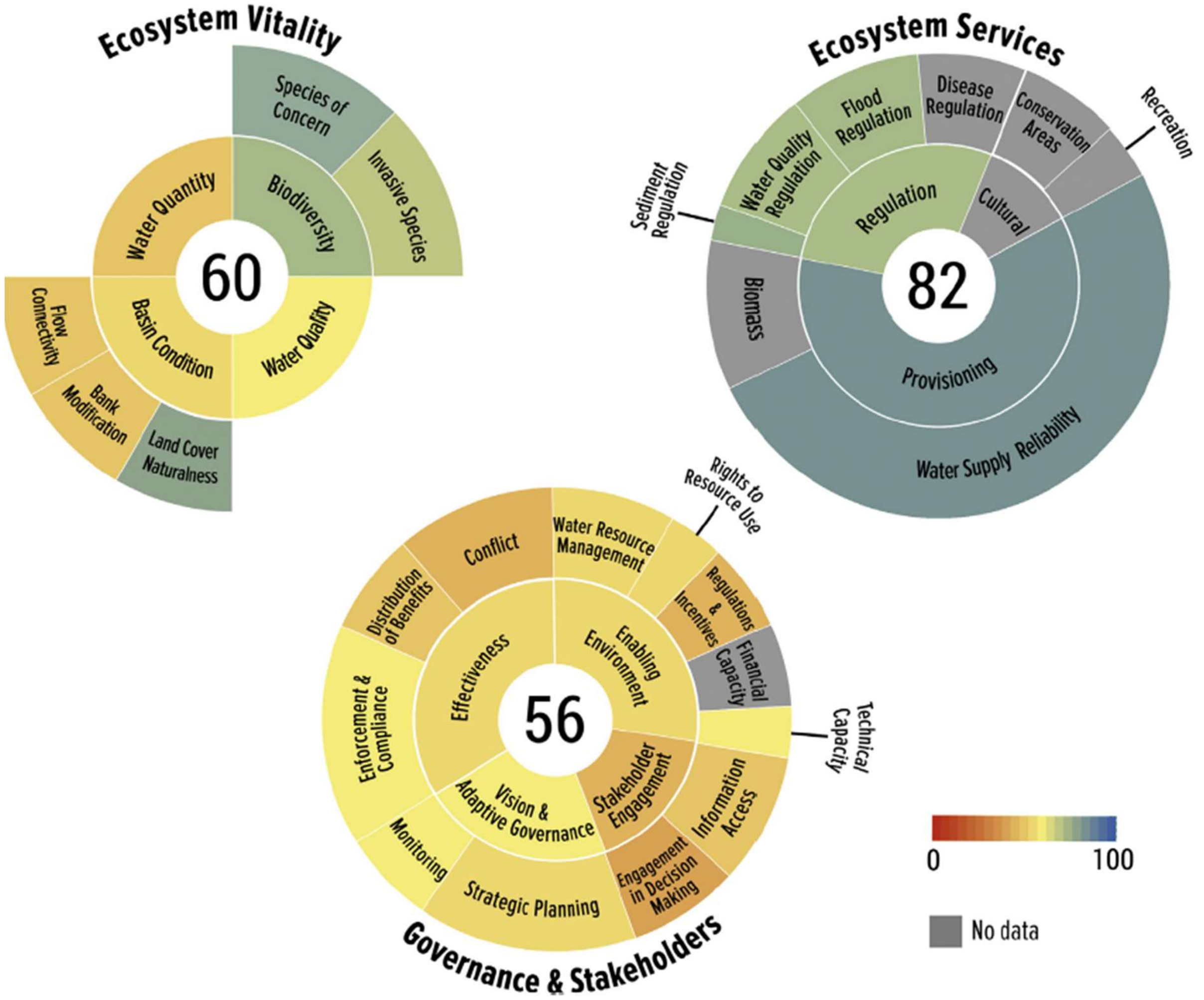


Figure 2.6 Example of the FHI results for the Dongjiang River Basin, Colombia (Vollmer et al, 2018)

Possibly the greatest benefit of the FHI is that it is truly integrative, including aspects of ecology, water use and governance, and acknowledging the important links between them. This enables the discovery of the driving causes of poorly performing indicators, and thus highlights areas for improvement. For example, a low value for ecosystem vitality but high values for other components may indicate where water use is ecologically unsustainable. A low value for Ecosystems Services signals that societal water needs are not being met. A low value for the Governance & Stakeholders can elucidate processes that stakeholders can change to realize improvements in Ecosystem Vitality and Ecosystem Services. It is thus a highly interactive system, suitable for IWRM monitoring and for testing future scenarios. In terms, of a global RH assessment, it is advantageous in its adaptability to existing datasets, especially large-scale spatial data. The nature of score calculation and representation also means that missing data may be omitted. In such cases, weighting values are adjusted accordingly in the score calculation, whilst wedges in the pie chart may be left grey. Rather than a hindrance, this can help highlight data gaps where data collection is required.

From the global RH perspective, the greatest shortfall of the approach is the limited inclusion of ecological and in particular biological data beyond a measure of biodiversity and invasive species. Other shortfalls include the necessity to undertake questionnaire surveys for the governance and stakeholder aspects, making them unlikely in a global assessment. Also, the consideration of biophysical indicators (water quantity, water quality, drainage-basin condition and biodiversity) as having equal weighting in the calculation of the ecosystem vitality index is arbitrary as their contributions are not likely to be equal. Rather, an empirically based weighting system would be more advantageous in a global assessment. The ES and GS indices, weightings, in contrast, are highly subjective but this conveys information of stakeholder perspectives. In addition, the down-scalability of the index to smaller spatial extents (from data gathered at sub-basin or basin scales) is limited.

2.2.1.5 Planetary Boundaries (PB) Framework

The Planetary Boundaries (PB) framework identifies and quantifies boundaries of critical Earth-system processes that regulate the stability of the Earth system and are subject to human influence, which must not be transgressed to avoid large-scale run-away environmental changes that could have disastrous consequences for humanity (Rockström *et al.*, 2009; Steffen *et al.*, 2015). They thus define a “safe operating space for humanity”, as the maintenance of a Holocene-like state of Earth systems, which led to the development of human civilisation, as a precondition for sustainable development. They identified nine Earth-system processes: Climate Change, Biosphere Integrity, Land-system Change, Freshwater use, Biogeochemical flows, Ocean acidification, Atmospheric aerosol, Stratospheric ozone depletion and Novel entities. Of course, all these processes are interconnected, however, an analysis of the interactions between them found two core systems – 1) climate change and 2) biosphere integrity – that underpin all others, such that a transgression of either has the potential to drive the Earth system into a new (less desirable) state. Originally, the framework (Rockström *et al.*, 2009) only considered boundaries at the planetary level. However, given the link between global change and regional processes and the fact that many of the identified global processes are spatially heterogeneous, the updated framework (Steffen *et al.*, 2015) included a two-tiered approach analysing boundaries at the regional and global scales for the five with strong regional operating scales (biosphere integrity, biogeochemical flows, land-system change, freshwater use, and atmospheric aerosol loading). However, the scale at which they were analysed are not the same but dependent on the process involved e.g., biosphere integrity is analysed at the biome scale and freshwater use at level of the river basin.

For each process, the boundary value is set by scientific evidence with targeted input from expert research communities. It is important to realise, however, that boundaries are not equivalent to thresholds, or tipping points, which are values of a variable that if crossed could trigger irreversible run-away changes through feedbacks in the Earth system itself. For example, the melting of the Greenland ice sheet because of global warming would speed up the warming process so is a Climate Change tipping point. Rather, boundaries are set upstream of thresholds (at "better" level) according to the precautionary principle to account for the unknown, inertia in Earth-systems, and to provide early warning of changes. The current human impact is thus set at three levels: The first (Green) is the “Safe Zone”, where human impacts are within the safe operating space. The second (Yellow) is the “Zone of Uncertainty”, characterised by increasing risk that encapsulates both gaps in the scientific knowledge and intrinsic uncertainties in the functioning of Earth systems. As one moves from lower to higher values in the zone of uncertainty, the uncertainty thus increases, and one transitions from a relatively “safe” to “danger” zone in terms of the probability of permanently changing an Earth system. Finally, the third zone (Red) is the “High-risk Zone” beyond the zone of uncertainty, where changes are likely to occur to an Earth-system due to human impacts. For processes, where thresholds have been studied (e.g., Climate Change), boundaries can thus be readily proposed, and the zone of uncertainty is relatively narrow. However, for many of the process (e.g., Biosphere integrity), where thresholds are still unknown or poorly studied, the zone of uncertainty is thus large but with further research the level of uncertainty can be expected to diminish.

The processes most relevant to an assessment of global RH are biosphere integrity, freshwater use, biogeochemical flows, land-system change, and climate change. **Biosphere integrity** includes terrestrial, freshwater and marine ecosystems, and comprises two variables. The first is the **extinction rate, E/MSY** or extinction per million species per year, which measures the **loss of genetic diversity**. In the first framework (Rockström *et al.*, 2009), this was the only biotic variable considered, due to the lack of readily available data for more sophisticated variables. The boundary is set at <10 E/MSY with an ambitious target of 1 E/MSY, which is the upper estimate of the average extinction rate in the fossil record. However, there is large uncertainty over what level of loss would trigger non-linear irreversible changes, so the zone of uncertainty is set from 10-100 E/MSY. Nevertheless, this has undoubtedly been transgressed with a global value between 100-1000 E/MSY. The second is the **Biodiversity Intactness Index** (BII), which measures the loss of functional diversity as the change in population abundance because of human impacts across a wide range of taxa and functional groups at a biome or ecosystem level using pre-industrial era abundance as a reference point. The index typically ranges from 100% (abundances across all functional groups at preindustrial levels) to lower values that reflect the extent and degree of human modifications to populations of plants and animals. The boundary is preliminarily proposed at 90 % BII, although the zone of uncertainty from 30 – 90% reflects the large gap in our understanding of the links between biodiversity intactness and Earth-system functioning. The BII has presently only been applied to Southern Africa but observations are that decreases in BII adequately capture increasing levels of ecosystem degradation (defined as land where the land-cover type has not changed but there is a persistent loss of productivity). They also estimated the *mean species abundance of original species* (MSA) at 84 % globally as an approximation of aggregated human impacts on the terrestrial biosphere but have not yet disaggregated this by functional groups or considered aquatic ecosystems. They write that in the long-term, the concept of biome integrity – the functioning and persistence of individual biomes – offers a promising and robust approach. GLOBIO-aquatic takes the concepts of the BII into aquatic ecosystems, but is limited in that the data on biodiversity is limited to mining of information from published papers that purposively presented a comparison between reference and present species abundance. Such papers proved to be scarce (Janse et al, 2015).

The Freshwater use indicator also involves two indicators. The first is the estimated '**maximum amount of blue water consumption'**, which compares human water use with the ecological flow requirements of rivers' at the global scale with the boundary set at 4000 km^3 yr^{-1} (Falkenmark, 1997). This is presently in the safe zone, with a global consumption value of ~2600km^3yr^{-1}. The second considers **environmental water flows (EWF)**, which operates similarly to the above but at the basin-scale, defining the boundary as lower estimate of the amount of water required in a river system to maintain a fair-to-good ecological state and avoid regime shifts in the functioning of flow-dependent ecosystems. This utilises the *variable monthly flow* (VMF) approach, which considers the EWF as a percentage of the *mean monthly flows* (MMF) for low, intermediate, and high flow periods, separately. Boundaries are set at 25–55, 40–70, and 55–85% of MMF for the low- , intermediate-, and high-flow regimes. Water use, at the basin-scale, was calculated from grid cell–specific estimates of agricultural, industrial, and domestic water withdrawals, based on observations and hydrological models. This regional assessment shows areas where ecological changes due to water use are beyond the zone of uncertainty.

For the biogeochemical flows sector, the authors consider anthropogenic influences on the ratios of elements in the environment, as this can drastically alter other Earth-systems, particularly biosphere integrity. They only consider N & P, as these are the most important nutrients affecting productivity and eutrophication/anoxia of aquatic systems. C is also important but already accounted for as an indicator for Climate Change. For N, they consider the level of '**Industrial and intentional biological fixation of N'** at the global level with the boundary set at 6.2 Tg yr^{-1} to prevent the eutrophication of aquatic eco-systems using the most stringent water quality criterion. This is presently at 150 Tg N yr^{-1}. For P, they consider the '**P flow from freshwater systems into the ocean'** as a global indicator of P pollution aimed to prevent a large-scale ocean anoxic event, with the boundary set at 11 Tg P yr^{-1}. This is presently at ~22 Tg P yr^{-1}. However, in the updated framework (Steffen *et al.*, 2015), they also included a regional indicator for P, as '**P flow from fertilizers to erodible soils'**, aimed to prevent the eutrophication of freshwater systems as a result of fertilizer application to erodible soils, which is the predominant source of P pollution with the boundary set at 6.2 Tg yr^{-1}. This presently stands at 14.2 Tg yr^{-1}. All three indicators suggest that human influence on biogeochemical systems is in the 'high risk' zone and likely to lead to widespread and irreversible environmental changes. However, the regional-level analyses of both P & N, show that the transgressions of these global boundaries is the result of fertilizer application to the world's croplands, amounting to a relatively small area of the world. In addition, this suggests that redistribution of fertilizers from areas where N & P are currently in excess to areas where soils area nutrient poor could both boost crop production and reduce transgression of the N & P boundaries.

Land-system change has clear implications to other Earth-systems. Originally, the indicator used to track this was '**percentage area converted to cropland**' (Rockström *et al.*, 2009) with the global boundary set at 15 %. However, the authors considered the 'biosphere integrity' boundary in the updated framework (Steffen *et al.*, 2015) to provide considerable constraint on the amount and patterns of land-use change across biomes. Therefore, they deemed this indicator redundant, and instead changed the focus to include indicators of land system changes more strongly linked to climate change, namely forests. Therefore, they include two forest cover indicators, the first is the '**percentage area of original global forest cover'**, which considers global changes, and the second the **'percentage area of potential forest cover for tropical, temperate and boreal forests',** biome-level changes of these three forest types which differ in their influence on climate. They add that these boundaries would almost certainly be met if the proposed biosphere integrity boundary of 90 % BII were respected.

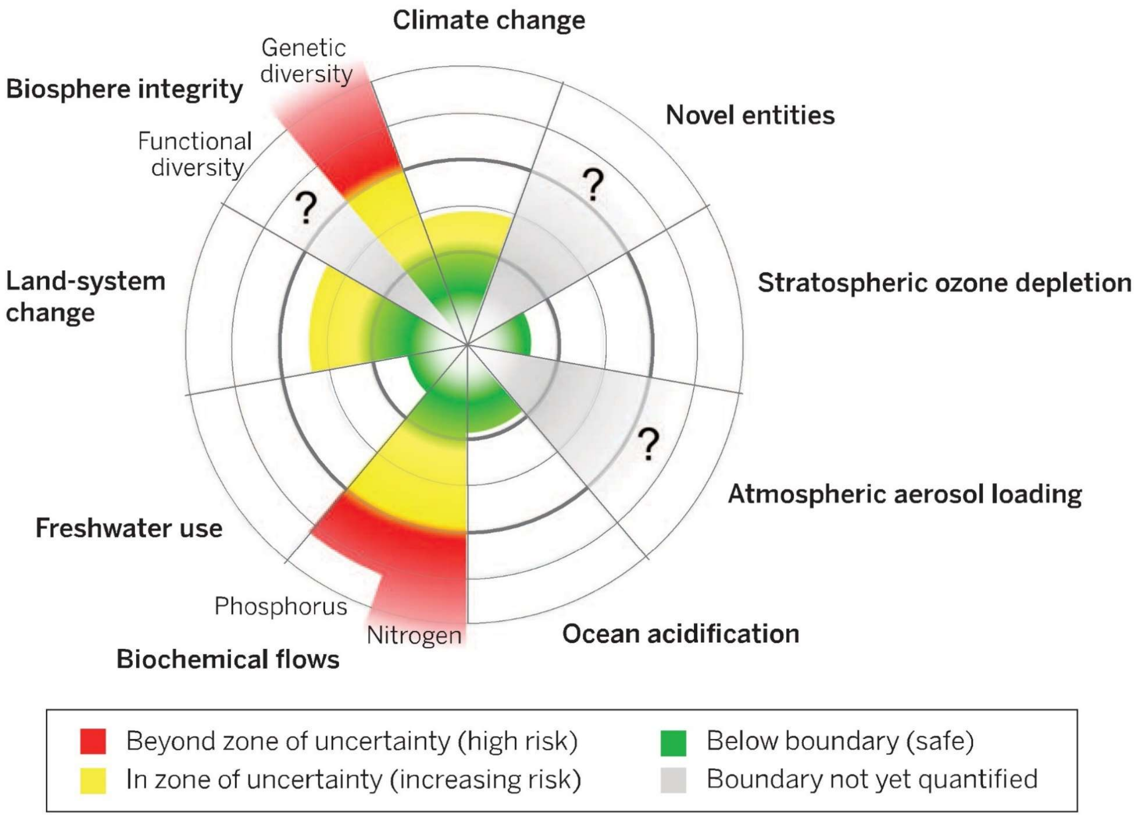
For Climate Change, the indicators include '**atmospheric CO_2 concentration'** with the boundary set at 350 ppm CO_2 but observed at 398.5 and '**energy imbalance at top-of-atmosphere'** with a boundary of +1.0 W m^{-2} and observed at 2.3 W m^{-2}.

A massive advantage of the Planetary Boundaries framework is that by considering only whether certain key variables are sustainable (below quantified boundaries above which change is irreversible) it reduces our reliance on detailed *in situ* monitoring systems and the definition of reference conditions. These are the core aspects of the most robust national/ regional-scale frameworks but the major inhibiting factor to their global application. The different perspective of the PB, with the identification and quantification of boundaries based on the best available scientific evidence, and minimal monitoring, makes it more suitable for a large-scale global assessment. The inclusion of a zone of uncertainty is also highly advantageous in accounting for knowledge gaps and intrinsic uncertainty, by the precautionary principle, making the final assessment of whether a boundary has been transgressed all the more robust and meaningful. The two-tiered approach, of defining boundaries both globally and regionally, is also highly advantageous in enabling the identification of regional impacts with global implications to environmental conditions. The best example being the global P & N imbalances, which are largely due to fertilizer application in only certain areas of the globe, thereby also revealing a potential solution to the problem, through regional redistribution of fertilizer use. The PB framework also makes more progress towards utilising large-scale biodiversity indicators than any other global-scale framework we have assessed, through the inclusion of extinction rates and the Biodiversity Intactness Index (BII), which is the first large-scale use of a functional diversity index. Although there is still much that needs to be done to create indices that are both practical and robust at broad scales and applicable to freshwater ecosystems, this offers a promising example of an appropriate method.

There are also several shortfalls to the PB framework. Firstly, in its current form, it is not down-scalable. Although the regional-level analyses in the 2015 version are a sure improvement from the 2009 assessment, these are still above the scale ideal for a global RH assessment (i.e., reach to basin). Secondly, there is also a lack of understanding about many of the tipping points, making the 'zones of uncertainty' particularly large. This is especially true for biosphere integrity as we have little idea of what level or types of biodiversity loss could trigger non-linear irreversible changes. Nevertheless, even with this uncertainty accounted for, this particular sector is clearly transgressed, and further research to quantify the boundary values could simply define by how much. There are also concerns over the representativeness of certain indicators, such as elements other than N & P in the biogeochemical sector and biomes other than forest in land-system change. Most the rest of the shortfalls are to do with individual indicator variables themselves. In terms of biotic integrity, the authors acknowledge the usefulness of the BII as a functional diversity index but lament its lack of application beyond Southern Africa. We would add to this, the fact that it is also only applied to terrestrial biomes. Similarly, the use of extinction rates as an indicator is less accurate than phylogenetic species variability as it considers only extinction at the species level and is subject to a time lag between impacts and extinction. However, fast progress in genetic extraction, analysis and phylogenetic are likely to change in the near future. The biogeochemical flows sector also neglects other important substances, although there is also potential to include these, whilst the indicators of P & N themselves are not fully comprehensive, neglecting other important sources of these pollutants, such as domestic sewage, in favour of fertilizers, which are the largest, but not the only source. For land-system changes, we strongly criticise the narrow focus of the indicators used, which only consider changes to the area of forest biomes, as this neglects the role that other biomes play in Earth-system processes, especially biosphere integrity. The argument by the authors is that the link between biosphere integrity and land-system change is so tight that the inclusion of both is redundant and that meeting the biosphere integrity target will inevitably also meet the target for land-use systems. Therefore, they intentionally only included forest, which they consider to be the most important for climate regulation. However, this reveals a bias of considering climate change more important than biotic integrity in the framework design.

There are several lessons we can learn from the PB framework to develop a global RH assessment. The first is the completely different approach of identifying and quantifying boundaries in natural systems (in this case, Earth-systems but one could imagine a similar approach at a regional level (i.e., basins or ecoregions)) that would lead to large-scale, irreversible change likely disastrous for humanity, and determining the risk of transgressing these boundaries. Although this faces challenges of limitations in the understanding of boundaries/ systems, it is highly beneficial to a large-scale assessment in the reduced reliance on comprehensive *in situ* data. Even with these limitations to understanding, however, the application of the precautionary principle (i.e., zones of uncertainty), make it relatively robust, as under such circumstance crossing the boundary of uncertainty almost certainly means the boundary has been crossed (e.g., biosphere integrity), although it may conceal potentially serious transgressions. The consideration of each process independently (instead of integration into a single index) is also valuable, as each is fundamentally different by nature with its own thresholds, potential impacts and management or mitigation options, so maintaining freshwater ecosystem components separate may be beneficial to foster and understanding of the impacts present and the potential management solutions.

In terms of the specific indicators used, the BII has high potential for application as a widespread indicator of biotic integrity of ecosystems, not reliant on extensive *in situ* sampling and has proven successful in terrestrial systems. The development of similar indices for aquatic ecosystems would thus be very useful. In addition, as noted by the authors, a biome-level approach to such indices (considering their innate biotic differences) would be especially useful for a global-level assessment. On the other hand, the strong relationship between land-use change and biodiversity loss suggests that it could be used as a proxy for biotic integrity. The use of environmental water flows (EWF) is already usable at the regional scale, so could be included as is in a global assessment to represent the 'water quantity' or 'hydrological' component.



2.2.1.6 Incident Threat Indices (ITI) to Human Water Security and River Biodiversity

The Incident Threat indices by (Vörösmarty *et al.*, 2010) quantifies the threats to freshwater resources using data on multiple stressors (i.e. driver indicators in DPSIR classes) from the perspective of a) human water security and b) biodiversity over a broad range of scales. It is based on the premise that integrated water management strategies depend on striking a balance between human water use and ecosystem protection. This requires an understanding of the spatial distribution of incident threats to human water security and biodiversity. The word 'incident' refers to exposure to a diverse array of stressors at a given location. This incorporates all major classes of anthropogenic stress and includes 23 geospatial drivers (with globally available information that has sufficient fidelity and spatial resolution) under four themes: **1) catchment disturbance** (cropland, impervious surfaces, livestock density, wetland disconnectivity), **2) pollution** (soil salinization, N loading, P loading, mercury loading, pesticide loading, sediment loading, organic loading, potential acidification, thermal stratification), **3) water resource development** (dam density, river fragmentation, consumptive water loss, human water stress, agricultural water stress, flow disruption), and **4) biotic factors** (non-native fishes (%), non-native fishes (no.), fishing pressure, aquaculture pressure). Making use of global high-resolution imagery, drivers were mapped onto a 30’ latitude/longitude grid. Driver loadings were then routed down river networks, accounting for new inputs and dilution or concentration from tributary mixing, based on changes in river discharges from precipitation and abstraction. Driver values for grid cells were then standardised using a cumulative density function to between 0 (no stress) and 1 (maximum expression of stressor) to reflect the relative stressor level on each cell across the globe. The scaled drivers were then combined into overall incident threat indices for human water security and biodiversity perspectives using two-tiered relative weighting matrices derived from expert assessment (first among drivers, then themes). For human water security, they then did the same procedure for drivers alleviating human water security, including supply stabilization, improved water services and access to waterways (i.e., water-related capital and engineering investments). Subtracting this factor from the original threat index and re-scaling globally produced a second map of ‘‘adjusted human water security threat’. The adjusted incident threat is often much lower in highly developed regions with high incidence threat but large investments in water infrastructure (i.e., much of North America and Europe), relative to developing countries. Unfortunately, there is insufficient information available for a similarly meaningful adjusted index to be computed for biodiversity. These values can then be mapped and are applicable over a broad range of scales from the local to the global level and enables the identification of threats at sub-national scales. The results show very strong relationships between incidence threats and ecological state indicators by independent studies (e.g. NARS).

Major advantages of the IHI framework are that it is applicable across spatial scales. The standardisation of stressor data also limits the impacts of uncertainty/ inaccuracies in stressor data. Major shortfalls, however, include the limited spatial resolution, which prevents its application to smaller streams (Strahler order ≤ 5; scale ≤ 1:62,500). However, with new advances in satellite technology and future missions, this can be expected to improve. The threats considered may also not be fully comprehensive, excluding important factors such as mining, inter-basin water transfers, and pollution by pharmaceutical compounds etc. In addition, the inability to compute a globally meaningful estimate of adjusted biodiversity threat from investments to counter biodiversity loss limits the meaningfulness of the index.

The major lesson for application to a global RH framework, is that it is possible to use drivers of ecological health to measure threats to it across spatial scales up to the global level. The readily available nature of the relevant data (EO) certainly makes data acquisition easier. However, (Vörösmarty *et al.*, 2010) also showed the difference that investments and technological developments can make to reducing the threats to water resources (for human water security), creating an adjusted index that is more strongly representative of the actual situation. However, once again, the lack of available data restricted the application of this to the biodiversity threat index, making it less representative. Nevertheless, it may even be possible to use the Incidence Threat Indices themselves as driver indices, alongside other ‘ecological state’ indicators, in the calculation of an overall index of RH for application at the global level.

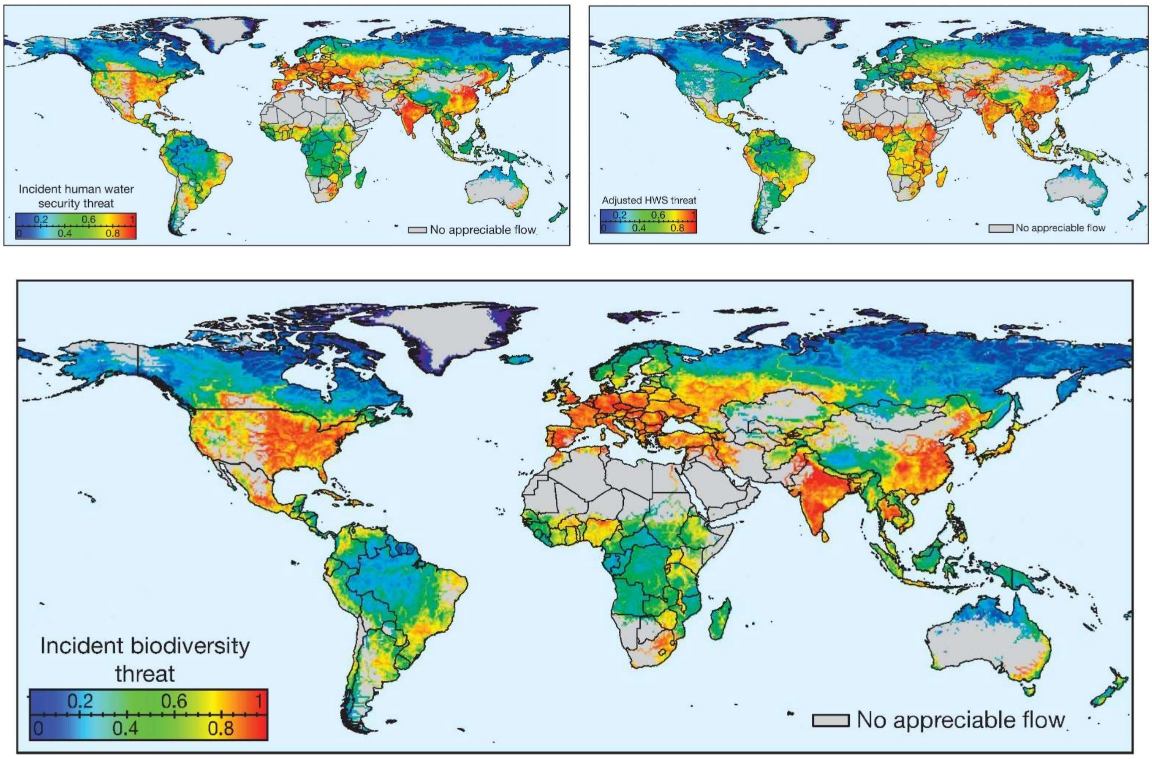


Figure 2.7: The global distribution of incident threat to human water security (above left), adjusted human water security (above right) and biodiversity (below) (adapted from Vörösmarty et al. (2010)).

2.2.1.7 System of environmental economic accounting - ecosystem accounts (SEEA EA)

The System of Environmental Economic Accounting – Ecosystem Accounting (SEEA EA) framework is part of the wider System of Environmental Economic Accounting (SEEA) that aims to generate understanding over the links between the economy and environment and to describe changes in the stocks of environmental assets, expressed in monetary terms. The SEEA was adopted as an international standard for environmental-economic accounting by the UN Statistical Commission (United Nations *et al.*, 2014) meant for use at the level of countries or geopolitical regions. It integrates economic, environmental and social data, providing a more comprehensive view of the relationships between them. The key feature of the SEEA EA is thus that it allows for ecosystem assets and contributions to society to be expressed monetarily, providing useful information to decision-makers regarding the importance of environmental assets. This provides a means to monitor pressures exerted by the economy on the environment and *vice versa*; how the economy responds to conservation and resource management. It consists of three parts. Part 1 is the Central Framework (United Nations *et al.*, 2014), which provides agreed upon concepts, definitions, classifications, accounting rules and tables for producing internationally comparable statistics and accounts. It does not propose specific indicators, opting instead for a more holistic multi-purpose information system that can be analysed in various ways, so leaving the choice of indicators entirely to the country/ region in question. This generates a wide range of statistics, accounts, and indicators with variable potential applications, making it flexible to the needs of different countries, whilst providing a common framework to make data between countries comparable, facilitating better-informed decision-making. The thematic areas considered include, agriculture, forestry and fisheries, air emissions accounts, energy, environmental activity accounts, ecosystem accounts, land accounts, material flow accounts and water. Data are reported using accounting tables but also maps, as the benefits arising from ecosystems are dependent on their position in a landscape relative to the beneficiaries. Part 2 is the SEEA Ecosystem Accounting, which we deal with in more depth below. Lastly, Part 3: The SEEA Applications and Extensions, illustrates to the users of the frameworks how the information can be used in decision-making, policy review and formulation, analysis and research.

The SEEA Ecosystem Accounting (SEEA EA) is the part of most interest to us and consists of five core account types. 1) ***Ecosystem extent accounts*** provide information on the extent of different ecosystem types within an area, providing an important starting point for ecosystem accounting and aligning with targets of environmental agreements (SDGs etc.). 2) ***Ecosystem Condition accounts*** measure "the overall quality of an ecosystem asset and captures, in a set of key indicators, the state or functioning of the ecosystem in relation to both its naturalness and its potential to supply ecosystem services" (United Nations, 2018a). The accounts are based on condition maps built from spatial data (although the underlying source of this data may be *in situ* in nature). Like most regional frameworks, conditions are quantified relative to ecosystem specific reference levels for representation on a common scale. This can then be categorised into qualitative descriptor categories for simple communication of results (see examples in Figure 2.9). They also specifically support the utilisation of biodiversity indicators as a component of ecosystem condition accounts, mentioning as examples the British Woodland Butterfly and Woodland Bird Indices that indicate British woodland conditions at the national scale. However, they do not provide any indication of indicators of freshwater biodiversity. 3) **Ecosystem Services Accounts** measure "the supply of ecosystem services and their corresponding use and beneficiaries, classified by economic sectors in national accounts". These can be compiled in both physical and monetary terms (United Nations, 2018a). Ecosystem services are defined as "the contributions of ecosystems to benefits used in economic and other human activity." These generally fall within the three categories of provisioning (e.g. material and energy contribution), regulating (e.g. hydraulic, geomorphological, biochemical or biological processes) or cultural services (spiritual, psychological and recreational). 4) **Monetary Ecosystem Asset Accounts "**record the monetary value of opening and closing stocks of ecosystems assets within an ecosystem accounting area, including additions and reductions to these stocks". Ecosystem assets are calculated by capitalising the flow of ecosystem services (determined above) over a given period. This allows the calculation of natural capital asset values for a region and enables the more comprehensive assessment of its wealth that includes natural, financial, human and social capital. Indicators such as wealth per capita (and its change over time) thus provide an indication of whether a country is developing sustainably. The links between these core account types are shown in Figure 2.8. However, the framework also supports the development of 5) **Thematic Accounts** that measure stocks and changes of specific policy-related themes, such as specific species or habitats, biodiversity or carbon capture etc.

During the piloting phase of the framework development, experimental accounts of each account type were compiled in over 40 countries (Figure 2.8) for a wide variety of objectives. These included thematic carbon accounts for peatlands in Indonesia and a species account of the Shea Nut tree in Uganda. An ecosystem services account for the EU in 2012 covering six services: crop provision, timber provision, carbon storage, pollination, flood control and nature-based recreation, which together accounted for a total of €124.87 billion (United Nations, 2018), see Figure 0.8, below. However, the most relevant to this review was the National River Ecosystem Account for South Africa (Nel and Driver, 2015). In this case, rivers were chosen due to the availability of comprehensive national datasets with two national surveys carried out by the Department of Water and Sanitation (DWS) in 1999 and 2011, partly using the EcoClassification framework mentioned above. For the basal extent account, the authors measured river lengths classified by water management area (WMAs), longitudinal zone, and ecoregions. For the ecosystem condition accounts, they then used four ecological condition indicators, flow, water quality, instream and riparian conditions. Data on these variables were gathered at the quaternary scale for main rivers in 1999 and quinary catchment scales for main rivers and tributaries in 2011. Therefore, although originally gathered at the site scale, conditions for each could modelled and mapped for the whole country. They then integrated values of these four indicators into overall ecological condition categories and an ecological condition index ((see Figure 2.9 top). What they found was no change in river lengths but a 10 % decline in the overall condition between 1999 and 2011, providing useful information for the National Water and Sanitation Master Plan.

A strength of the SEEA EA framework is its flexibility, as it can be used to assess the condition of virtually any ecological aspects in virtually any context. The mapping of results also makes it applicable at various scales within the country or region where it is applied. However, the lack of identification of the core components of ecosystem health for any ecosystem, including freshwater ecosystems, means that it lacks consistency at the global level if one is interested in a single aspect (e.g., freshwater health), as the indicators chosen can be vastly different between countries, making comparison problematic. It is also highly dependent on data availability so would have limited applicability in data poor regions. For example, the South African Rivers Account was only carried out due to the availability of this data, even though it still left out the biological component due to it being less comprehensively sampled, so similar accounts would not be possible in countries without such data.

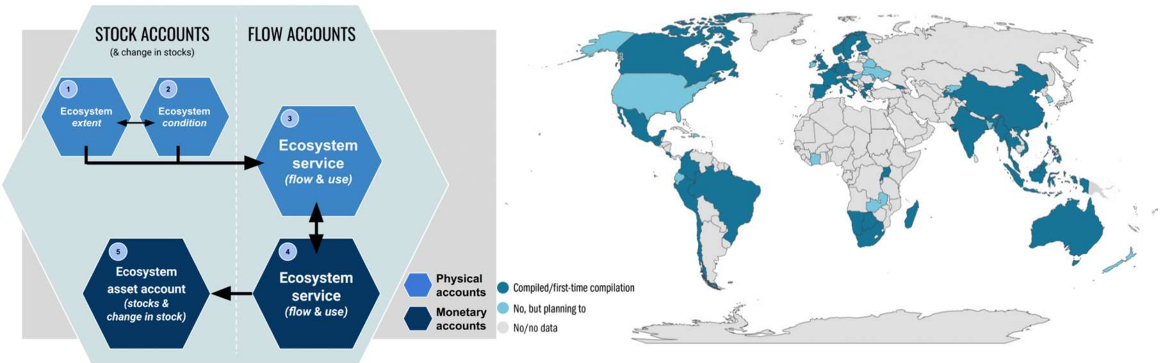


Figure 2.8: Left, links between SEEA account types (sourced from (ONU - The Committee of Experts on Environmental-Economic Accounting, 2021)*. Right, map of countries compiling or planning to compile SEEA Ecosystem Accounts.*

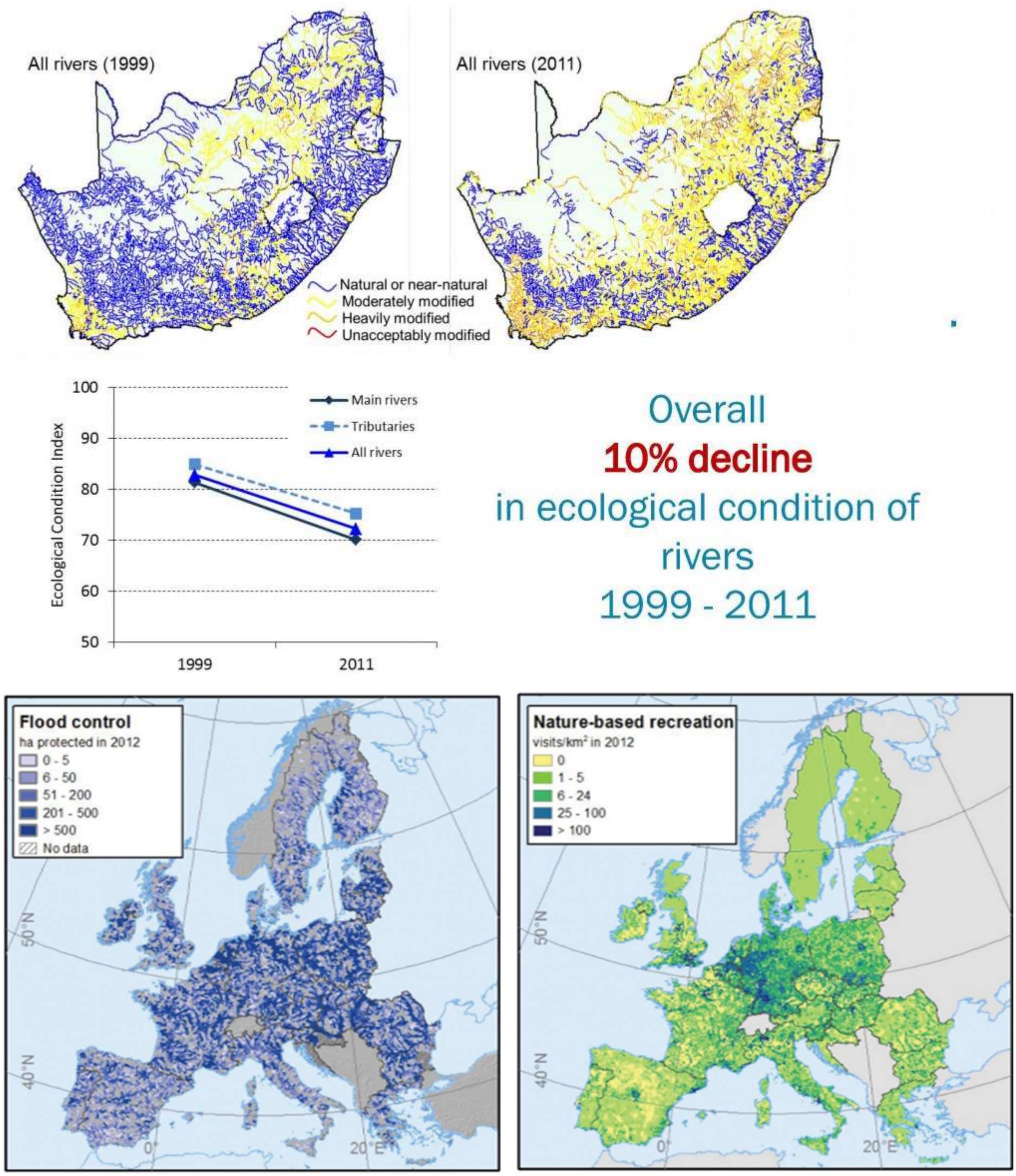


Figure 2.9: Top, change in national ecological river conditions in South Africa between 1999 and 2011 in the National River Ecosystem Account in South Africa (Source: Presentation by Mandy Driver, SANBI at Third Forum on Natural Capital Accounting for Policy Decisions, November 2018 in Nel, J.L. & Driver, A. 2015). Below, maps of two ecosystem service (flood control and nature-based recreation) across the EU in 2012 for calculation of the EU ecosystem services account (Source: see United Nations (2018)*).*

2.2.1.8 The Environmental Performance Index (EPI)

The EPI ranks countries' performance on high priority environmental issues in relation to a) protection of human health and b) protection of ecosystems. It provides a common framework to compare countries' performances and to track changes in their performance over time, in a similar way to the SDGs. This includes scores for 9 issues comprised of > 20 indicators, which measure the country's proximity to meeting internationally established targets, or, in the absence of targets, relative to one another. Relevant to us, are the inclusion of ‘Water Resources’ and ‘Biodiversity & Habitat’. To calculate the EPI, raw data are first standardised by population, land area, GDP etc, and normalised by statistical transformation. Performance indicators are then calculated using a ‘proximity-to-target’ method by assessing how close a country’s measured data are to identified policy targets. Targets are high performance benchmarks defined by policy goals or scientific thresholds (Srebotnjak *et al.*, 2012). Scores are then scaled from 0 to 100 with 0 being the farthest from the target and 100 the closest (see Section 3.2.6 below). The score for each issue is then calculated as the weighted-average of indicators, according to their relevance and reliability, and the overall scores for ‘ecosystem vitality’ and ‘environmental health’ as the weighted-average of issue scores. In the initial versions of the EPI (2008 & 2010), the performance of ‘water resources’ was measured using a globally-relevant Water Quality Index (WATQI) (see section 3.4.1.7 below). However, this was excluded in the 2016 EPI due to concerns over gaps and biases in the data (Hsu *et al.*, 2016). It was replaced by ‘wastewater treatment’, which tracks the proportion of wastewater that is treated before release into the environment (weighted by the human population), as wastewater is one of the main drivers of poor water quality worldwide with readily available data for most countries, providing a good proxy for water quality. However, although it is a major contributor to poor water quality, it overlooks all other sources of poor water quality and other aspects of ecosystem health not related to water quality (e.g., water quantity) so is not fully representative of actual conditions. As for ‘biodiversity and habitat’, the indicators included are specific to terrestrial and marine ecosystems and exclude freshwater ecosystems. Furthermore, the indicators used, including species protection (global and national), terrestrial biome protection (global and national) and marine protected areas, measure the proportion of taxa or habitats protected, providing little information on actual ecosystem conditions. Therefore, the EPI themselves recognise freshwater quality, as well as related issues of species loss, and wetlands loss as key gaps in monitoring of environmental issues (Hsu *et al.*, 2016). The major advantage of using ‘wastewater treated’ as a proxy of water resource conditions is that it is based on readily available data and is one of the key drivers of poor water quality globally. It is also directly related to SDG target 6.3.1 to ‘halve untreated wastewater by 2030’. However, it has many shortfalls and as mentioned above overlooks all other sources of water quality impacts and all other components of ecosystem health providing only a very narrow definition of water resource conditions.

Lessons from the EPI include being an example of a broad-scope dashboard indicator, that is little value for RH reporting as water resources are represented by mainly wastewater. Biodiversity is dominated by terrestrial indicators with no direct measure of freshwater ecosystems.

Figure 2.10: The 2016 EPI Framework showing the 9 issues and > 20 indicators in relation to ecosystem vitality and environmental health. Note the use of 'water treatment' as the only indicator of the condition of water resources.

2.3 Bending the curve of global freshwater biodiversity loss: an emergency recovery plan

Freshwater biodiversity is declining faster than in terrestrial or marine systems. This is partly because traditional conservation actions are less effective or relevant to freshwater ecosystems, which integrate impacts from multiple sources. Furthermore, coordinated action to reverse the decline is lacking. Therefore, in 2019, led by the WWF, a panel of experts set out to define the priority actions needed to bend the curve of freshwater biodiversity loss. The result was an ambitious but pragmatic Emergency Recovery Plan (ERP) for global freshwater biodiversity based on proven methods that involves six points (Tickner *et al.*, 2020). These include:

1. Allowing rivers to flow more naturally (accelerate implementation of environmental flows).
2. Reducing pollution to improve water quality
3. Protecting and restoring critical habitats, particularly wetlands.
4. Managing the exploitation of freshwater ecosystem resources, especially species and riverine aggregates (i.e., ending overfishing and unsustainable sand mining).

5. Controlling non-native species invasions, and
6. Safeguarding and restoring connectivity

These are aligned with the Convention on Biological Diversity (CBD) Aichi targets and SDGs with the recommended targets and indicators given for their incorporation by both agreements.

Although the ERP does not involve the calculation of an index, it highlights important considerations for the development of a global RH assessment. The major lesson from the ERP is that it clearly identifies the main issues threatening freshwater biodiversity, and hence ecosystems, and the actions that must be implemented to reverse the decline. A truly robust and flexible global RH assessment could therefore include indicators of these factors: flow regulation, pollution, overfishing, wetland protection, invasive species, and connectivity. Although these may be driver (not state) indicators, they are vital to be able to understand the links between human actions and ecological integrity, especially to inform management actions to 'bend the curve' of freshwater biodiversity loss.

3 CONCEPTUAL REVIEW OF KEY ATTRIBUTES AND BEST APPROACHES OF SUCCESSFUL FRESHWATER ECOLOGICAL OR RIVER HEALTH ASSESSMENT FRAMEWORKS

3.1 Key attributes of successful frameworks

In this section, we summarise the key attributes of successful frameworks from the lessons learned in the review of some of the most widely used regional and global frameworks, above. This is heavily informed by the FBEHF report and IECA manual, which undertook similar concept reviews, but to develop frameworks at the regional scale. The key attributes (Clapcott *et al.*, 2018) considered here include:

- **Consistency** - understanding of what constitutes ecosystem health and how to measure it
- **Representativeness** - includes measurement of a full range of the core components of ecosystem health
- **Robustness** - rigorous science with justified selection of components and indicator variables based on empirical evidence
- **Informativeness** - easily understood
- **Flexibility** - can be meaningfully applied across a wide range of waterbodies
- **Scalability** - application remains consistent across spatial scales
- **Feasibility** - not highly demanding on time, labour or money

3.1.1 Consistency

A consistent framework is one that provides a clear understanding of what constitutes ecosystem health and how to measure it (Clapcott *et al.*, 2018). Consistency is a core feature of all the national-scale frameworks assessed here and some of the more successful global frameworks, such as the FHI and ITI. Such frameworks include the clear definition of terms and objectives and the standardisation of protocols on how to collect, analyse and report data, to ensure comparability between different areas. This also facilitates cooperation between jurisdictions in the assessment of freshwater ecosystem health. At the very least, consistent frameworks define what components must be measured to define the RH (see Section 3.2.3) but leave the selection of indicators and methods employed open. To make results comparable, this is generally accompanied by guidance on the standardisation/ harmonisation of data (scaling from 0-1 or 0-100). As we will see below (Section

3.2.6.1:), this often involves comparison to reference conditions. However, there are strong criticisms of the limitations to the 'reference conditions' approach, especially at the global scale, given difficulties in implementation and shifting baselines but alternatives, such as the comparison of relative conditions, proximity to targets/ boundaries or analysis of trends, do exist. An example of a high-level consistent framework, the WFD, which requires members states to achieve 'good ecological status' (clearly defined) across all 'natural' waterbodies and offers clear guidance on the cross-calibration (between methods) and inter-calibration (between ecological class boundaries) of data. For the most part, this has enabled cross-calibration between methods and inter-calibration between ecological class boundaries between European freshwater ecosystems across national boundaries (Poikane *et al.*, 2014). At their most prescriptive, consistent frameworks define the precise indicators and method required. Examples of prescriptive frameworks include NARS and the EcoStatus Reports, which require the standardised collection of given indicators for the multiple components considered by each. Although, in both cases, this allowed for consistent assessments at the national level, it makes cross-calibration with other countries/ regions difficult and restricts the inclusion of new technologies as they emerge. The higher-level approach is thus preferable for a global framework, which must balance the needs and abilities of all countries involved. ***For the global RH assessment to be consistent, it must therefore have a clear purpose, agreed-upon terminology, and support methods of data standardisation.***

3.1.2 Representativeness

A representative framework is one that includes the measurement of a full range of the core components of ecosystem health for a fully integrative assessment (Clapcott *et al.*, 2018). What constitute the 'core components' is addressed in Section 3.2.3.

The links between components are important to understand for a robust and informative framework. They are also very useful for the identification of key components and appropriate indicators by revealing key nodes within the network. Examples of frameworks using conceptual models of freshwater ecosystems include the REMP and SRA.

Key components for assessment of river health, generally include biological, physical and chemical components. However, there may be benefits to considering other values in a specific region or context. A framework's representativeness is relevant at two stages of the design and application, namely in the 1) 'sampling network design', and 2) 'data aggregation and integration'. In the first case, the sampling network must ensure a balanced spatial representation in the RH assessment. Most national frameworks take one of two approaches to this, either sites are selected at random (e.g., NARS) or they are selected based on a stratified risk-based approach that considers e.g., access and ecoregion representativity. However, indicators operate at different spatial and temporal scales. Therefore, a truly representative framework should allow for variability in the sampling network design such that different indicators can be measured at different scales (i.e., integration of *in situ*, earth observation and modelled data). The differences in scale can then be addressed during 'data aggregation and integration' (see Section 3.2.6.3).

An example of a representative framework is the IECA, which recommends the core components to be measured and allows for the aggregation of individual indicator variable values at different spatial scales (using averaging, modelling or summing depending on the units involved) before their integration (using mathematical and non-mathematical rules) into composite indicators of each component. ***For the global RH assessment to be representative, it must therefore clearly define the core ecological health components to be assessed and support the standardisation of methods of data aggregation and integration.***

3.1.3 Robustness

A robust framework is one informed by rigorous science with justified selection of components and indicator variables based on empirical evidence (Clapcott *et al.*, 2018). This is vital for facilitating public and government trust in its application. Transparency over the definition of concepts and the scientific knowledge used to inform the development of a framework are thus essential. Some of the most successful frameworks (e.g., WFD, NARS, EcoStatus) have been developed through partnerships between scientists, resource managers, government, and multinational (governmental and non-governmental) organisations bolstering transparency and trust. Another consistent feature supporting the longevity of the successful frameworks is the commitment by all the parties involved from the early stages of framework development, as this creates shared-knowledge and supports adaptive management programmes (Bunn *et al.*, 2010). This also requires adequate resourcing during the development, testing and implementation phases of the framework. Many frameworks also add a 'confidence' value to support transparency over the robustness of the reported results. ***For the global RH assessment, the selection of components and indicators as well as conceptual and analytical methods employed must therefore be empirically supported.*** It is also recommended that the development of the framework involve a partnership between scientists, resource managers, governments and international organisations, should support early commitment by the parties involved and source adequate funding for the development, testing and implementation of the framework.

3.1.4 Informativeness

An informative framework is easily understood (Clapcott *et al.*, 2018). It should provide the necessary context to interpret information. This is best achieved using conceptual models, illustrating ecosystem components and the mechanistic links between drivers and indicator variables and trends e.g. models used by the SRA and IECA. This is crucial to contextualise the indicators and approaches used and make explicit the link between human actions and ecological responses expressed by the indicator variables. The development of conceptual diagrams can sometimes also help in the identification of potential indicators during framework development (Clapcott *et al.*, 2018). Another aspect of an informative framework is that knowledge gaps (missing data & uncertainty) are made explicit. This is expressed during the reporting phase and the use of a tiered or hierarchical stacking approach to reporting supports an informative framework and facilitates reporting at various levels. This involves going from detailed reporting of raw data and assessments of individual indicators and components at the basal level to synthesised reports on integrated scores at sites or river reaches or greater (See Figure 3.8). ***For a global RH assessment, the development of conceptual diagrams with explicit links between drivers and indicators and the adoption of tiered or hierarchical reporting (with transparency regarding knowledge gaps) is encouraged.***

3.1.5 Flexibility

A flexible framework is one that suits varied application and can be meaningfully applied across a wide range of waterbodies (Clapcott *et al.*, 2018). This is particularly relevant to application of the framework to *different ecosystem types* as they differ in the relative importance of different ecosystem components, the specific component indicators that are relevant and their values under natural (reference) conditions. For example, the assessment of the hydrological component might involve flow volumes in rivers but hydraulic residence times in lakes, but between river systems, the natural flow values differ naturally. However, the importance of the hydrological component itself also varies between river systems and changes to hydrology in some may be more impactful to the biota and ecological conditions than in others. Relevant indicators and their desired values can also differ under *different management contexts* (e.g. water quality targets vary depending on whether a waterbody is used for human water consumption, irrigation or for the general environment (Department of Water Affairs, 2013)) and according to the *information available*. The later point is particularly relevant to a global framework as information is unevenly distributed across regions. For example, reference-based biotic indices are only available in a limited number of countries (usually wealthier countries with functioning water monitoring programmes) and would be ideal to include in a global assessment, but alternatives are required for countries where such indices or indeed monitoring programmes are not widely implemented. However, the framework must still allow for comparability between waterbodies under different contexts of ecoregions, management and information availability. This is achieved by cross- and inter-calibration between different indicators through the standardisation/ harmonisation of data to a set value scale, as mentioned in Section 3.1.1: Consistency. This also facilitates the inclusion of new indicators as they arise. To account for variability in the relative importance of different components within a waterbody and different indicators to the components, most national-scale frameworks utilise a weighting method, which adds flexibility to a framework. *Frameworks with variable or 'toolbox' approaches that include a mix of compulsory and optional indicators, recognising the variability in data availability and institutional capacity between regions (along with intrinsic natural and management context differences), that weight variables and/ or components differently relative to their importance, and make use of data* cross- and inter-calibration *are thus some of the most successful.* Flexible national-scale frameworks include the WFD, IECCA, FBEHF, and RHI, which leave indicator selection up to jurisdictions, weight indicators and components differently and provide guidance on cross-calibration. The two most flexible global frameworks are also two of the most utilised, the FHI and SEEA, as they allow different indicators to be measured for different components (although with a broader focus than just ecosystem conditions) but the calibration of data still allows for comparability between distant regions. ***For the global RH assessment, the protocol should thus make use of a 'toolbox' approach, which recognises that all freshwater bodies consist of the same core components but allows for the variable selection of indicators under different contexts and supports the use of weighting of indicators and components and data cross- and inter-calibration to compare different environment types.*** This is possibly the most important of the attribute of widely used frameworks for the assessment of RH.

3.1.6 Scalability

A scalable framework is one whose application remains consistent across spatial scales (from river reach to sub-basin, basin, regional, national, and international scales). Of importance for consideration here, is the impact of different spatial scales on the relevance of indicators used, as each responds differently at different scales, some respond more to catchment-scale effects and others to local effects. For example, migratory fish are strongly affected by basin-scale effects, such as loss of connectivity, whilst primary production is more affected by local factors (Sheldon *et al.*, 2012). This may necessitate the use of different indicators at different scales. This is generally informed by a robust scientific understanding (backed by literature) of the mechanisms by which drivers affect indicators. Conceptual models are thus highly useful in this regard and may help to illustrate the response of indicators at different scales. Once the links between drivers and indicators across scales are understood, this may then inform how measures of ecosystem health are combined for an overall assessment of ecological health. Crucial to enable this, is the standardisation of data (grading to a single scale). The use of different indicators also thus necessitates the complexity of reporting, which would be facilitated by a hierarchical design (see Figure 3.8, above). Nevertheless, most frameworks regard the watershed-scale, the most appropriate for monitoring and reporting ecological conditions as it is both the scale at which most ecological processes occur and at which most anthropogenic pressures impact river health and can thus be mitigated. It is the preferred scale of the WFD, NARS, REMP, IECA and SRA. At the global scale, the FHI also operates at the river-basin scale. ***For the global RH assessment, the framework should therefore allow variable indicator selection by scale and include a hierarchical reporting system.***

3.1.7 Feasibility

A practical framework is one that is not highly demanding on time, labour or money. The lack of resource, institutional capacity and conflicts in many countries are prohibitive to them having comprehensive *in-situ* monitoring systems. This includes some of the world's most important regions in terms of water security. Even in wealthy countries, the costs associated with a detailed monitoring framework are prohibitive, especially when the gains are not perceived to be large, as was the case with the scrapping of the NRHP in Australia. As another example, the protocols for the highest level of assessment in the South African REMP has never even been fully completed as they are extremely labour and financially intensive so less-comprehensive protocols are always chosen in preference. ***For the global RH assessment, the framework should strive to be effective in terms of labour, time and financial costs. It should rely as little as possible on*** **in-situ** ***data acquisition and the need for reference site monitoring and use EO or modelled datasets wherever possible.***

3.2 Characteristics of successful frameworks

3.2.1 A clear policy-driven purpose

A clear purpose is the foundational element of any framework as it influences decisions to all subsequent aspects of the framework development, from the definition of terms, to choice of data acquisition methods, processing and reporting. It is generally a short phrase summarising the goal(s) of the framework. The purpose is thus generally directed by policy. For example, the WFD framework, is based on the requirement of 'good ecological status' (clearly defined) across all 'natural' waterbodies in EU member states and the protocols of the framework are written into the directive. Generally, however, frameworks emerge out of policy requirement for a standardised monitoring framework, for example the REMP EcoStatus Reports, which were developed in response to the South African National Water Act, and the NARS, in response to the adoption of the Clean Water Act. A link to policy also incentivises the uptake of a framework and encourages cooperation between jurisdictional authorities. For a global RH assessment, a goal(s) that includes the objectives of the relevant international ambitions would thus be most beneficial and should align with the targets of the SDGs, CBD Post-2020, IPBES, etc.

Table 3.1: Purpose of the reviewed regional and global frameworks.

Framework	Purpose	Objectives	Policy basis
	Regional		
WFD	"…to prevent further deterioration and enhance protection and improvement of aquatic ecosystems, whilst promoting sustainable water use and maintaining socioeconomic systems."	"…to achieve at least 'good' ecological status for all inland surface waters (natural and artificial) across all members states by 2015 and at the latest 2027." The Directive also provides objectives for 1) the characterisation of river basins, and 2) monitoring and 3) classification of ecological status.	Water Framework Directive 2000
NARS	"…restoring and maintaining the chemical, physical, and biological integrity of the waters."	"…to assess the quality of the nation's coastal waters, lakes and reservoirs, rivers and streams and wetlands using a statistical survey design."	USA Clean Water Act 1977

Framework	Purpose	Objectives	Policy basis
REMP	"…to gain insights and understanding into the causes and sources of the deviation of the present ecological state of biophysical attributes from the reference condition. This provides the information needed to derive desirable and attainable future ecological objectives for the river."	1) Determine reference conditions for each component. 2) Determine the Present Ecological State for each component as well as for the EcoStatus. The EcoStatus refers to the integration of physical changes by the biota and as reflected by biological responses. 3) Determine the trend (i.e., moving towards or away from the reference condition) for each component as well as for the EcoStatus. 4) Determine causes for the PES and whether these are flow or non-flow related. 5) Determine the Ecological Importance and Sensitivity (EIS) of the biota and habitat. 6) Considering the PES and the EIS, suggest a realistic and practically attainable Recommended Ecological Category (REC) for each component as well as for the EcoStatus. 7) Determine alternative Ecological Categories (ECs) for each component as well as for the EcoStatus for the purposes of providing various scenarios	South African National Water Act, (Act No 36 of 1998)
RHI	"…to establish a standard quantitative approach applicable nationwide to assess both freshwater health and the effectiveness of the RCS "		Chinese National Water Standards, River Chief System (2016)
NRHP	"…to provide the information needed to reverse the degradation of our inland waters."	1) Provide a sound information base on which to establish environmental flows; *2) Undertake a comprehensive **assessment** of the health of inland waters, identify key areas for the maintenance of aquatic and riparian health and biodiversity, and identify stressed inland waters;* 3) Consolidate and apply techniques for improving the health of inland waters, particularly those identified as stressed; 4) Develop community, industry, and management expertise in sustainable water resources management and raise awareness of environmental health issues and needs of our rivers.	Australian National River Health Program
IECA	"…primarily to assess the condition of aquatic ecosystems within the defined assessment unit but may also serve other purposes, such as: • A part of a broader project involving high ecological value aquatic ecosystems; • To establish a benchmark of condition against which change can be assessed;		Australian National Water Initiative (NWI)

Framework	Purpose	Objectives	Policy basis
	• To determine changes in condition over time; • To fulfil and/or set specific planning or reporting requirements; or • To inform a management intervention program		
MIF	The mission of the MRC is to “promote and coordinate sustainable management and development of water and related resources for the mutual benefits of the lower Mekong countries and the people's well-being”. The purpose of the Indicator Frameworks is “…to provide an overall picture of the Mekong River Basin in terms of its ecological health and the social and economic circumstances of its people, and the degree to which the cooperation between riparian countries envisaged under the 1995 Mekong Agreement is enhancing these conditions”.	To deliver outcomes in four key areas: 1) Enhancement of national plans, projects and resources from basin-wide perspectives; 2) Strengthening of regional cooperation; 3) ***Better monitoring and communication of the basin conditions (Member Countries strengthen basin-wide monitoring, forecasting, impact assessment and dissemination of results for better decision-making)***; and 4) Leaner River Basin Organisation. The purpose of the SOB reports is “…to assesses conditions within the basin and the impacts, both positive and negative, that development and use of the water and related natural resources are having.”	Mekong River Commission (MRC) 1995
	Global		
SDG 6	ensuring global water access and safety by 2030 by investing in adequate infrastructure and protecting and restoring water-related ecosystems	Through its targets, to ensure safe water access and sanitation services. To reduce wastewater impacts and protect water quality as well as water quantity. To ensure productive use of water resources, implement good governance through IWRM at national and transboundary scales, and also to protect water-related ecosystems.	UN Agenda 2030
CBD Pos2020	Sets out an ambitious plan to implement broad-based action to bring about a transformation in society’s relationship with biodiversity, ensuring that by 2050 the shared vision of ‘living in harmony with nature’ is fulfilled.	Ensure that at least 30 per cent globally of land areas and of sea areas, especially areas of particular importance for biodiversity and its contributions to people, are conserved through effectively and equitably managed, ecologically representative and well-connected systems of protected areas and other effective area based conservation measures, and integrated into the wider landscapes and seascapes. Prevent or reduce the rate of introduction and establishment of invasive alien species by50%,and control or eradicate such species to eliminate or reduce their impacts. Reduce nutrients lost to the environment by at least half, pesticides by at least two thirds, and eliminate discharge of plastic waste. Use ecosystem-based approaches to contribute to mitigation and adaptation to climate change. Redirect, repurpose, reform or eliminate incentives harmful for biodiversity. Increase financial resources for biodiversity.	UN CBD

Framework	Purpose	Objectives	Policy basis
FHI	This framework helps to operationalize a truly integrated approach to water resource management by recognizing the interplay between governance, stakeholders, freshwater ecosystems and the services they provide.	Includes indicators of Ecosystem Vitality, Ecosystem Services, Governance and Stakeholders.	Vollmer et al 2018
PB	Define a "safe operating space for humanity" as a precondition for sustainable development.	Describe the tipping points or boundaries of nine Earth System Processes that include biodiversity, water, biogeochemical, chemical pollution and climate change, all of which connect to RH.	Rockstrom et al, 2009
ITI	Global-scale analysis of threats to freshwater that considers human water security and biodiversity perspectives simultaneously within a spatial accounting framework.	IWRM strategies depend on striking a balance between human resource use and ecosystem protection. To test whether this objective can be advanced globally, ITI maps Incident Threats to human water security and biodiversity.	Vorosmarty et al, 2010

3.2.2 Definitions of 'freshwater ecosystem / river health' and other terms

The clear definition of terms is a key feature of consistent frameworks. Most important is what is meant by "freshwater *health*" or more specifically "freshwater *ecosystem health"*, although the term "health" is used interchangeably with *"conditions"*, "*quality*" or "*status*" (see Table 3.2 below).

Freshwater *is water with low concentrations of dissolved salts and other solids (< 1000 mg/L)* and a ***freshwater-related ecosystem*** *is a dynamic complex of plant, animal, and microorganism communities and the non-living environment dominated by the presence of flowing (lotic) or still (lentic) water, interacting as a functional unit."* (MEA (2005) adapted by Dickens and McCartney (2021)). These include rivers, estuaries, floodplains, wetlands, lakes, ponds, pools, pans, swamps and peat. Groundwater is sometimes also included in this definition but refers to water held underground in spaces between the soil and rock, so it is more often the interaction of ground and surface water ecosystems that is relevant. For the purposes of this review, we will consider freshwater ecosystems as including only lotic systems, in other words rivers, although the concepts can extend to other surface waters. Wetlands, lakes and groundwater are excluded as they are fundamentally different in structure and functioning and require a completely different approach to assessment. Any reference to "aquatic ecosystems" incorporates both freshwater and marine ecosystems.

"Aquatic ecosystem health" (condition, quality, integrity or status), has been defined as: *"the capability of supporting and maintaining a balanced, integrated, adaptive community of organisms having a species composition, diversity, and functional organization comparable to that of natural habitat of the region"* (Karr and Dudley, 1981). This definition clearly applies to rivers as well, is an abstract concept that cannot be measured from a single variable and all frameworks agree that a multiplicity of features are required to measure ecological health. However, which components to include is what varies the most between frameworks

One of the main differences between the definitions used by different frameworks is the inclusion of ecological structures (biota & abiotic features), functions (or processes) and/or non-ecological or human orientated components in the definition of "health". Although the components included by the various frameworks are reviewed in the following section, it is important to note how the definition of ecological health influences both which components are measured and how the final index of ecological health is calculated.

All frameworks consider the condition of ecological structures in the definition of ecosystem health, and most consider at least some aspect of the biota as core to the structure of ecosystems. Some (e.g., NARS, REMP, SRA, NRHP, MRC) consider ecosystem health to be best represented by the biotic components alone as the biota integrate changes across all aspects of the ecosystem. For example, the NARS consider "biological integrity" representative of "ecological health". Nevertheless, all these frameworks also consider abiotic (driver) components separately, as most human pressures to freshwater ecosystems are abiotic in nature so changes to abiotic aspects can provide useful information to enable the interpretation of impacts to the biota. Others (e.g., WFD, RHI, FBEHF, FHI) consider both biotic and abiotic features under a single definition of ecological health. The WFD and FBEHF also include ecological functions in the definition of ecological health. However, ecological functions or processes (e.g. biogeochemical cycles and biotic interactions) are notoriously difficult to measure so, even though included in the definition, they are not measured at all by the WFD and only partially by the FBEHF. Furthermore, some frameworks define "freshwater health" more broadly to include non-ecological factors. The FHI, for example considers "Ecosystem Vitality", "Ecosystems Services", and "Stakeholder Values & Governance" to be components of freshwater health (Vollmer *et al.*, 2018) (see Table 3.2).

Several frameworks include human values in the definitions of "health". This is based on the idea that a healthy waterbody can keep its ecological functions while maintaining the needs of society (Meyer, 1997). Many more-recent frameworks therefore include the human dimension in the definition and use a broader definition of "**freshwater health**" that encompasses both *ecological values (ecological health and resilience to stress) and human values (social services and benefits).* It is a key feature of the Chinese RHI, which is explicitly based on the need to assess the effectiveness of water governance as well as freshwater ecological health. The inclusion of human values is more common among global frameworks, being a key feature of the FHI, whilst a human focus is also evident in the PB, ITI, and SEEA, although many of the global frameworks do not clearly define "ecological health".

Table 3.2: Definitions of freshwater ecological health (or related terms) used by regional and global frameworks. We also include a breakdown of whether the definition includes biotic, abiotic, ecological process, and/or human (services, stakeholders & governance) aspects.

Framework	**Term**	**Definition**	**Aspects considered**
		Regional	
WFD	*Ecological status*	*An expression of the quality of the structure and functioning of aquatic ecosystems*	Biotic/ Abiotic/ Processes
NARS	*Biological integrity*	*Freshwater ecosystems' ability to support and maintain a balanced community of organisms comparable to those in natural condition*	Biotic
REMP	Ecological Status	*The totality of the features and characteristics of the river and its riparian areas that bear upon its ability to support an appropriate natural flora and fauna and its capacity to provide a variety of goods and services*	Biotic/ Abiotic

Framework	Term	Definition	Aspects considered
RHI	River (lake) Health	*A healthy waterbody can keep its ecological functions while maintaining the needs of the society, it should therefore contain both ecological values (ecological integrity and resilience to stress) and human values (social services and benefits).*	Biotic/ Abiotic/ Processes/ Human values
NRHP	River Health (or Condition)	*"The ability of the aquatic ecosystem to support and maintain key ecological processes and a community of organisms with a species composition, diversity and functional organisation as comparable as possible to that of natural habitats within a region"*	Biotic/ Processes
IECA	Ecological condition	*The state or health of individual animals or plants, communities or ecosystems as they relate to values and ecosystem services with reference to specific management goals or objectives and assessment against a defined baseline. Condition indicators can be physical-chemical or biological and represent the condition of the ecosystem. They may also be surrogates for pressures and stressors acting within the ecosystem.*	Biotic/ Abiotic/ Human values
SRA		*The ability of the aquatic ecosystem to support and maintain key ecological processes and a community of organisms with a species composition, diversity and functional organisation as comparable as possible to that of natural habitats within a region*	Biotic/ Abiotic
FBEHF	Ecological Integrity	*Ecological integrity refers to the ability of an ecosystem to support and maintain structure and function over time in the face of external stress.*	Biotic/ abiotic
		Global	
SDG 6.6.1	Health	*The ability of ecosystems to maintain their structure and function over time in the face of external pressures*	Biotic/ Abiotic
CDB	Biodiversity	*"Biological diversity" means the variability among living organisms from all sources including, inter alia, terrestrial, marine and other aquatic ecosystems and the ecological complexes of which they are part; this includes diversity within species, between species and of ecosystems*	Biotic/ Abiotic
FHI	Freshwater Health	*"…the ability of freshwater ecosystems to deliver ecosystem services and benefits, sustainably and equitably, through effective management and governance".*	Biotic/ Abiotic/ Human values
PB	Planetary Boundaries	"safe operating space for humanity"	Biotic/ Abiotic

Framework	Term	Definition	Aspects considered
ITI	Incident Threat Indices	Incident threat refers to the exposure to a diverse array of stressors at a given location (threats to river biodiversity and human water security)	Biotic/ Abiotic/ Human values
SEEA	Ecosystem Condition	“The state and functioning of an ecosystem in relation to both its naturalness and its potential to supply ecosystem services”	Biotic/ Abiotic/ Human values
EPI	Ecological Vitality & Environmental Health	Ecological Vitality:	Abiotic/ Human values

The decision of what to include or exclude in a definition of river health (RH) should be determined by the use for the indicator. In this report, where it is the health of the river alone, as an ecological entity, that is being considered, the following definition is adopted:

Recommended definition of River Health (adapted from Karr and Dudley, 1981):

"The ability of the river ecosystem to support and maintain key ecological processes and a community of organisms with a species composition, diversity and functional organisation as comparable as possible to that of natural habitats within a region".

Several of the frameworks reported have included human values in their definitions, which has the advantage of promoting the concept of RH into society. However, for this RH framework, it is proposed that it is the ecosystem alone that is of relevance, and that this indicator can then provide a contribution into other indicators that include the human perspective.

Definition of other commonly used terms:

Ecosystem components are the factors constituting an ecosystem. For example, these include hydrological, geomorphological, physico-chemical, and biological components in freshwater ecosystems. However, in some frameworks i.e., the Freshwater Health Index or Chinese River Chief System, the definition of “components” may be extended to include non-ecological factors, such as social services or governance. This is used in the EPA assessments (USEPA, 2006, 2020a) and EcoStatus Reports (Kleynhans and Louw, 2008) and is synonymous with “*elements*” in the EU WFD (CEC, 2000) and “*Themes*” in the Australian SRA (Davies *et al.*, 2010). It is equivalent to "*sub-categories*" in the Chinese framework (Xie *et al.*, 2020).

Measures / variables / metrics are individual factors that can be measured directly and usually contribute to the estimation of the status of components.

Environmental Stressors or simply “**stressors**” are anthropogenic factors causing detrimental ecological change.

3.2.3 Using conceptual models

The links between components are important to understand for a robust and informative framework. They are also very useful for the identification of key components and appropriate indicators by revealing key nodes within the network. Perhaps the pre-eminent conceptual model for ecosystems is the DPSIR framework (see Figure 3.2) that clearly separates the drivers of ecosystem change, from the actual changes that represent the "state" and the resulting societal impacts. Examples of frameworks using conceptual models of freshwater ecosystems include the REMP and SRA where there is clear separation of the drivers from the ecological responses. It all comes down to the purpose of the framework, and if it is RH then as discussed elsewhere, the state of the aquatic ecosystem itself should dominate with or without reference to the drivers of change. Consideration of the scale of the assessment as shown in Section 3.3 and the types of data available will also be important to draft the conceptual model.

3.2.4 Key components for assessment of river health

All frameworks agree that freshwater (ecological) health is made up of multiple components, and for a framework to be representative, it must include measurements of all these components. However, the components considered relevant differs between frameworks. We thus summarised the components considered by all the reviewed frameworks in Table 3.4 below to determine which are the most widely used and appropriate at the regional vs global scales.

We found the four most commonly included components are 1) **biology** (aquatic life), 2) **physico-chemistry** (water quality), and 3) **hydrology** (water quantity & dynamics), and 4) **morphology.** The *biological* component includes all measures of aquatic life and biodiversity, including invertebrates, fish, macrophytes, periphyton, zooplankton, phytoplankton, and other groups where relevant, as well as biodiversity overall and invasive species. In many models the relationships of taxa to adverse conditions are also included. The *physico-chemical* component relates to water quality, including physical aspects such as temperature and water clarity, and chemical aspects, such as pH, salinity and concentrations of various pollutants (salts, nutrients etc), and sometimes disease-causing pathogens. The *hydrological* components include aspects of water (flow) quantities and dynamics with a particular focus on seasonality. The *morphological* component includes aspects of physical form (channel, bank & bed conditions, geomorphological process (e.g., sediment dynamics) and riparian zone/ vegetation conditions. Of the components considered less often, the ***ecological processes*** component includes of the processes through which ecological structures interact, such as biogeochemical cycles and biotic interactions. However, although this features prominently in the definitions of "freshwater ecosystem health" of several frameworks (see e.g., WFD), it is notoriously difficult to quantify and lacks suitable indicators so only features in the more recent FBEHF. Indeed, in a review of how ecosystem health is defined and measured in 119 published studies, 80% used a combination of two or more physical (form and flow), chemical and biological indicators to assess ecosystem health (O'Brien *et al.*, 2016), whilst only 30 % of studies included indicators of ecosystem processes. Another aspect of freshwater ecosystems that only sometimes features as a distinct component is ***connectivity***, which is a measure of the connection of water as well as living organisms and substances (e.g. nutrients or sediment) and their ability to move freely within an area in all spatiotemporal dimensions. In addition, only a few frameworks considered the ***ecosystem (social) services*** component, which includes the provisioning (water, food, energy etc.), regulation and support (flood control etc.), and cultural/ aesthetic values (spiritual significance, recreation etc). ***Human health*** is only considered by NARS. Similarly, the ***governance*** component, which recognises the importance of governance in influencing ecological conditions and ecological services and considers the government's vision and ability to adapt, whether it fosters an enabling environment, and effectiveness, is featured infrequently.

Karr (1999) and Meyer (1997), and some of the approaches reported above (Table 3.4 i.e. IECA, MIF, RHI, NRHP, FHI, SEEA, ITI) propose that human values should be included in river health assessments (see Figure 3.1). This ranges from ecosystem services, to monetary value and social services. Advocates supporting that RH assessment should be based purely on ecosystem criteria include many authors such as (Haskell et al., 1992) and also the majority of those approaches in Table 3.4. Inclusion of social and economic values in a RH assessment introduces indicators for the purpose of describing a situation in its broader context, how an ecological system fits within a social system. This however is responding to a different mandate, that takes it beyond a pure estimation of ecosystem health. Thus, it is important to decide upfront on the purpose of the RH assessment, as either it is an assessment of the state of river ecosystems, or of ecosystems integrated with society.

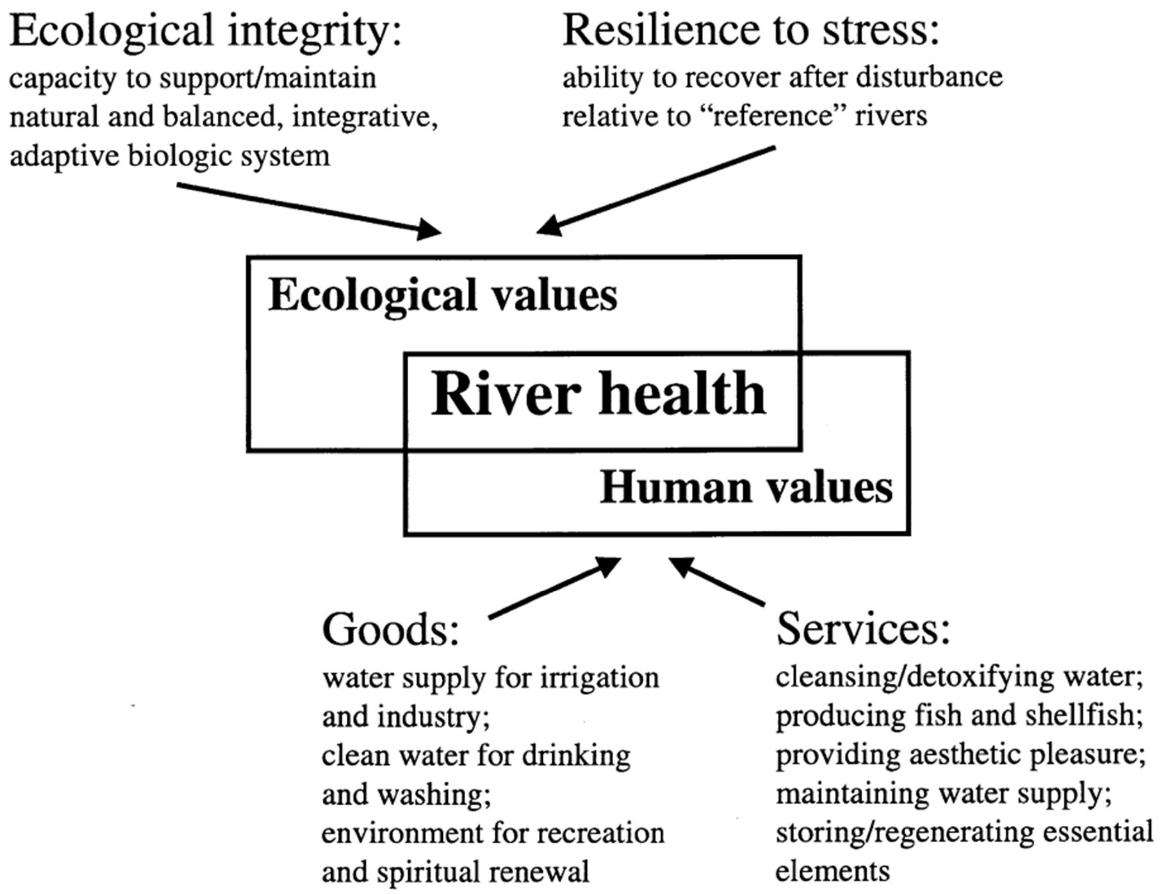


Figure 3.1 Proposed integration of ecological and human values in the definition of river health (Karr, 1996 and Meyer, 1997)

3.2.4.1 Grouping of components between frameworks

There are, however, important differences in the grouping of components between frameworks. This mostly involves the different treatments of physical form, riparian vegetation, sediment dynamics, habitat quality, and connectivity, which we have here considered as two components: 'morphology' and 'connectivity'. These are variably lumped, or split based on perceived differences or similarities in the mechanisms by which they influence the biota. Most regional frameworks include some measure of physical form (i.e., channel, bank, & bed conditions), which affects the habitat suitability for organisms. Sometimes these are grouped together with sediment dynamics (e.g., under 'Physical Form' in the SRA or the 'Geomorphological Element' of the REMP), as sedimentation or scouring can also negatively affect the availability of instream habitat, as well as having direct effects on the biota and water quality. The 'physical' component also often includes the riparian zone (e.g., NARS), although riparian vegetation is often also considered an independent component (e.g., REMP and SRA), since changes in the riparian vegetation can influence all other components including the instream biotic conditions. Both the REMP and SRA considered riparian vegetation to be a biological component, given that plants respond to changes in abiotic drivers (especially flow) and can indicate sustained long-term impacts, especially to the hydrological regime.

Given that most the above features affect the biota through habitat availability, they may be considered together as a single ‘habitat’ component (e.g., the ‘physical habitat’ components of the RHI and FBEHF). Given the importance of ‘habitat’, the REMP offers direct ‘habitat’ assessment as an alternative to the more detailed assessment of geomorphological, physicochemical and hydrological components. Connectivity, on the other hand differs fundamentally from the above-mentioned components in mechanisms by which it impacts ecosystem conditions (especially the biota) and is a major driver of losses in biodiversity and ecosystem function (Tickner *et al.*, 2020) so we preferred to consider it an independent component (e.g., IECA), although given that it is ‘structural’ in nature it is often lumped together with the morphological component.

In larger-scale frameworks, operating at the basin rather than reach scale, in particular the geomorphological, hydrology and connectivity aspects are often lumped into a single ‘basin conditions’ component. For example, in the WFD, the grouping extends even wider to include hydrology, geomorphological and connectivity together as a single ‘hydro-morphological component. In the FHI, physical form and land cover are lumped together as ‘drainage-basin conditions’ and in the ITI, landuse and disconnectivity of critical habitats are grouped as ‘catchment disturbance’, and connectivity, flow and social services as ‘water resource development’. In the end, however, the grouping is less important than the role the respective components play in the calculation of “freshwater health”, and this can be effectively dealt with using a weighting system (see Section 3.2.6.7: below), which relies heavily on having robust conceptual models.

The main reasons for differences in the choice of components are 1) the definition of what is meant by “freshwater health” and the conceptual framework, and 2) scale limitations. In terms of the role of the definition of “freshwater health”, all frameworks consider the biological conditions as most representative of “ecological health,” so their consideration as an independent component is clear. The difference is usually in the treatment of abiotic components. In frameworks where the biotic conditions are the focus, abiotic features are considered as secondary and only with regard to their importance as drivers of biological conditions. For example, the NARS and NRHP include loose sets of morphological and physico-chemical indicator variables to assist in interpretation of biological conditions. The lack of detailed conceptual models in both examples is certainly important in the looseness of the definition of these components. However, for frameworks with explicit conceptual ecosystem models, the abiotic components are usually more clearly defined. Most of these include some grouping of hydrology, physical form, and physico-chemical conditions. In addition, ecosystem services and governance are only included in frameworks that include human values in their definition of “freshwater health”. Therefore, we do not consider these useful for a global review of “freshwater **ecological** health”, but they may play a role in **river health**.

3.2.4.2 Influence of scale

In terms of influence of the scale of analysis on the choice in component, the above-mentioned components (*biological, physicochemical, hydrological, and morphological*) are generally the preferred options for inclusion in representative frameworks and are typical of regional frameworks in areas where this data is readily available. However, the lack of readily available data becomes a challenge at the global scale especially because many of these indicators require *in situ* measurement. This is particularly the case for the biological component, which is noted as a knowledge gap by SDG 6 and EPI and where biodiversity indicators are included (FHI, PB, ITI), they are poorly representative of biotic conditions (see Table 3.3).

Table 3.3 Approaches to data acquisition for by the reviewed freshwater ecological health assessment frameworks.

	Approach to data acquisition				
	In situ	**Earth Observation**	**Modelled**	**Government Statistics**	**Integrative**
			Regional		
WFD	X				
NARS	X				
REMP	X				
RHI	X	X	X	X	X
NRHP	X		X		
IECA	X	X	X		X
SRA	X	X			X
FBEHF	X				
MIF	X	X	X	X	X
			Global		
SDG 6	X	X	X	X	X
CBD	X	X	X		X
FHI	X	X	X	X	X
PB		X	X	X	X
ITI		X	X	X	X
SEEA	X	X	X		X
EPI				X	

Below, we summarise some of the main types of indicators and sources of data available for the key components identified above and review some potential large-scale options for application. This is only intended as a guide to the range of options available and is by no means exhaustive, so further investigation is required to propose indicators for a global RH framework.

Biological Indicators below). In cases where the inclusion of large-scale biological measures is attempted, the variables used are generally broad-scale biodiversity indices that are not representative of actual biotic conditions at smaller scales (e.g., modelled distributions of threatened and invasive species in the FHI).

Proxies have been used to substitute for the lack of *in situ* data especially the biological data. The two most common are 1) the ***spatial extent of freshwater ecosystems*** and 2) **Basin Conditions (or Landuse)**. *Spatial extent* is a measure of the overall extent of an ecosystem. Given that the spatial extent of an ecosystem encompasses all its components, it is used as an alternative in some global frameworks, most notably the SDG 6. However, as discussed in Section 3.2.5.8 (Spatial Extent Indicators), it is inadequate at indicating ecological health in freshwater ecosystems (besides wetlands), as most impacts do not result in a change in area. On the other hand, *basin conditions* may include changes in landuse within the catchment and are based on EO data. Given that human activities within the catchment ultimately affect the freshwater ecosystem condition, it is a possible proxy for “ecological health”, although the representativeness depends on the indicators included.

Table 3.4: Summary of components included in the reviewed freshwater (ecological) health assessment frameworks at regional and global levels with total frequencies in the final column. The 4-5 most frequently used at each scale are indicated in bold text. Sub-components of the ‘morphology’ component are left-indented and written in grey below it. Components recognised as knowledge gaps but in existing frameworks are also indicated as ‘GAP’s.

Freshwater Ecosystem Health Components	DPSIR	Regional									
		WFD	NARS	REMP	RHI	NRHP	IECA	SRA	FBEHF	MIF	
Biology (Aquatic Life)	**S**	**X**	**X**	**X**	**X**	**X**	**X**	**X**	**X**	**X**	**9**
Physico-chemistry (Water Quality)	**D/S**	**X**	**X**	**X**	**X**	**X**	**X**		**X**	**X**	**8**
Hydrology (Water quantity & dynamics)	**D/S**	**X**		**X**	**X**	**X**	**X**	**X**	**X**	**X**	**8**
Connectivity	D/S	X			X		X				3
Morphology (Physical Habitat)	**D/S**	**X**	**X**	**X**	**X**	**X**	**X**	**X**	**X**	**X**	**9**
Physical Form	*D/S*	*X*	*X*	*X*	*X*		*X*	*X*			*6*
Riparian Vegetation	*D/S*	*X*	*X*	*X*	*X*	*X*		*X*			*6*
Sediment dynamics	*D/S*		*X*	*X*	*X*			*X*		*X*	*5*
Habitat	*D*		*X*	*X*	*X*	*X*			*X*	*X*	*6*
Ecological Processes	S				X				X		2
Spatial Extent	S						X			X	2
Basin Conditions (Land use)	D									X	1
Human Health	I		X				X				2
Ecosystem (Social) Services	I				X		X			X	2
Stakeholders Values	I				X						1
Water Governance	R				X						1

	Global								
	SDG6	CBD	FHI	PB	ITI	SEEA	EPI	ERP	
Biology (Aquatic Life)	**GAP**		**X**	**X**	**X**		**GAP**	**X**	**6**
Physico-chemistry (Water Quality)	**X**		**X**	**X**	**X**	**X**	**X**	**X**	**7**
Hydrology (Water quantity & dynamics)	**X**		**X**	**X**	**X**	**X**		**X**	**6**
Connectivity			X		X			X	3
Morphology (Physical Habitat)			X			X			2
Physical Form			*X*						*1*
Riparian Vegetation						*X*			*1*
Sediment dynamics					*X*				*1*
Habitat						*X*		*X*	*1*
Ecological Processes									**0**
Spatial Extent	**X**					**X**	**GAP**	**X**	**4**
Basin Conditions (Landuse)			**X**	**X**	**X**			**X**	**4**
Climate Change				X					1
Human Health									0
Ecosystem (Social) Services			X		X	X			3
Monetary Value						X			1
Stakeholders Values			X						1
Water Governance			X						1

3.2.5 River Health indicator types

Indicators are measures of features that provide clues to matters of larger significance or enable the perception of trends or phenomena that may not be immediately detectable. Thus, their significance extends beyond what is actually measured to larger phenomena of interest (Hammond, 1995). Therefore, ecological indicators provide insights into ecosystem conditions. For example, measuring the concentrations of N or P in a waterbody not only provides information on their current levels but, when elevated, can signal a change in trophic state or eutrophication, which has a cascade of effects on the state of biological communities and water quality and even human health. Thus, the choice of indicators used by an assessment framework is crucial to its success. They are central to ensuring the robustness of the information obtained, whilst flexibility in the choice of indicators is key to enabling the flexibility and scalability of the frameworks itself i.e., its application under different contexts and at variable scales. There are thus several conditions to consider when choosing indicators.

Of first consideration are the ***types of indicators*** to use. The OECD (1993) defined three categories of indicators, namely measures of environmental pressures or drivers, state and (societal) responses (i.e. DSR or PSR). This is further refined by Smeets and Weterings (1999) into five indicator classes, known as the DPSIR framework, used by the European Environment Agency. This distinguishes between indicators of *driving forces* (e.g. indirect social and economic developments that result in pressures on the environment), *pressures* (forces acting on the environment e.g. pollution discharges, land-use changes etc.), *state* (changes in the environment e.g. the conditions of biological communities, water quality, habitat etc.), *impacts* (of environmental state on other features e.g. human health, ecological systems or materials), and *responses* (societal reactions to changes in the environmental state e.g. policy change) (see Figure 3.2). Most of the reviewed ecosystem health frameworks include a combination of state and pressure indicators (see Table 3.5), where state indicators form the core of 'freshwater ecosystem health' assessment, as this is fundamentally a measure of 'state'. Most frameworks also include some aspect of pressure indicators, as these enable interpretation of changes in the ecological state. Only the ITI focuses entirely on the drivers and pressures, as their objective was to determine the threat to water resources (not measure the state). Some go further, to show the link between changes in the state and further impacts, such as to human health and recreation in the NARS or ecosystem services in the RHI, IECA, and FHI, whilst from the SEEA further quantifies the monetary impacts of the losses of ecosystem services due to a reduction in ecological state. However, the RHI and FHI are the only examples that include response indicators in the form of governance and stakeholder values indicators, in an attempt to show the role that government and societal attitudes play in influencing the ecological state. The choice of indicator class(es) used by a framework is, therefore, heavily influenced of its definition of "freshwater health", which itself depends on the management objectives of the assessment so the sound consideration of these aspects during framework development are fundamental to influencing the choice of indicators.

Another important distinction to make is between ***direct and proxy indicators***. Direct indicators provide measures of a component in question, such as temperature or pH as parameters of water quality. However, in cases where there are no suitable indicators of key phenomena, measures of closely related variables that are causally linked to the variable of interest or "proxy indicators" provide a viable alternative. In this way, N & P are proxy indicators of eutrophication as they are the main cause of eutrophication and are easier to measure than direct measures of algal growth. At a higher level of analysis, the EPI, for example, uses 'wastewater treatment' (percentage of the total wastewater produced that is treated) as a proxy measure of water quality. This was due to the lack of suitable globally applicable direct indicators of water quality and the fact that untreated wastewater is one of the main drivers of poor water quality worldwide for which data are readily available, making it a good indicator of general water quality at the international scale. However, the use of proxy indicators has various shortfalls, namely in that they overlook other aspects affecting the ecological state being measured. For example, in the EPI example, focusing on wastewater overlooks all other sources of poor water quality. Therefore, the closer the causal relationship of the proxy to the variable of interest, the better. The identification of the most relevant indicators (direct or proxy) is best established using causal networks that show the relationships between all relevant indicators (Niemeijer and de Groot, 2008).

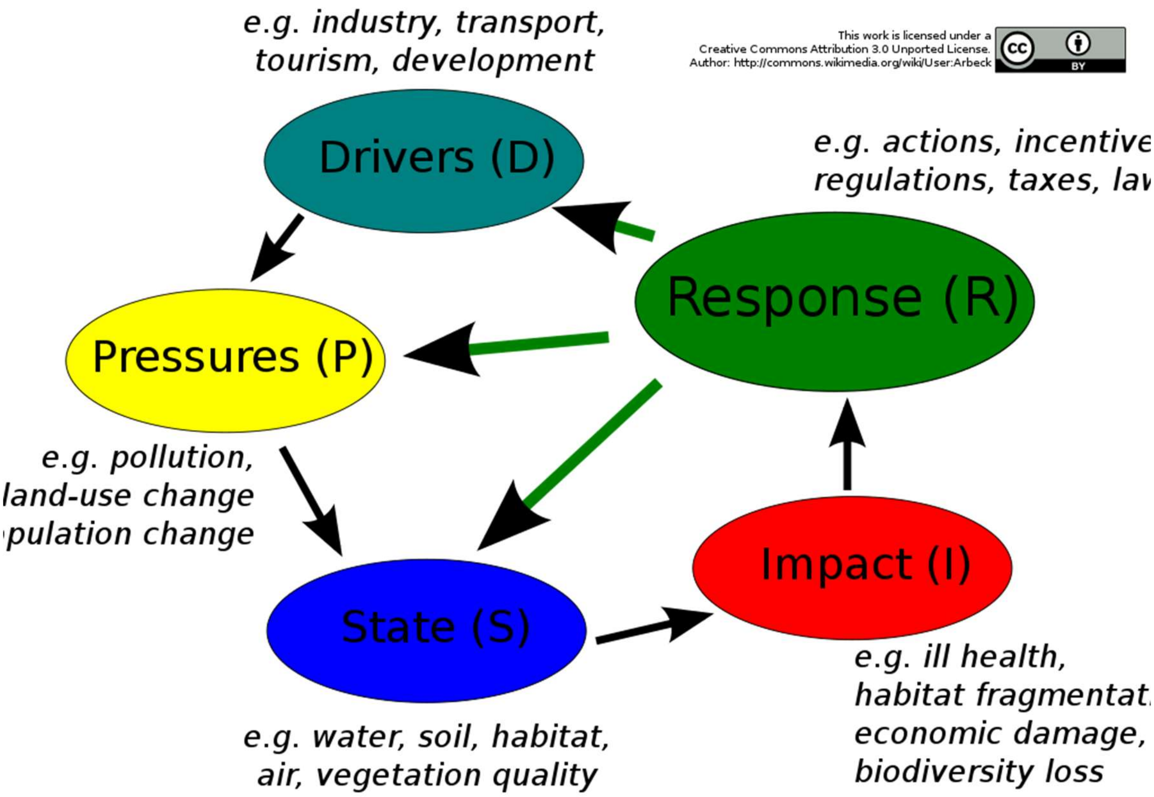


*Figure 3.2: Model of the DPSIR framework of environmental indicators, consisting of **d**riving forces, **p**ressures, **s**tates, **i**mpacts, and **r**esponses, as used by the European Environmental Agency* (Smeets and Weterings, 1999).

Table 3.5: The use of different indicator classes by the reviewed regional vs global frameworks.

	Drivers	**Pressure**	**State**	**Impact**	**Response**
			Regional		
WFD		Direct	Direct		
NARS		Direct	Direct	Direct	
REMP		Direct	Direct		
RHI		Direct	Direct	Direct	Direct
NRHP		Direct	Direct		
IECA		Direct	Direct		
SRA		Direct	Direct		
FBEHF		Direct	Direct		
MIF		Direct	Direct	Direct	
			Global		
SDG 6			Proxy		
CDB					
FHI			Direct	Direct	Direct
PB			Direct		
ITI	Direct	Direct			
SEEA		Direct	Direct	Direct	
EPI			Proxy		

3.2.5.1 *Criteria for good indicators*

A widely recognised standard for the selection of indicators is the SMART framework. The origins of the SMART indicator framework are unknown, but it has developed over the past six decades and is foundational in management circles.

- **Specific** - the indicator needs to be narrow and accurately describe what needs to be measured.
- **Measurable** - regardless of who uses the indicator it would be measured in the same way.
- **Achievable (or attainable)** - means that collecting the data should be straightforward and cost-effective.
- **Relevant** - closely linked to the relevant outcome.
- **Time-bound** - there should be a timeframe linked to the indicator (such as the frequency with which it is collected or measured).

However, in relation to RH reporting there are further considerations to consider that provide more support to RH programme implementation:

Sensitivity (Specific): The sensitivity of potential indicators is of course paramount as in order to be effective they must be sensitive to anthropogenic impacts. They should also respond to anthropogenic impacts in a predictable manner and be anticipatory of the ecological responses (such that they can signify impending changes in ecological systems). The indicator should also be closely related to the component of interest, whilst the full suite of indicators chosen for a component should be fully integrative (i.e., cover all key ecological gradients e.g., soil, vegetation, temperature changes etc.). One should also aim to use indicators with the least variability in response to natural changes to maximise the signal to noise ratio of human impacts to natural variation given that their sensitivity to human impacts makes many indicators highly sensitive to natural disturbances. For example, macroinvertebrate communities may be radically altered by flooding. This can be aided by using rules, to avoid sampling macroinvertebrates in the period during or after a flood event in the example above.

Scale (Relevant): Considerations of spatial and temporal scales are also important. Most indicators are more sensitive to impacts at certain scales and not others so the most suitable indicator for a particular assessment may vary according to the scale employed. For example, an *in situ* based invertebrate example may be appropriate for monitoring below a sewage outfall, while an Earth Observation technique would be better suited for land-use impacts. Therefore, one should aim to measure indicators at the scale that provides the most useful information for the management objectives. In other words, the spatial and temporal scales at which the impacts and ecological changes occur and at which management interventions can be made. Generally, this involves using indicators measured at the scale of the assessment or below, where there is the possibility of upscaling. A dashboard approach, where different indicators are recommended according to scale, would therefore be most beneficial.

Feasibility (Achievable): Perhaps the most important considerations at the global scale are around feasibility. Indicators with standardised methods, especially those already in use are generally preferable to those without a standard protocol or not yet applied, as they are generally more accessible, reduce the need for further investment in research and development, and have a proven record of accomplishment. This includes whether reference conditions are defined, as this is key to ensuring consistency in the application of an indicator. The methods of assessment and analysis should also be easy to use and understand as well as inexpensive, as this is crucial to the feasibility and cost-effectiveness of the resulting framework. However, at the global scale, many of these questions of feasibility, come down to the means of data acquisition.

3.2.5.2 Approaches for RH assessment

There are several different approaches for RH assessment, the selection of which is dependent on the objectives for monitoring.

Bottom-up (in situ) approach: The bottom-up approaches involve taking measurements *in situ*. Historically, this has been the approach used in past frameworks. It is still the most representative source of data at the local scale (i.e., individual river reaches). Hence, it is still the preferred option for most regional frameworks, especially those operating down to the local scale. However, it is very resource intensive, requiring large amounts of time, labour and money, and usually also requires the establishment of extensive sampling networks (generally also including reference sites where *ecological ratios* are used). This makes it inappropriate for many parts of the world, particularly poorer countries and remote regions (i.e., regions that presently lack monitoring systems and for which a global assessment is most desired). However, the development of new technology (such as eDNA) and automation of monitoring systems (e.g., for hydrological and physicochemical indicators) is reducing the cost of obtaining *in situ* measurements. These are still more likely to be carried out in wealthier countries and are not nearly widely-enough implemented for global implementation. Another major challenge with the bottom-up approach is how to scale up the results. In other words, how to determine the conditions of areas not directly measured. Usually, this involves inference of the conditions at the nearest (upstream) site, providing they remain constant (i.e., are not subject to major changes in biophysical features or additional human impacts). The problem lies with inferring the ecological condition of sites where conditions may differ from those at the monitoring site(s). For this reason, most frameworks using the bottom-up approach require the establishment of extensive monitoring networks to cover all geographically distinct regions of a river/ lake basin. Furthermore, the selection of monitoring sites not fully representative of the basin can also lead to biased results. Therefore, although the bottom-up approach may provide the most representative locally, there are still has large challenges in the representativeness of *in-situ* data when selecting sampling sites and up-scaling results, and EO or modelled approaches may be just as accurate, if not more so for some indicators. Nevertheless, the lack of appropriate top-down or modelled alternatives for several variables, most notably the biological component, mean it will likely continue to play a role in a global RH assessment.

Top-down (Earth Observation) approaches: The top-down approach involves taking measurements most commonly from satellite imagery and detection, also known as 'earth observation'. This data is readily available data at the global or continental scale and at a much lower cost with less bias than *in situ* monitoring systems. It is thus an obvious source of data for a global assessment. It is applicable to various components, particularly to measure changes in landuse, vegetation and basin condition. Progress is also been made for its use to measure physicochemical indicators (e.g. AquaSat (Ross *et al.*, 2019) and (Ritchie, Zimba and Everitt, 2003), including Total Suspended Sediment (TSS), turbidity, clarity (Secchi Disk Depth (SDD)), Dissolved Organic Carbon (DOC), Chlorophyll a, and temperature. However, it is not suitable for several important indicators and components, particularly the biological component, for which monitoring of Chlorophyll a as a measure of productivity is the only readily measurable indicator, and major knowledge gaps are still present. Methods are also presently in their infancy for measurement of hydrological indicators. Another drawback with using EO is the that it is dependent on limited time series due to set orbits of satellites or because of data processing limitations that take only infrequent slices of the available data, which means that impacts are often missed or underestimated in severity, especially when they are short lived or linked to cycles not corresponding to the satellite time series. Unfortunately, rapid change is a feature of lotic systems, suggesting that the use of such methods may not be adequate at capturing impacts (Hsu *et al.*, 2016). Nevertheless, improvements in satellite data are steadily advancing and proving useful for monitoring global trends (Hsu *et al.*, 2016).

Modelling Approaches: The modelling approach involves taking measurements of indicators from models developed for the relevant systems. Models are informative representations of systems and can take various forms. They may be empirical (i.e., where the relationships are entirely determined from actual measurements), partially empirical (more conceptually designed but confirmed by actual measurements), or machine learning-based (relationships are automatically adjusted based on continuously updated data). Modelled data reduces the reliance on physical measurements (*in situ* or EO), with the associated benefits of being far cheaper than *in situ* systems and less susceptible to site selection bias, whilst being applicable at large scales with suitable resolution across scales (depending on the model resolution). The greatest concern is potential shortfalls in the accuracy of the data. This depends on the robustness of the model, which can vary dramatically, so the choice and design of models used is crucial to obtaining representative data. Appropriate large-scale or global models for inclusion in a global assessment are already in place for the hydrological component, and can also be used for indicators of sediment dynamics e.g. the SedNet model (Wilkinson et al. 2004) in the SRA (Davies *et al.*, 2010), and for connectivity (the CSI (Grill *et al.*, 2019)). Modelling of species distributions also provides a possible avenue for the quantification of biological conditions reducing the need for *in situ* monitoring. For example, the Australian SIGNAL (Chessman, 1995, 2003)) uses species distribution models to predict the species that should naturally be present (against which observed communities can be compared). These can also be used to determine comparative biodiversity indices, although these have not yet been developed at the global scale.

Integrative Approaches: An integrative approach utilises a combination of *in situ*, EO and/ or modelled data sources. Its greatest benefit is that it makes use of the approach best suited to each indicator, maximising the total benefits. For example, hypothetically, one could combine measurements of basin conditions (landuse) based on EO, biological indices based on *in situ* data, and hydrological conditions based on modelled data. It is thus flexible to the variability in data availability, institutional capacity and natural conditions between regions, which is an important feature of successful frameworks, and can provide a more robust and representative measure of ecological conditions than frameworks based solely on *in situ*, EO, or modelled data. The integrative approach has become increasingly utilised in recent frameworks both at regional and global scales (i.e., IECA and RHI, and FHI and SEEA, respectively).

For selecting globally indicators relevant at large scales, therefore, one will generally end up making trade-offs between the sensitivity of a potential indicator (i.e., its robustness) and the feasibility of gathering the relevant data at large-scales. This is because the most accurate indicators are often measured at the site level but appropriate survey methods and institutional capacity for implementation are lacking for large areas of the world making their widespread application not feasible. On the other hand, indicators that are easily obtainable at large scales (e.g., from EO or modelling) may be inappropriate for indicating local conditions.

Table 3.6: Approaches to data acquisition for by the reviewed freshwater ecological health assessment frameworks.

	Approach to data acquisition

	In situ	Earth Observation	Modelled	Government Statistics	Integrative
	Regional				
WFD	X				
NARS	X				
REMP	X				
RHI	X	X	X	X	X
NRHP	X		X		
IECA	X	X	X		X
SRA	X	X			X
FBEHF	X				
MIF	X	X	X	X	X
	Global				
SDG 6	X	X	X	X	X
CBD	X	X	X		X
FHI	X	X	X	X	X
PB		X	X	X	X
ITI		X	X	X	X
SEEA	X	X	X		X
EPI				X	

Below, we summarise some of the main types of indicators and sources of data available for the key components identified above and review some potential large-scale options for application. This is only intended as a guide to the range of options available and is by no means exhaustive, so further investigation is required to propose indicators for a global RH framework.

3.2.5.3 Biological Indicators of RH

Biodiversity provides ecosystems with resilience to change, as the genetic material organisms contain provides the "information bank" that determines the potential for all life to coevolve with the abiotic environment. The diversity of organisms (and genetics) thus provides the long-term capacity for ecosystems and the biosphere to persist under scenarios of change (Steffen et al., 2015). Biological indicators measure ecological health based on the biodiversity present at sites. The biota integrate the influence of various drivers of change (water quality, flow, physical form, habitat etc.) and therefore provide a single convenient component for assay of ecological conditions that is representative of the whole ecosystem. They are thus widely used in official RH assessments carried out at the regional and global levels in the European Union, Japan, Republic of Korea, South Africa, and the USA and at the state/province level or in major catchments in Australia, Canada, China, New Zealand, and Singapore (Feio *et al.*, 2021). For some of these (i.e. NARS, REMP), they constitute the sole basis by which ecological conditions are characterised, with driver variables only being used for interpretation. However, there are important considerations involved in the choice of biotic indices to be included in a framework, including which group(s) of organisms are most suitable for use as indicators, given innate differences between regions/ habitats, and how they can be used to quantify ecological conditions.

Taxonomic groups used as biological indicators

Various groups are used as indicators of ecological state. The most widely used are ***benthic macroinvertebrates***, which are ideally suited to indicating ecological conditions, especially in smaller lotic systems, because they are a dominant component of these ecosystems, have diverse life histories and sensitivities to stress, and can respond rapidly to change. They are also relatively easy to sample and identify to an appropriate level. However, they are less effective indicators in larger rivers and lakes, given that they are a less prolific part of these ecosystems (compared to say plankton or fish) and it can be difficult to sample them in deep waters. ***Fish*** are another prominent group that are used by many frameworks. They are also sensitive to a large array of human impacts including water quality deterioration, changes in flow, physical habitat degradation, loss of connectivity, invasive species and over-exploitation. Therefore, they are well suited as biological indicators and are complementary to macroinvertebrates since they are more robust in larger rivers and lakes and display different sensitivities e.g., to fragmentation). The challenge in using fish is sample collection that can be expensive in terms of time and labour. ***Plants*** are also used as indicators. By far the most widely used group are the ***periphyton*** (algae communities growing on submerged surfaces), with periphyton-based indices, specifically for diatoms (algae with silica cell walls) growing quickly in popularity around the world. As primary producers, they are sensitive to changes in physical, chemical and biological factors, sometimes to a very specific degree, making them ideal indicators for specific impacts (e.g. nutrients, specific pollutants, pH, temperature etc.), sometimes indictable in other groups (e.g. herbicides), whilst their rapid life cycles enable them to indicate short-term or sudden changes (Li, Zheng and Liu, 2010). They are also very easy to sample, as one only requires a small sample of substrate, with simple to advanced methods available for analysis. The next most used group are the ***riparian macrophytes*** or **riparian vegetation**, which are mostly affected by changes in the flow regime and to the physical form of the river, thus are better at indicating long-term geomorphological and hydrological changes. The condition of the riparian zone also has an important influence on the in-stream conditions. On the other hand, ***aquatic macrophytes*** are generally only monitored in large lotic systems, especially where floating species (e.g., water hyacinth or *Azolla*) can become problematic, especially when nutrient levels are elevated. In lentic systems, especially lakes, ***zooplankton*** and ***phytoplankton*** are also usually used as indicators, as they form the base of the food chain in these environments and are susceptible to multiple impacts. Only occasionally are other groups, such as ***waterbirds***, ***dragonflies***, or ***amphibians and reptiles*** used, usually when their conservation is a target of water resource management. In many cases, especially at larger scales one may also find generalised ***biodiversity indicators***, which are inclusive of a variety of organismal groups.

Approaches to Biotic Indices

Changes to the biota because of anthropogenic impacts can be measured in various ways, including:

Diversity Metrics**:** Diversity metrics measure how many different types of taxa (families, genera or species) are present in a sample/ community and can take into account phylogenetic affinities and the abundance of the taxa present. They are based on the premise that environmental stress results in the loss of sensitive species, resulting in a reduction in the number of taxa present and the number of taxa present in high numbers. The simplest diversity metric is ***taxonomic richness***, which is simply a count of the number of taxa (families, genera or species) of a certain taxonomic group present and is normally associated with better ecological conditions. One of the most widely used taxonomic richness indicators in freshwater ecosystems is the number of Ephemeroptera, Plecoptera and Trichoptera number of families (EPT), e.g. (Baker and Sharp, 1998; Álvarez-Cabria *et al.*, 2017). These three orders of aquatic insects are common in benthic macroinvertebrate communities of lotic ecosystems and are made up of taxa with widely variable life histories and sensitivities to stress, so their richness is directly related to ecological health. However, given that environmental stress also results in a reduction in the abundance of sensitive species, not just the diversity present, ***taxonomic diversity***, takes into account both richness and abundance with values such as Shannon or Simpson diversity and is sometimes termed ‘taxonomic evenness’. In the Mekong River Commission biomonitoring surveys, biotic indices included such indices as the

Shannon-Wiener Diversity Index and the Berger-Parker Dominance Index (Vongsombath *et al.*, 2009; MRC, 2019b). The assumption is that natural ecosystems possess a greater number of abundant organisms (i.e., higher evenness), whilst stressed ecosystems are dominated by high abundance of only a few tolerant taxa (i.e., low evenness). However, such diversity indices are highly dependent on factors such as a sufficiently high level of natural taxonomic richness (species pool), low taxonomic turnover (i.e., natural absence), and the availability of suitable habitats, so these indices are less effective in areas with naturally low taxonomic richness, high turnover or homogenous habitat.

Phylogenetic species variability (PSV): Phylogenetics is the inference of evolutionary history and relationships between organisms, either using genetics of morphology. This approach could provide the most accurate indicator and may be applicable at local or regional levels. However, shortfalls in the global phylogenetic knowledge and of many important groups makes it unsuitable at present as an indicator. However, as genetic technologies are improving at an exponential rate, this may soon change. In the meantime, however, the **extinction rate** (of well-studied organisms), E/MSY or extinction per million species per year, is used as the next best alternative as it is better known. What makes it is less ideal is the loss of accuracy, due its sole focus on losses at the species level and the fact that it is subject to a time lag (between impact and extinction) making it difficult to use as a management tool (Steffen *et al.*, 2015).

Composition Metrics & Observed-Expected Ratios: Composition metrics are commonly used in the reviewed frameworks. They consider changes in the balance between different groups of organisms, either taxonomic or ecological, to reflect changes in environmental conditions. For all metrics involving composition or proportions, the most common means of quantifying change is as a ratio of the observed versus expected groups. Hereby, the observed values are compared to those that would be expected under natural conditions and expressed on a standard scale (e.g., 0 – 1 or 0 – 100). The greatest challenge is in the definition of the reference conditions to determine the expected proportions of each group, although generally these can be obtained through measurements at nearby reference sites, historical data, expert opinion or species distribution models.

The first type of composition metric is ***taxonomic composition***, which measures changes in the proportional abundance of different taxonomic groups within a community/ sample. Given that different taxa respond differently to stress, the proportions of taxa indicative of natural versus stressed conditions serve as good indicators of ecological condition. Taxonomic composition metrics are included in several frameworks. In the initial US Wadeable Streams Assessment (USEPA, 2006), for instance, the Macroinvertebrate O/E Ratios of taxa loss provides the central measure of ecological health.

The second type of composition metrics are ecological in nature i.e., comparing groups of organisms with different ecological similarities or life history traits that can reveal different types of ecological impact. One of the most common in lotic RH assessments, for example are ***functional feeding group (FFG) metrics,*** which measure the relative abundance of organisms grouped by feeding strategy (e.g., filtering, scraping, grazing or predation for macroinvertebrates). FFGs are closely tied to ecological conditions, particularly nutrient loads and factors affecting physical habitat (e.g., sedimentation), so the ratio of FFGs will change accordingly. A quantitative comparison of changes (e.g., Observed to Expected Ratio) to the proportions of FFGs thus provide a measure of impact. Other ***habits/ habitat* metrics** compare changes in the relative abundance of organisms according to other aspects of their life histories (i.e., how/ where they live, move and breed). Ecosystems with diverse habitats will support greater diversity while where habitat is lost (e.g., through siltation) the diversity of organisms with different habits will decrease. In the US NARS, for example, different habitat, feeding groups and spawning habitats are taken into account for fish, with feeding groups and habitats/ habits for macroinvertebrates. One group that have proved especially useful at providing information on ecosystem health are migratory species as their ability to move freely is highly dependent on ecological conditions. This includes the connectivity of a site to the wider system, which can be physically blocked by dams etc. or by poor ecological conditions unfavourable to the passage of the species involved, as well as the availability of suitable habitat and ecological condition of the site in question. Sites that are themselves disturbed or occur within a disturbed system would be expected to have lower proportions of migratory species than natural sites. The ***percentage migratory taxa*** is thus related to ecological condition. By contrast, the **p*ercentage native species*** or 'nativeness' of the community is a direct representation of ecological health. Alien invasive species (AIS) are generally associated with poorer ecological conditions as they are both generally more tolerant of environmental stress and are themselves the cause of stress to native species through competition and/or predation, which can lead to bottom-up changes in macro-ecological structures and processes. The percentage alien versus native organisms thus provides a direct measure of ecological conditions and is, again, most widely for fish, as either the percentage native species, as used in the NARS, SRA and EcoStatus Reports, or percentage native abundance or biomass in the SRA. Migration is particularly relevant to fish, which are often migratory by nature and whose habits are often better understood than invertebrates, and are included in the US NARS and South African FRAI (Kleynhans, 2008). Fish migration is also employed with telemetry, using fish to signal a variety of river conditions based on attached instruments (Burnett et al, 2021).

Functional diversity indices capture the role of organisms in ecosystem (or biosphere) functioning by measuring the loss if biodiversity components as changes in the value, range, distribution, and relative abundance of the functional traits of the organisms present in an ecosystem or biota (Steffen *et al.*, 2015). Several functional diversity indices for application at the local level are presented in Mason et al 2013 (in Steffen et al., 2015). However, application of these at larger scales is challenging and has yet to be undertaken.

Tolerance Metrics: Tolerance metrics are perhaps the most widely used type of biotic indices in the reviewed frameworks. They utilise tolerance (or sensitivity) scores of the taxa recorded in a sample to calculate indicator variables that are directly related to ecological conditions. Such scoring systems were initially created for macroinvertebrates in lotic ecosystems, e.g., the Biotic Index for South African Rivers (Chutter, 1972) and British Biological Monitoring Working Party Score System (BMWP, 1978; Walley and Hawkes, 1996). The robustness and ease of use of these systems led to the creation of multiple similar systems in other regions, such as the Spanish Biological Monitoring Water Quality (BMWQ) (Camargo, 1993), Wisconsin Biotic Index (BI) and Family Biotic Index (FBI) (Hilsenhoff, 1987; Hilsenhoff *et al.*, 1988), Australian Stream Invertebrate Grade Number Average Level (SIGNAL) scoring system (Chessman, 1995, 2003), New Zealand's Macroinvertebrate Community Index (MCI) (Stark, 1985, 1998), and South African Scoring System (SASS) (Dickens and Graham, 2002). The later was since adapted to various other African countries as the Namibian Scoring System (NARS) (Palmer and Taylor, 2004), Tanzanian River Scoring System (TARISS) (Kaaya, Day and Dallas, 2015), Zambian Invertebrate Scoring System (ZISS) (Dallas *et al.*, 2018), and Rwenzori Score (RI) (Musonge *et al.*, 2020). Also common are Periphyton (Diatom)-based Specific Pollution sensitivity Indices, SPI (e.g., Karthick et al, 2010) that claim to provide high-resolution indication of sources and types of pollution. Less common are tolerance indicators created for fish, for instance the Fish Response Assessment Index in South Africa (Kleynhans, 2008).

Multimetric Indices: Many of the reviewed frameworks aggregate several different metrics into a single multimetric index, which is thus more robust and integrates multiple sources of information. The 2013-14 US NARS survey (USEPA, 2020a) makes use of such indices to assess ecological conditions of macroinvertebrates and fish by ecoregion. For benthic macroinvertebrates, this included the development of a Macroinvertebrate Multimetric Index (MMI) that includes measures of taxonomic richness, taxonomic composition, taxonomic diversity, feeding groups, habits/habitats, and pollution tolerance, that varies between ecoregions. For fish, it included taxonomic richness, taxonomic composition, pollution tolerance, habitat and feeding groups, spawning habits, number and percentages of migratory taxa, and percentage native taxa. The exact metrics vary between the nine US freshwater ecoregions. For the Australian SRA (Davies *et al.*, 2010), for fish, the two indicators include (1) Expectedness and (2) Nativeness. Both are compositional metrics. Expectedness provides information on species richness relative to the reference condition based on the metrics a) Observed to Expected Ratio and b) Observed to Predicted Ratio. Both compare the number of native species predicted to occur under reference conditions and those collected, although the first corrects the number of species 'expected' to occur downward by taking into account rare species. Nativeness provides Information on the proportions of native versus alien species in the a) biomass, b) abundance, and c) species richness. For Macro-Invertebrates, samples are collected by standardised kick-sampling and weep-netting using the AUSRIVAS protocol (Davies 2000) where condition is determined using two indicators. The first metric considered is the O/E ratio comparing the 'observed' to 'expected' families. The reference condition is developed through application of filters, based on traits determining family distributional limits for temperature, hydrology, geomorphology and biogeography. The second is a SINGAL O/E Metric, which also compares the 'observed' vs. 'expected' taxa' but after the application of a tolerance scores using SINGAL (Stream Invertebrate Grade Number Average Level), which reflects the sensitivity of macroinvertebrates to disturbances (0 = high tolerance, 10 = high sensitivity). Using expert-rules, the two indicators were combined to provide a single metric of macroinvertebrate conditions. Other multimetric indices for other regions around the world include the Multimetric Index of Zio River Basin (MMIZB) in Togo (Tampo *et al.*, 2020).

***eDNA*:** This is a recent and rapidly developing field, advancing as the ability to monitor DNA in complex mixtures such as a water column is increased. Carraro et al, 2020 gave guidance on the use of eDNA for monitoring of an entire river catchment that is most appropriate for the development of a global RH programme, describing how the eDNA sampled at a river's cross-section results from the aggregation of the dynamics of particle transport from a number of upstream sources (i.e., the locations of the target species) along the river network toward the sampling site. The DNA analysis at that sampling site shows all of the species upstream in a somewhat indiscriminate way and presently lacks the ability to indicate the abundance or status of any of these species, making it difficult to interpret RH beyond an indication of total biodiversity.

3.2.5.4 Water Quality Indices

Measures of "water quality" are the original indicators of ecological conditions and are crucial for understanding and measuring ecological conditions yet a consistent global assessment somehow still evades implementation. Interest in water quality was first raised over concerns of the impacts of pollution on human health (CCME, 2002). However, this soon translated into an awareness of the impacts of pollution on freshwater organisms, paving the way for later indices of biotic or ecological conditions. Indeed, water quality is closely linked to ecological health, namely due to the toxic effects of pollution on organisms, including humans. Most countries around the world, therefore, have programmes in place to monitor water quality parameters of water resources, providing one of the largest sources of data available for a global RH assessment. However, expanding such indicators to larger scales is a major challenge primarily because water quality definitions vary widely depending on sources, location, and the intended use of the water, such that no single definition is appropriate for global measurement. It also involves many different chemical, physical, and biological parameters, whose relevance and/or status depends on the context, posing a massive challenge to create indices applicable at the international level**. Major ecological impairment or human health risk issues include oxygen depletion, nutrient pollution, acidification, salinization, faecal contamination, sediment loading, turbidity/ clarity and toxicity**. Nutrient pollution causes eutrophication or excessive growth in algae, affecting aquatic life and consist of nitrogen or phosphorous compounds. These are mostly commonly measured as the concentrations of total nitrogen, nitrates, nitrites, and ammonia, and total phosphorous or orthophosphates. Excessive organic material can also cause excessive bacterial growth and, along with the eutrophication, can cause a reduction in dissolved oxygen (DO), seriously affecting the survival of aquatic organisms. This can be measured directly or as dissolved organic carbon (DOC) or total organic carbon (TOC). Temperature, pH, and salinity also influence the survival of aquatic organisms and geochemical processes and can be measured directly. Various other chemicals also influence the survival of aquatic organisms; these can either be measured directly. However, the total dissolved solids (TDS), most often measured as electrical conductivity (EC) provides a more general indicator of pollution by dissolved substances. Often, concentrations of specific toxins, particularly those detrimental to human or animal health, such as pesticides or heavy metals, may be measured directly. Given the health risk posed to the transmission of waterborne diseases by faecal matter, the concentration of *E. coli*, a bacterium associated with faecal material provides a proxy of faecal contamination. Sediment loading includes contamination by inorganic (soil) and organic sources and has multiple impacts on ecological conditions by smothering or scouring of aquatic habitats, directly affecting the survival of aquatic organisms and influencing biogeochemical cycles, such as primary productivity. This is most accurately measured as the total suspended solids (TSS) but is closely related to water turbidity and clarity, which is also influenced by eutrophication and chemical discolouration.

Nevertheless, two integrated measures have made major advances in developing widely applicable indices (see below). Therefore, for a global RH assessment, the importance of water quality at influencing the condition of biological communities, the commonplace nature of the data, potential for future remote measurement of select variables, and flexibility to local conditions (namely the CWQI approach), make it highly desirable for consideration in a comprehensive integrative value. As a sole indicator of ecological health, however, it is inadequate as it fails to represent impacts to other components of the freshwater ecosystem that may not affect water quality.

3.2.5.5 Hydrological Indicators

3.2.5.5.1 Flow alteration indices

Several indices have been developed over the years to measure the degree of flow alteration, particularly due to dam development. The proportion of the mean annual runoff (MAR) of a river captured by a single reservoir or a cluster of reservoirs is a rough estimate of the degree of potential impact on downstream flows (Lehner et al., 2011a). A commonly used metric in this regard is referred to as Flow Regulation (Dynesius and Nilsson, 1994; Nilsson et al., 2005) or Degree of Flow Regulation (DOR) (Lehner et al., 2011b). Other metrics hydrologically equivalent to DOR are 'Change in Residence Time (CR)', 'Water Aging (WA)' (Vörösmarty et al., 1997) and 'Regulation Degree (RD)' (Steinmetz and Sundqvist, 2014). A similar metric is Hydraulic Dam Size (HD), introduced by Kibler (2017), to encompass not only water storage, but diversions and return flows as well. While all the above indicators may be used to accumulate the degree of regulation over an entire basin, the River Regulation Index (RRI) proposed by Grill et al. (2014) quantifies the overall impact on the basin by assigning higher weights to reservoirs placed on main stems of rivers than to those on upstream reaches.

Besides the indices on flow regulation, a number of other indicators are able to measure the degree of flow alteration in rivers. They include the well-known suite of 'Indicators of Hydrologic Alteration (IHA)' (Richter et al., 1996) which measures changes in 33 flow statistics encompassing the primary components of a flow regime (flow magnitude, frequency, duration, timing and rate of change). The method measures the central tendency and dispersion for the 33 flow parameters using time series of daily flow data under present day conditions which can then be compared with those under natural conditions using the Range of Variability (RVA) approach (Richter et al., 1997). The RVA quantifies the change in the range of variation of the 33 IHA parameters from the pre-impact to the post-impact period. The change in each parameter is categorized into high, medium or low categories in comparison to pre-defined targets, and a hydrologic alteration category is calculated based on the percentage of years the RVA target range is not attained. This results in a hydrologic alteration category for each of the 33 parameters expressed as high (68%-100% alteration), medium (34%-67% alteration) or low (0%-33% alteration). This quantitatively robust approach can indicate the degree to which different important flow components are being affected due to changes in flow regimes. However, since this method relies on long-term daily flow time series, its application is limited in data-scarce regions. Also, the large number of intercorrelated metrics may be redundant and complicated to estimate. Based on the original 33 IHA parameters other global alteration metrics have been developed by Black et al. (2005), Shiau and Wu (2008) and ISPRA (2011).

A different approach is used by the non-dimensional metrics Eco-surplus and Eco-deficit (including the suite of Eco flow statistics) which reflect the overall loss or gain in streamflow due to regulation on a flow duration curve (Vogel et al., 2007; Gao et al., 2009). Alternatively, Sengupta et al. (2018) and Gippel et al (2012) present flow alteration indices consisting of 39 and 9 different flow metrics respectively.

3.2.5.5.2 Environmental Flow Metrics

Environmental flows (EF) are defined as the '*the quantity, timing, and quality of freshwater flows and levels necessary to sustain aquatic ecosystems which, in turn, support human cultures, economies, sustainable livelihoods, and well-being*' (Arthington et al. 2018). The quantity, quality and timing of present-day water flows in a given river, and more specifically the quantity, quality and timing of flows that can be allocated as environmental flows, provides a measure of the hydrological component of RH. A simple comparison between the percentage of the mean annual runoff (MAR) that can is allocated as environmental flows and what should be allocated for the river to be in a "pristine" state may be used as a hydrology- based indicator of RH. An online tool that may aid this comparison is the Global Environmental Flow Information System (GEFIS) developed by IWMI (http://eflows.iwmi.org/; accessed on 13/12/2021).

Matthews and Richter (2007) developed a method to evaluate changes in ecologically important flow components which are represented by 34 flow statistics. These flow statistics complement the original 33 IHA parameters and are based on time series of daily flow data (Jumani et al. 2020). The ecologically important flow components are grouped into low flows, extreme low flows, high flow pulses, small floods and large floods. This approach can indicate the degree to which different ecologically important flow components are being affected due to changes in flow regimes. The method is supported by open-access desktop software developed by the Nature Conservancy.

Similarly, Bejarano et al. (2017) also presented a suite of ecologically relevant metrics to characterize the short-term effects of hydropower plant operations on within-day flow regimes

Flow-ecology response curves (embedded in environmental flow assessment methods such as ELOHA and DRIFT (Poff et al. 2010) and PROBFLO (O'Brien et al, 2018) attempt to account for ecological responses to flow alteration. Flow-ecology response curves combine hydrology, hydraulics, ecology and social processes to build links between hydrology and ecology based on river types (Jumani et al. 2020). These relationships are built for baseline conditions and present-day conditions using time-series of flow data, ecological data and expert opinion. The ability to be applied across a broad region, the ability to account for all water uses in a river basin, and the ability to be continually improved are some of the advantages of this approach. However, the requirement for extensive hydrological and biological data across large regions and the subjective nature of targets decided by stakeholders and decision makers are some of the drawbacks of this approach.

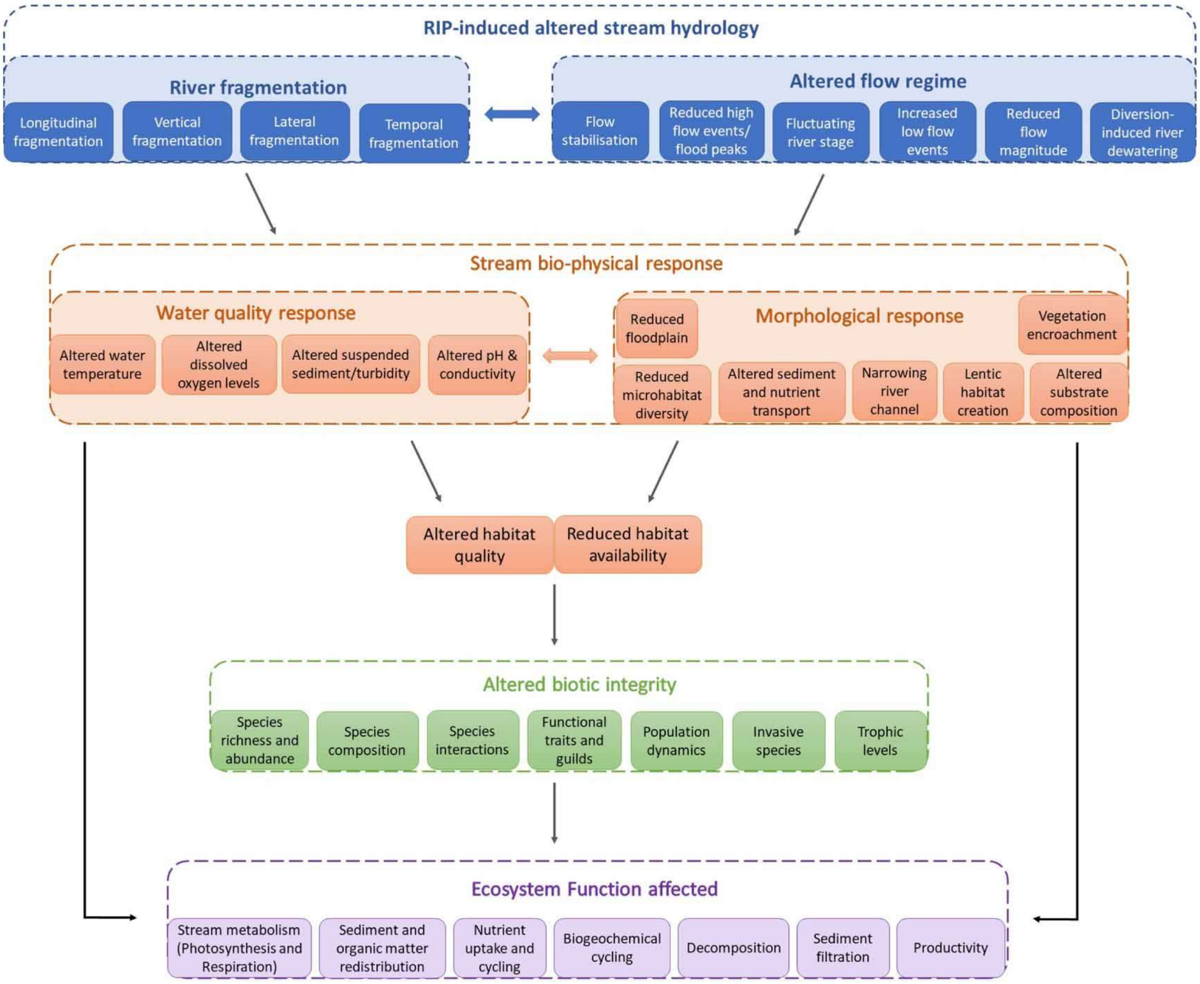


Figure 3.3 Schematic model illustrating effects of flow alteration and river fragmentation by river infrastructure projects (RIPs) on indicators of freshwater health. (Source: Adapted by Jumani et al. (2020) based on (Poff et al 1997)

3.2.5.6 Connectivity Indicators

River connectivity is critical for maintaining freshwater biodiversity due to the inherent linear structure of rivers and few alternatives to improve their status if longitudinal and/or lateral connectivity is lost (Tickner *et al.*, 2020). Several indicators have been developed quantifying different aspects of river connectivity, however only the Connectivity Status Index (CSI) includes all dimensions so is the most comprehensive and is presented below.

3.2.5.7 Habitat Indicators

Index of Habitat Integrity (IHI): There have been many developments of general habitat indicators. Profiled here is a summary of the two tiers of the Index of Habitat Integrity (IHI) that form part of the EcoStatus framework from South Africa (Kleynhans *et al*., 2008). The habitat integrity of a river refers to the maintenance of a balanced composition of physico-chemical and habitat characteristics on a temporal and spatial scale that are comparable to the characteristics of natural habitats of the region The IHI methodology assesses the habitat integrity by considering the current condition of instream and riparian zones. The assessment of the integrity of each zone is based on the appraisal of metric groups, each of which has a number of sub-metrics.

The contents of the IHI are summarised below, here describing a stream that is "largely natural with few modifications: A small change in natural habitats may have taken place but the ecosystem functions are essentially unchanged " i.e., a "B" class or category on an A-F scale.

- Physical drivers:
    - Hydrology: The flow regime has only slightly been modified
    - Geomorphic: limited to slight sediment changes
    - Physico-chemical changes: Water clarity may sporadically be slightly influenced. At worst, only sporadic traces of toxics present. Salts may sporadically be slightly increased.
- Associated habitat conditions:
    - Instream: Very little change in habitat types and their dimensions and frequency. Connectivity between habitats virtually unchanged.
    - Riparian: Riparian habitat close to natural in terms of biophysical characteristics. Very little modification and use of riparian zone. Virtually no fragmentation.

The assessment is based on an interpretation of the deviation from the reference condition (i.e., least-impacted condition). Deviation from reference conditions is determined using an impact-based approach where the intensity and extent of anthropogenic changes are used to interpret the impact on the habitat integrity of the system. This information is obtained via site visits, surveys and / or other available data sources. Changes are interpreted in terms of modification of the drivers of the system, *viz.* hydrology, geomorphology and physico-chemical conditions and how these changes would impact on the natural riverine habitats.

The advantage of the approach to global RH is that a high-level IHI can be carried out using Earth Observation and layers of mapped geospatial data. Accuracy is increased by adding *in situ* perspectives.

IFIM and PHABSIM: The Instream Flow Incremental Methodology (IFIM) and its Physical Habitat Simulation Model (PHABSIM) were devised by the United States Fish and Wildlife Service to assist in the assessment of instream flow requirements of rivers (Bovee, 1982). The model sets out to describe the physical microhabitat for selected target species, especially for fish.

It acts by simulating hydraulic conditions over a range of discharges which are then linked to habitat information on selected riverine species. Foundational for this is the development of libraries of species and "habitat curves" showing the preference of a named species for specific water depth, water velocity and substrate and cover conditions. Such libraries have been compiled in the United States of America where the model has been widely applied (e.g., Hatfield and Bruce (2000). The model was in the early days rejected for use in South Africa because of its complexity (King and Tharme, 1994). The model was also recommended for the UK (Spence and Hickley, 2020).

PHABSIM is recognised for its contribution to the development of the science, however the approach has met with some resistance (Railsback, 2016; Williams, 2010) where two general problems are identified. First, PHABSIM is a habitat selection model (HSM)—but does not conform to modern practices of ecological modelling, with a particular weakness identified as its weak ability to adapt to changed spatial scales which are driven by hydraulic considerations and not biological preferences. Also, HSMs, in general, are not well suited for many instream flow decisions as they cannot consider variation in flow over time, whereas dynamic flow regimes are now considered essential, HSMs also do not make testable predictions of fish population responses. An empirical examination of PHABSIM for salmon found that the results did not match with measured habitat preference (Beecher et al, 2011). Alternatives to PHABSIM include analyses based on explicit understanding of species ecology, individual-based models, and more powerful modern habitat selection modelling methods. The likes of Poff et al. (2010) are proposed as alternatives.

For a global RH programme, the PHABSIM approach is founded on *in situ* data that needs to be collected at an intensity that is inappropriate for large-scale implementation. The model also is not amenable to upscaling, the priority for development of a global framework. Its focus is also on environmental flows and not on river health.

Land-use changes and river habitat: Many freshwater ecosystem impacts are driven by changes in land use. Given the readily available nature of land use data with global coverage, tracking changes in land uses or covers are thus often used as driver indicators of freshwater ecosystem conditions in macro-scale assessments. For example, the FHI (Vollmer *et al.*, 2018; Bezerra *et al.*, 2021) considers 'land cover naturalness' one of three indicators of basin conditions, as one component of ecosystem vitality, alongside water quality, water quantity and biodiversity. They found strongly relationships between other components of RH (e.g. biodiversity and water quality) and naturalness, indicating it is useful as a proxy of RH. Land cover is also used in the Incident Threat Index (Vörösmarty *et al.*, 2010). The MIF also considered, as a component of RH, the status of environmental assets, which included the 'condition and status of ecologically significant areas' (i.e., percentage of original area of forests and grasslands). Land use change are sometimes more specifically applied for the riparian zone, such as 'riparian vegetative cover' and 'riparian disturbance' as physical indicators of freshwater ecosystem health in the US NRSA (USEPA, 2020a), and riparian vegetation disturbance in the Murray-Darling Basin SRA (Davies *et al.*, 2010).

The advantage of considering the impact of land-use on RH is that data is readily available at a global scale, however, at smaller scales land use indicators become less useful as indicators of instream RH. They should thus only be used at larger scales or where there is a lack of direct indicators. The major limitation of this approach is that it does not measure RH directly.

3.2.5.8 Spatial Extent Indicators

The spatial extent of water-related ecosystems is the main indicator used to indicate the condition of aquatic ecosystems in the SDGs, in particular SDG 6.6.1 (UN Water, 2017), where information on the current global status is available using the online portal (https://www.sdg661.app/home). This reporting is supported by readily available remote sensing data on the extent of open water bodies at the global level, however limitations to the resolution of publicly available data mean that the measurement is not appropriate for most rivers.

Although modifications to the physical form of rivers is a cause of ecological deterioration, it is far less of a threat to these ecosystems than changes to flow, pollution, over-exploitation, fragmentation, invasive species and impacts to instream or riparian habitats that do not cause major changes to their spatial extent (Tickner *et al.*, 2020; USEPA, 2020b). For example, in the SEEA National River Ecosystem Accounts for South Africa (Nel and Driver, 2015), although the overall ecological conditions (determined from flow, water quality, instream and riparian conditions) of the country's rivers deteriorated by around 10 % between 1999 and 2011, there was no change in river length. River width could not be reported due to lack of data.

The exception will be for very wide shallow rivers where the spatial extent would be reduced by a drop in discharge, and also for floodplains and riparian wetlands the extent of which may also be changed by reductions in discharge.

The major limitation of this indicator for a global RH framework is that it does not account for impacts that do not cause changes in the area, such as pollution, invasive species, over-exploitation or fragmentation, such that the use of ecosystem extent will always underestimate actual ecosystem conditions. The use of extent as the predominant indicator of freshwater-related ecosystem conditions by SDG 6.6.1 is thus likely to result in a drastic underestimation of conditions. Spatial extent is thus inappropriate at indicating the conditions of most rivers and should only be used to contribute to an overall assessment.

Wetted perimeter: The wetted perimeter of a river is a sub-set of the spatial extent and is commonly used in environmental flow assessments to indicate an important aspect of RH and the role of discharge in maintaining the habitats of a river. Its implementation is described by Gippel and Stewardson (1998). This approach may not be appropriate for a global RH framework as useful data is normally gathered at an *in situ* scale, however advances in remote sensing may allow at least larger rivers to be evaluated in this way. Required would be the surface water extent coupled with a digital elevation of the river-bed.

3.2.5.9 Ecological Processes

Ecological processes includes the interactions between life forms and the environment and can be represented by a combination of indicators of the status of dynamic biochemical processes (Xie *et al.*, 2020). Ecological processes are thus included in the definition of ecological conditions for several regional frameworks, including the WFD, FBEHF and RHI. However, they are notoriously difficult to quantify and measure so are more often than not excluded from measurement, including in the WFD and RHI.

Probably the most used indicator is primary production, as a measure of the growth of autotrophic organisms (plants, algae & phytoplankton), which form the basis of many freshwater food chains. The most used indicator of this in aquatic ecosystems is Chlorophyll *a*, used as an indication of biomass and productivity, for which data is already recorded by satellite and readily available for application at the global scale (forming part of SDG 6.1.1). However, its value is hampered by the coarse spatial resolution of satellite sensors unable to record values for smaller rivers and streams. The maximum spatial resolution is currently 10-30 m but with advances in satellite technology and new planned launches, specifically aimed at measuring Chlorophyll *a*, it is expected to improve. A further limitation of this index is that in turbid systems, light limitations mean that Chlorophyll production is suppressed, giving a false indication of water quality status.

Globally, the 'Planetary Boundaries' framework includes the assessment of human impacts to the N & P cycles, which are the main nutrients involved in ecological productivity. The remote detection of elements central to biogeochemical cycles thus offers another means of assessing the condition of ecological processes.

For a global RH framework there is potentially a role for inclusion of ecological processes, but for most smaller rivers the data is not available.

3.2.5.10 Ecosystem (Social) Service indices

Contributions of ecosystems to human well-being and livelihoods are indisputable, with fundamental benefits derived from ecosystem services (ES) perceived to contribute significantly to making human life possible and worth living (Costanza et al., 1997; MEA, 2005). The Millennium Ecosystems Assessment (MEA 2005) classified ES into four major categories *viz.* **provisioning services**, **regulating services**, **supporting services** and **cultural services**, which although the subject of some criticism, continue to be widely used. These services are explained in Table 3.7. They have also been defined as the contributions of ecosystem structure to human well-being (Burkhard *et al.*, 2012; Burkhard and Maes, 2017) while most recently, the Inter-governmental Panel on Biodiversity and Ecosystem Services (IPBES) have publicised the term for ES as “nature's contributions to people” (IPBES 2018).

Table 3.7: CLASSIFICATION OF DIFFERENT TYPES OF ECOSYSTEM SERVICES (MEA 2005)

PROVISIONING SERVICES	Tangible products obtained from ecosystems, including for example, genetic resources, food and fibre and freshwater.
REGULATING SERVICES	The benefits obtained from the regulation of ecosystem processes, including, for example, the regulation of climate, water and some human disease
CULTURAL SERVICES	The non-material benefits people obtain from ecosystems through spiritual enrichment, cognitive development, reflection, recreation, and aesthetic experience, including, e.g., knowledge systems, social relations and aesthetic values.
SUPPORTING SERVICES	Ecosystem services that are necessary for the productions of all other ecosystem services. Some examples include biomass production, production of atmospheric oxygen, soil formation and retention, nutrient cycling, water cycling and provision of habitat

It is indisputable that ecosystems are intimately tied to human wellbeing, and some RH monitoring frameworks have already included ecosystem services, e.g., the FH, SEEA, RHI, IECA. Such reporting however serves a wider objective that includes the human perspectives of RH, going beyond the direct measure of ecosystem health. Inclusion in a global framework of RH would thus depend on the objectives of the monitoring framework. It could be argued that a future global RH framework should be limited to a direct evaluation of the ecosystem alone, which data could then be related to social needs as a separate exercise.

3.2.6 Approaches to data processing

The methods used to process data are core to the functionality of successful large-scale monitoring frameworks. Data processing usually involves three steps. 1) The aggregation of raw data to the appropriate scale for each metric. This has important implications for the representativeness and scalability of the framework. 2) Data are standardised to a common scale, to ensure consistency and flexibility. This often involves the comparison to reference data. 3) The integration (or combination) of data at the indicator, component, or overall ecological condition levels for reporting. This often involves the incorporation of weighting and has important implications for flexibility and representativeness. We also consider means to deal with missing data.

Generally, aggregation is carried out before integration, as it is more efficient (prevents 'over-sampling' of some indicators), simpler and can reduce the confidence intervals around estimated values (Robinson W, 2017), see Figure 3.4. Indicators can also be sampled independently of one another, using scales and site networks most appropriate to them. The downside is that this can result in more logistical effort for fieldwork. Integration before aggregation has the benefit of simpler coordination of sampling efforts. However, it can be problematic to integrate indicators that are relevant at different scales and requires data at the same scale for all. It is also more susceptible to missing data. Aggregation-before-integration is therefore the approach preferred by most frameworks.

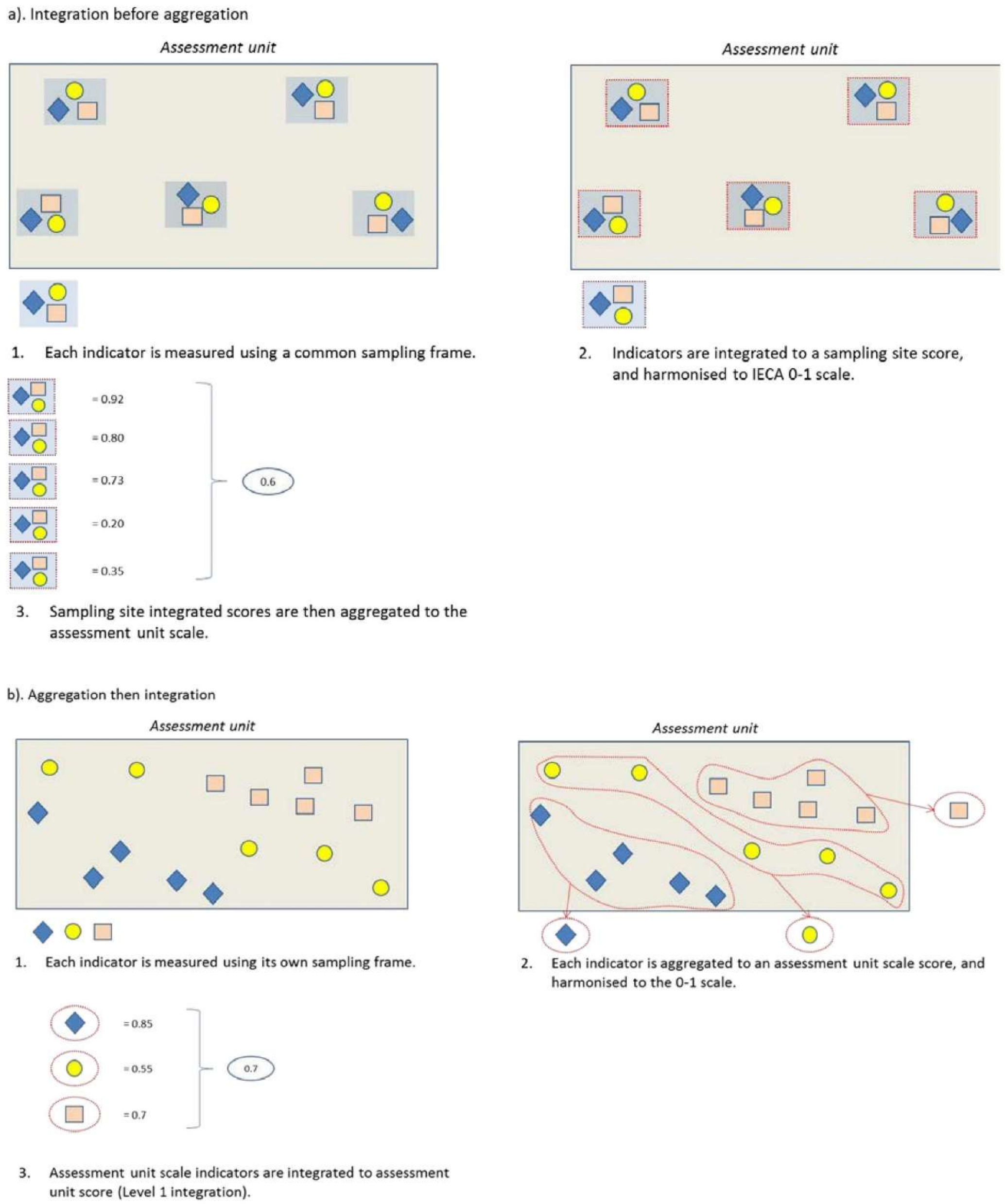


Figure 3.4: Graphic representation of methods for a) integration before aggregation and b) aggregation before integration (sourced from IECA (Department of the Environment and Energy, 2017).

3.2.6.1 *Data aggregation and scaling*

Aggregation is the compilation of data for a specific indicator variable. Scaling is its aggregation to a specific scale that differs from the scale at which it was recorded. This is vital to a successful global RH assessment as different indicators are generally recorded by independent sampling systems with variable sampling scales but require methods for manipulation to the assessment unit scale, or the scale required for reporting (e.g. catchment, management unit, national, regional etc). There are various means of aggregation, most involving some degree of inference and inaccuracy. Therefore, the method(s) chosen are important for a robust and representative framework and depend on a) the nature of the variable in question, b) the scale at which the data are collected, and c) the assessment unit scale (Hsu et al. 2013; Department of the Environment and Energy, 2017; Clapcott et al., 2018). In most cases the data are gathered at scales smaller than those required for reporting, in which case they must be *upscaled*. However, when data are gathered at scales larger than the scale required for reporting, they must be *downscaled*.

Upscaling

Methods of upscaling (aggregation to larger scales) include averaging, extrapolation (i.e., modelling) and summing of data. The choice of method depends largely on the measurement scale of the data involved.

Averaging: Averaging is the calculation of an average score for the assessment unit from smaller scale units. This is well suited for data gathered at a series of points. When all data points are collected at the same scale with a known selection probability, then one may simply average the results. However, when sites are not selected using equal probabilities then the baseline conditions of the assessment unit scale must reflect these probabilities when setting the reference conditions. For example, if for a biotic integrity index, 5 out of 10 reaches of ecotype A (inclusion probability of 0.5) are assessed but only 5 out of 50 of ecotype B (inclusion probability of 0.1), then the baseline reference condition must reflect this. Generally variables should be aggregated by ecologically consistent units i.e. ecotypes (Clapcott *et al.*, 2018). It is thus advantageous to aggregate data from sites selected with unequal probability and to calculate confidence intervals for the assessment unit. The main drawback is that scores can tend to a central (mean) value, so the range and variability can be lost. This is particularly true if the categorisation of ecological units is too coarse for the scale of assessment.

Modelling (Extrapolation): Extrapolation is the prediction of a score value at a scale larger than that at which the data are collected. It is best suited to indicators measured at an assessment unit scale. For example, in freshwater it is commonly used to predict hydrological values (e.g., flow volumes) from data collected at a single station or even using hydrological models created at the sub-catchment scale. Another example is the estimation of the total number of species present in the assessment unit from species accumulation curves for sub-units with a set scale. The disadvantages of this method are that is it heavily reliant on the quality of the data collected (i.e., the 'rubbish in, rubbish out' principle), it adds modelling errors to all predictions (over and above sampling predictions), and is complex if site selection probabilities are unequal.

Summing: Summing is the calculation of a total score for the assessment unit from the sum of the sub-unit scores. This is suited to indicators with discrete values (counts etc.) and a baseline scale larger than the scale of measurement. For example, total species richness for an assessment unit can be calculated by adding together lists from sub-units. Also, proportions of a feature in the assessment unit can be upscaled by summing raw frequency or presence/ absence data. The disadvantages of this method are that it requires measurements for all sub-units of the assessment unit, it can be complex if site selection probabilities are not equal, and the confidence intervals are wider than for the averaging method.

Downscaling

Modelling (Interpolation): Interpolation is the prediction of score values at a scale smaller than the scale of measurement. For example, flow volumes may be predicted at scales smaller than the scale at which they are modelled using geo-static techniques (Lehner and Grill, 2013). This is inherent in the 'Natural Discharge' estimates of WaterGAP v2.2. (Döll, Kaspar and Lehner, 2003), which downscaled flow data from a resolution of 0.5 degrees to 15 arc-seconds. The disadvantages of this method, like extrapolation, are that it is reliant on the quality of the measured data and model errors are inherent in the predicted values.

Disaggregation: Weighted division is the calculation of scores for assessment units by the division of the single value measured a larger scale. This assumes that features are evenly distributed in the landscape. Generally, however, features are not evenly distributed within an area so values may be further refined by weighting with related features. For example, pollutant volumes may be calculated using the ratio of population sizes, whilst area-related features may be weighted by area etc. The Red List Index also uses disaggregation to determine regional red lists by weighting by the fraction of each species' distribution occurring within a particular region, building on the methodology published by (Rodrigues *et al.*, 2014).

3.2.6.2 Data standardisation

Standardisation is simply the conversion of measurements from different series to a common scale, most commonly 0-100 or 0-1. It is necessary to enable the comparison between indicators with different measurement units or across different spatial scales and is a crucial step before the later data integration. In many cases, the scale is further broken up into categories of ecological conditions with a corresponding colour-scale, which is easy to understand.

The most used means of standardisation are ***Ecological Condition Ratios.*** These represent the actual ecological state by expressing an observed value as a ratio or percentage of a reference condition, which is usually represented on a scale of 0 – 100 % (100 % being the reference condition) or 0 – 1 (1 being the reference condition). They thus require benchmark conditions to be set, against which observed values can be compared. Usually, this is the natural conditions of an ecosystem but can also be applied to artificial systems or varied according to management scenarios (see Section 3.2.6.3: Defining Reference Conditions (or Benchmarks below). The greatest strength of this method is that the scale is representative of actual conditions. For this reason, it is the preferred approach and used by all the regional-scale frameworks reviewed. However, it is restrictive in the need for knowledge of appropriate benchmark values, which often necessitate an extensive network of reference sites, which has thus far restricted its use at the global scale. The approach differs slightly depending on whether one is using raw data or indices (such as a biotic index) (Department of the Environment and Energy, 2017) and numerical or categorical data. For raw data, the values are simple to convert to a common scale, but the choice of the value is important to being able to compare conditions across spatial scales. For example, for 'fish abundance', a value of 'catch per unit effort' allows comparison across scales, whilst raw abundance data do not. For indices, on the other hand, the conceptual relevance of standardisation must be considered, as the interval units on a common scale may not be comparable between different indicators. This is particularly true for indicators with non-linear relationships to the ecological aspects they represent, as the intervals on a common scale (i.e., 0-1) are not comparable to indicators with linear or different non-linear relationships. This may, however, be dealt with statistical transformation to a normal distribution. However, to ensure the conceptual relevance of the scale for different indicators an oversight committee should be in place to confirm this.

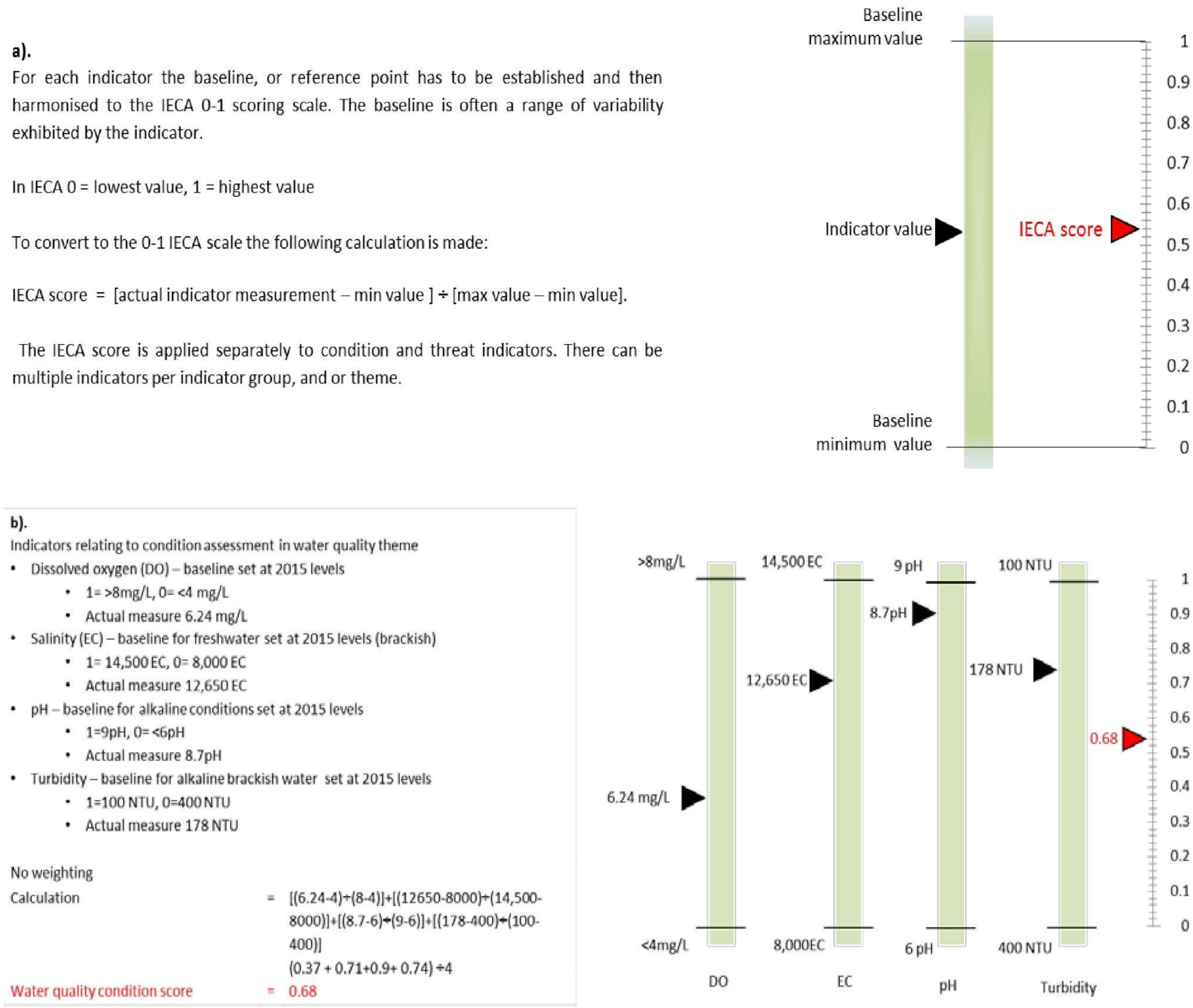


Figure 3.5: Hypothetical example of calculation of an ecological condition ratio (i.e. the IECA score) for water quality in the IECA framework (Department of the Environment and Energy, 2017)*. Although the example is for an estuarine ecosystem, the means of calculation remain the same.*

The **boundary crossing** approach used by the Planetary Boundaries framework is unique in quantifying the risk of irreversible and runaway environmental change if tipping points in biophysical systems are crossed (e.g., 350ppm CO_2 as the boundary to runaway climate change, or an extinction rate of <10 **E/MSY** for biosphere integrity). The greatest advantages of this method are that is does not rely on reference data and is less reliant on *in situ* sampling (with many variables measured remotely). However, it requires a sound understanding of the tipping points present in any given system. Although this is relatively well established in the global framework, it is complicated to understand at smaller scales and heavily reliant on knowledge about the relevant tipping points (both in literature and expert opinion). Uncertainty over the boundary values (both intrinsic and due to lack of understating) are thus a major weakness of this method. However, a conservative approach ensures that when a value enters the 'high risk' zone, it is almost certainly real, but this means it may overlook unknown boundaries.

Where the scale depends on the measured values, it is relative. One of the most commonly used relative approaches is the ***proximity-to-target*** method. This involves comparing a given measurement to a fixed target, usually based on the relevant local, national or international standards, which are often empirically based (Hsu *et al.*, 2016). As such, it is most suitable for pressure or driver indicators, and is commonly employed for monitoring compliance to water quality standards (e.g., pollutant concentrations etc) for which standards are widely available. It could, however, be applied to targets for any component, depending on the availability of suitable standards. Scores are then usually also represented on a scale from 0 to 100, with 100 being the target value and 0 the value farthest from the target. The scale thus depends on the 'worst performer', which makes it relative. The strength of this method is that it does not require knowledge of reference conditions so can be used in where reference conditions are unknown, yet it is still measured relative to set environmental standards, making it more robust and informative than purely relative means (see below). As such, it is the chosen method by the EPI as a measure of environmental performance. However, a major shortfall is that it is not representative of actual conditions and can be misleading as a means to indicate ecological state. In addition, since the low performance value varies between assessments, it makes comparison between assessments and trends difficult to see. However, it is possible to rank the performance between countries (i.e., quantify improvement or deterioration) and overall trends for a component can be calculated independently for the same data.

Where both the min and max limits are based on measured values, the method of standardisation may be referred to as ***Non-dimensional Scaling***. Values may also be given on a scale of 0 -100 or 0-1 but in this case, they are entirely relative to each other (not to reference conditions or standards) with higher values simply having 'better' and lower values 'worse' connotations. The advantage of this method is that it does not require any knowledge of reference conditions or even standards so is more widely applicable. It also enables the comparison between aspects of freshwater health with fundamentally different natures, as varied as ecological vitality, ecosystem services, stakeholder values and governance, and covering all DPSIR indicator classes (see Figure 3.2). This is the approach taken by the ITI (Vörösmarty *et al.*, 2010), which quantifies the relative levels of 'threat' of driver loadings to human water stress and freshwater biodiversity loss, and the FHI (Vollmer *et al.*, 2018; Bezerra *et al.*, 2021), which is aimed at enabling greater flexibility in application at a global scale. However, it is not representative of actual conditions. Nevertheless, if applied to areas with a complete spectrum in the scale of an indicator variable, such as areas with both natural and completely degraded conditions for a particular state indicator, the scale would be very similar to the actual condition ratio. However, where natural or completely degraded conditions are absent, the scale is then set by the 'best' and 'worst' performers. Also, like the proximity-to-target method, it is also not suitable to use for comparison between studies or to determine trends.

3.2.6.3 Defining Reference Conditions (or Benchmarks)

Reference conditions are standards or benchmarks for indicators against which changes can be measured without which indicators are meaningless. They are thus vital to ensure consistency in the assessment of freshwater ecosystems components and must be set for all indicators independently. A variety of measures and methods for determination may be used.

The '***reference state'*** for each indicator is usually the preferred option for measures of ecological state. Generally, this is defined by the ranges of indicator values occurring under natural conditions as the ***'natural state'***. However, given that most places are subject degree of human impact, several frameworks prefer to use ***'least' or 'minimally disturbed'*** sites for the reference condition and have stringent criteria for how these are defined and chosen. However, this can be problematic in that the levels of disturbance may vary between locations and assessments. However, the reference state may also be defined for artificial waterbodies in line with management objectives. The WFD defines reference conditions for artificial bodies as their '***maximum ecological potential***'.

On the other hand, '**guideline values**' can be used to measure ecological state as well as assess the threat or risk of ecological drivers or pressures on ecological state and thus have greater applicability. This may involve setting **target values**, which are the values that indicators are expected to achieve to meet the relevant management objectives. For example, water quality targets are usually based on the ranges of values for indicators to meet the specific management objectives (e.g., good ecological health, risk to human or organismal health, suitability for irrigation etc.). For assessments of driver or pressure indicators that assess that the threat they pose to ecological state, **trigger values** are preferred (e.g., ITI and PB). Trigger values are set upstream of values for ecological thresholds that, if crossed, acts as an 'early warning' detector that a threshold is being approached and intervention is required to avoid irreversible widespread changes in ecological state. ***Ecological thresholds*** or ***tipping points*** are values of indicators that, if transgressed, result in runaway changes in ecological state, driven by positive feedback loops. Close to a threshold, a small change in a driver can 'tip the balance', triggering a massive change of ecological state (generally catastrophic). Given this and the innate uncertainty in threshold values, trigger values are preferably set upstream of the threshold following the precautionary principle.

The values to be used for the reference conditions can be determined in multiple ways. One of the most used methods is ***representative sampling***, where a comprehensive sampling network is established across all the ecological regions in question to determine the natural conditions for different ecosystems against which impacts can be compared. This is the preferred method of most of the regional frameworks reviewed. However, an alternative that relies less on an extensive sampling network and monitoring is ***predictive modelling***, whereby models are created to predict the natural conditions of the relevant indicators for the ecological regions involved. For example, the AUSRIVAS models predict macroinvertebrate assemblages that should naturally be present at locations throughout Australia based on natural environmental variables. These are employed by the Australian NRHP and SRA but have proven inadequate for measuring ecological conditions (Bruce C. Chessman, 2021) due to reasons discussed in Section 2.1.1.5: *The National River Health Program (NRHP), Australia* above. When this is not an option for the location in question, one may use representative sampling or predictive modelling from ecologically similar ***natural or least-impacted sites elsewhere***. Alternatively, ***historical*** or even ***palaeoecological data*** may be used to set reference states or target values where site or model data on present natural conditions are unavailable (including due to the lack of locations in minimally disturbed condition). However, for many indicators of ecological state, knowledge of reference conditions is incomplete or altogether lacking in many parts of the world. This is a major limitation to the use of these indicators and the overall functionality of the framework! In such cases, one may instead use the available **policy guidelines, scientific literature, and expert judgement** to determine suitable guidelines values as a temporary placeholder whilst reference conditions are determined. Guidance by the WFD on the selection of the approach used for the determination of reference conditions according on the availability of data is given in Figure 3.6. However, for the assessment of the risk posed by diver or pressure indicators to ecological state, this is generally the preferred method employed (e.g., ITI and PB).

When setting reference conditions, it is also important to consider that baselines can be expected to shift over time (especially under a climate change scenario), affecting what is meant by the 'reference condition'. Two approaches could be followed here. The first involves changing reference conditions as the baseline shifts i.e., using contemporary conditions. This excludes the direct influences of climate change on the system and only consider non-climate change related impacts. The second would be to measure changes from a 'pre-climate change' reference. The point in time to serve as the reference period would need to be decided upon but popular options include 'preindustrial' conditions (e.g. SRA), 1970s (when the first satellite imagery became available and global warming started becoming evident), or 2000, when climate change impacts started becoming evident. In any case, one could then use past data or (where this is lacking) predictive models to determine reference conditions. This approach would include climate impacts to the ecosystem (relative to its past condition) along with non-climate impacts in the assessment of health. From a water management point of view, the first may be more useful. However, if a full spectrum of human impacts is desired, the second may be preferable. There is also no harm using both approaches to provide assessment in the current vs. historical contexts.

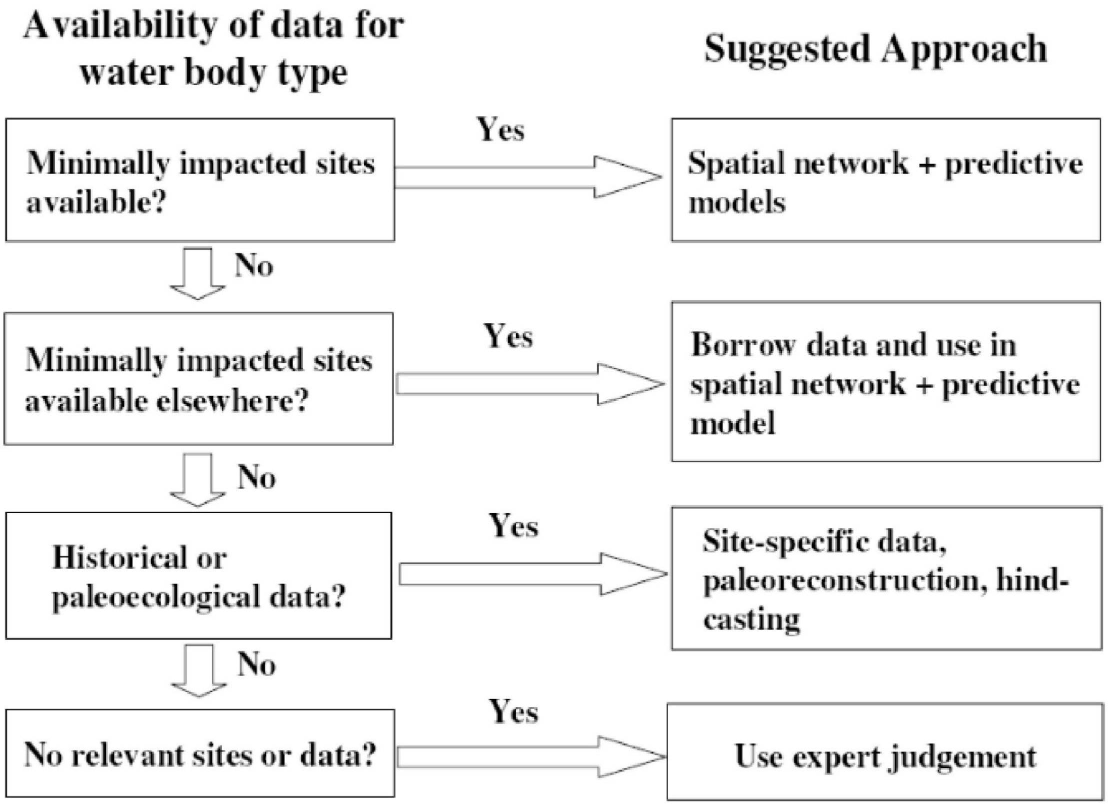


Figure 3.6: Guideline by the WFD for the selection of the best approach to determining reference conditions accordingly to the availability of data (from

3.2.6.4 Categorisation of conditions or risk

Categorisation is the conversion of numerical scores to categories. Values of condition or risk are present on a continuum so their division into discrete categories involves an inherent degree of subjectivity.

Ecological Condition Categories: Most frameworks that measure ecological state against reference conditions utilise *discrete ecological categories* to represent this. In these cases, category intervals are generally not equal, as impacts are generally non-linear in their influence on ecological conditions through the generation of positive feedback loops, so categories generally increase in the range covered with decreasing condition. For example, the SRA defines conditions as Good (80 – 100, near reference condition), Moderate (60 – 79, moderate difference), Poor (40 – 59, large difference), Very Poor (20 – 39, very large difference) and Extremely Poor (0 – 19, Extreme Difference). Similarly, the RHI classifies conditions as "very healthy" (RHI ≥ 90), "healthy" (75 ≤ RHI < 90), "subhealthy" (60 ≤ RHI < 75), "unhealthy" (40 ≤ RHI < 60), or "hazardous" (RHI < 40). Of course the limits of each category may vary according to the ecosystem involved so should vary accordingly. As such, the WFD classifies ecological conditions as five classes: 'high', 'good', 'moderate', 'poor' or 'bad'. These are precisely defined according to the degree of alteration but the exact values that correspond to these levels of change (e.g. slight changes for the 'good' category) are left up to the members states to decide (CEC, 2000). The REMP further recognises that there are often cases where there is uncertainty over which category a particular entity longs to, so they follow the "*fuzzy boundaries*" concept, where an entity can be considered to have characteristics of both classes e.g. A/B or B/C etc (see Figure 3.7.b). Although the processes of defining limits may seem arbitrary, it is in fact based on sound logical and statistical means. The process of defining category limits for any particular value is well represented by the NARS framework, where the categories: "Good", "Fair", or "Poor" are set according to the statistical distribution of the reference conditions (see Figure 3.7.c). Given that a small percentage of the "least-disturbed" sites used to determine references conditions would naturally be expected to have lower than normal levels of any particular indicator (i.e. some would be classified as "fair" and fewer still as "poor") simply due to variation and natural impacts, they therefore set percentiles to define the thresholds between categories. They consider the 25th and 5th percentiles to be the boundary between good and fair, and fair and poor conditions. Thereby measured values falling in the range of 75 % of the reference site values would be in "good" condition, those falling within the range of the 5 to 25 % percentiles would be in "fair" condition, and values below the 5th percentile would be classified as poor. This is effectively the process followed by all the above frameworks but not so explicitly explained. The selection of threshold values is where the subjectivity lies.

Pass or Fail: For pressure or driver variables for which standard limits are known (e.g., for pollutant concentrations), indicators can sometimes be indicated as simply a "pass or fail" or "above or below" the limit. This is included in the WFD for the parallel assessment of water quality as well as by the NARS. However, there is debate over what point a factor becomes 'impactful', especially for naturally occurring factors/ substances, whilst this also varies between ecoregions.

Risk and Trends: An alternative to the *ecological state categories* is the categorisation of risk to the desired conditions. For example, the PB considers the risk to changing earth systems relative to a defined boundary (upstream of a threshold) with an interval of uncertainty. Accordingly, the risk of destabilisation can be defined as "safe" when the observed indicator is below the threshold, "increasing risk", when it is in the zone of uncertainty, and "high risk" when it is above the boundary of uncertainty (see Figure 3.7.d). These are represented using the colours green, orange and red. The MIF uses a similar "traffic light" categorisation. However, this is based on both the risk (to the stability of the Mekong basin) and the observed trend and indicators are classified as having "immediate concerns", "some significant concerns to address" or "considerable concerns, urgent action needed" (see Figure 3.7.e). This is geared towards providing the information required for adaptive management but does not represent the actual present conditions. The ITI also utilises the concept of risk (or 'threat'), specifically to water security and river biodiversity, but represents this on the continuous scale without categorisation.

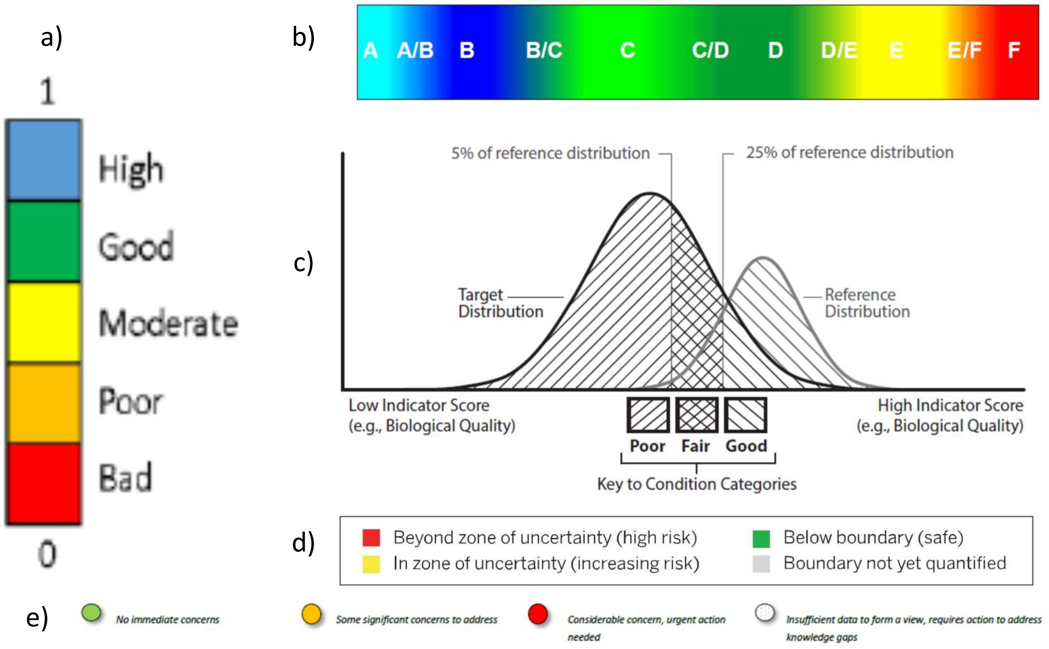


Figure 3.7: Examples of the categorisation of ecological conditions into a) fixed categories by the WFD, b) "fuzzy boundary" categories by the REMP, c) using thresholds by NARS, d) according to risk in the Planetary Boundaries, and e) risk & trends in the MIF.

3.2.6.5 Trends – an alternative to measured conditions?

A potential alternative to the need for standardised condition values is to look at trends in known driver, pressure or state indicator variables as this eliminates the need for complicated benchmarks. It is also informative, providing useful information to decision-makers and managers for the purposes of adaptive water resource management. Trends are thus included in virtually all frameworks (see Table 3.8), although usually alongside condition indicators. However, they are entirely reliant on the accompanying measures of condition or standards to convey any useful information. Without measures of condition or standards against which to compare indicator values, they are meaningless.

Table 3.8: Indication of which reviewed frameworks provide for the indication of trends for the indicators assessed.

	Regional									Global						
	WFD	NARS	REMP	RHI	NRHP	IECA	SRA	FBEHF	MIF	SDG	CBD	FHI	PB	ITI	SEEA	EPI
Trends evaluated	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	Y	Y	Y	Y

3.2.6.6 Data integration

Integration refers to the combination of different variables into single scores. In the context of a RH assessment framework, this is carried out at various levels. First, it may be used to combine multiple individual variables into single indicator scores, second multiple indicators into component scores, and third component scores into scores of overall ecosystem conditions. However, the IECA and FBEHF both advise against integrating across components (e.g., biological integrity and water quality) except for broad scale reporting. Integration can be done multiple ways.

'One out, all out': The 'one out, all out' approach involves taking the lowest score of the components as the overall score. See for example , where the overall condition is estimated as "poor" due to hydromorphological component scoring "poor", even though the biological and chemical & physicochemical components had "moderate" scores. This was the preferred method for the integration by the WFD (CEC, 2000), however, it is criticised by multiple authors as being far too conservative (Hering *et al.*, 2010), which limits its informativeness. It is also difficult to justify that a single component should define the quality of the whole ecosystem (Hering *et al.*, 2010).

Simple arithmetic mean: A simple averaging (arithmetic mean) of variable scores is one of the most used methods of integration. However, this assumes that variables are correlated. This is generally true between individual metrics of an indicator (e.g., indices 1, 2 & 3 of macroinvertebrate health) so is therefore the recommended approach by the IECA and FBEHF for integration at the indicator variable level. However, indicators and are generally assumed to act independently, so should never averaged in this way.

Weighted arithmetic mean: A weighted average on the other hand takes into account the varying degrees of importance of different variables in the set to be averaged. Therefore, this generally the recommended approach for integration of indicator variables into component scores and component scores into an overall score of ecological health, as it allows the relative importance of the indicators/ components to be represented e.g., the greater importance of biological conditions to overall ecological state. Deciding on the most appropriate weights to use and means of weighting are dealt with in Section 3.2.6.7: Weighting below. However, arithmetic means (simple and weighted) can obscure the performance of individual components (or indicators), especially poorly performing components (or indicators). For this, they components are usually not integrated at the assessment scale.

Geometric mean: The geometric means indicates the central tendency of a set of numbers from the product of their values (instead of the sum). This is offered as a potential solution to the masking of poor performers effect that the arithmetic mean has as it treats low scores more harshly (see Box below). A good geometric mean score, therefore, requires high performance across all components. However, the disadvantage is that it is difficult to understand and communicate (Hsu, Johnson and Lloyd, 2013).

Box 1: Hypothetical Example of Different Average Score Calculation Methods

Here we show the differences between the simple arithmetic, weighted arithmetic and geometric means for integration of 4 component scores for 3 hypothetical sites. For the weighted averages, component 2 was weighted twice the other components.

Table 3.9: Average Scores (Simple, Weighted & Geometric) of 4 components at 3 sites. For the weighted averages, component 2 was weighted twice the other components.

	Comp 1	Comp 2	Comp 3	Comp 4	Simple Arithmetic Mean	Weighted Arithmetic Mean	Geometric Mean
Site A	40	80	30	20	52.5	50.0	50.2
Site B	100	80	65	30	68.8	75.0	62.8
Site C	10	75	30	20	33.8	29.0	25.9

Calculations for Site A:

Simple Mean

$$\frac{40 + 80 + 40 + 50}{4} = 52.5$$

Weighted Mean

$$\frac{(2 \text{ x } 40) + (1 \text{ x } 80) + (1 \text{ x } 40) + (1 \text{ x } 50)}{5} = 50.0$$

Geometric Mean

$$\sqrt[4]{40 \text{ x } 80 \text{ x } 40 \text{ x } 50} = 50.2$$

3.2.6.7 Weighting

Weighting is the process of assigning a numerical factor to variables in a calculation based on their relative importance. Given that ecosystem components vary in their relative importance in contributing to overall conditions and indicator variables vary in their relative importance at indicating the conditions of ecosystem components by the locality, the weighting of indicators in components, and components in the overall measure of ecosystem conditions relative to their importance is crucial for a flexible and robust framework. It is thus used to denote greater or lesser importance to different indicators in components and components in an overall measure of ecosystem conditions. This is carried out during the process of data integration (see Section 3.2.6.6). The challenge with using a weighting system, though, is determining the appropriate weighting values to use, as the relative ecological importance of components and indicators varies by the type of ecosystem. For example, one river may be highly sensitive to disruptions in the hydrological regime and much less so to physicochemical impacts, whilst another may be highly sensitive to physicochemical changes and less so to hydrological changes. Within each component, there are also differences in the relative importance of indicators, for example, the physicochemical component of one river may be very sensitive to changes in nutrient levels, whilst another is less so. The challenge is to develop a weighting system that captures these differences in a consistent manner. Various systems are available and have been used by the various frameworks.

Most frameworks with integration use some sort of expert elicitation (Morgan, 2014) to weight components and indicators, as this allows for the greatest amount of flexibility in the framework.

This makes use of expert assessment to rank or prioritise variables and components. The REMP use the Multi Criteria Decision Analysis approach (MCDA) that allows for the development of consistent rating systems or indices for the categorisation of indicator variables in components and components in the overall condition and aggregates these mathematically in a theoretically justifiable way (Joubert, 2004). The principle they followed was one of 'ranking-weighting', where components (in the overall condition calculation) and indicators (in the calculation of component conditions) can be ranked by their importance for determining the ecological conditions. This can be determined by asking which indicator would most affect a component (or component the overall condition) if it changed from the best to worst ecological condition categories. These are ranked as 1 and given a weight of 100 %. Those with lower rankings (2, 3 etc) are dealt lower percentages, respectively. The difference in the change in percentage value is also changeable. In the REMP, this goes down to a resolution of 5 %. Where components are considered of equal importance, they are given equal weights (i.e., the two most important are both given weights of 100%). Similarly, the IECA recommends a similar 'sensitivity analysis', where components/ variables are prioritised on a scale of 1 - 3 according to their preference, importance or contribution to the objective (Robinson W, 2017). The variables (or components) may then be multiplied by a corresponding factor of 1 - 3 (i.e., 3 X the score for high priority components/ variables and 1 X for low priority ones). The expert opinion process is also recommended in the FBEHF. The ITI is the only global-scale framework to use expert assessment in the weighting of indicators within components and components to human water security and river biodiversity loss (Vörösmarty *et al.*, 2010).

Fixed systems are another possibility, whereby the weighting of variables and components is fixed beforehand by expert elicitation. This is the approach used by the RHI (see Section 2.1.1.4: The River Health Index (RHI), China above). Generally, this is decided beforehand based on expert opinion, but the inflexibility of weights makes the framework inflexible to regional differences. The FHI also uses fixed weights in calculating 'Ecosystem Vitality' but in this case, they give equal weights to the relevant indicators and components because they consider all to be important to the overall ecological condition. However, this is grossly inflexible to regional differences and reduces the representativeness of the results. For the non-ecological component, however, they use Analytic Hierarchy Process (AHP) (Saaty, 2005) to weight components/ indicators based on stakeholder perceptions of the importance of each (Vollmer *et al.*, 2018; Bezerra *et al.*, 2021).

Another means of weighting is through the *analysis of data* to base the weights on the values on the values of indicators and components themselves (i.e., a component with lower condition may be weighted higher). However, this is not actually implemented by any of the reviewed frameworks. The SRA, however, uses '*fuzzy logic'*, which supports a combination of weighting by data analysis and expert opinion. In other words, the weighting of each component in the overall ecosystem health score is based on its condition category and expert opinion over the relative importance of a reduction in condition by each category in all combinations of component conditions (Davies *et al.*, 2010). The 'Ecosystem Vitality' objective of the EPI also follows somewhat of a fuzzy logic' approach, as indicators are weighted based on the quality of the underlying data (i.e. more reliable data are weighted more) and their relevance to the component (Hsu, Johnson and Lloyd, 2013).

Oher frameworks leave the choice of weighting system up to each jurisdiction. This is also the principle followed by the WFD, IECA, SEEA etc. However, this raises questions over the consistency of the method and comparability of results from different studies and requires an oversight committee to guarantee that results from different jurisdictions are comparable. Furthermore, frameworks in which components are considered independently (not integrated), do not require a weighting system. These include NARS, NRHP, SDG 6, SDB, and PB. However, whichever method is used, what is most important is for the weights used to be clearly justified and the process transparent and open to review in order to foster trust in the framework (Department of the Environment and Energy, 2017; Clapcott *et al.*, 2018)

Table 3.10: Weighting systems used by the reviewed freshwater health assessment frameworks

	Approach to Weighting				
	Fixed Weights	Expert Opinion	Fuzzy logic	Variable	None - independent analyses
	Regional				
WFD				X	
NARS					X
REMP		X			
RHI	X				
NRHP					X
IECA		X			
SRA			X		
FBEHF		X			
	Global				
SDG 6					X
CBD					X
FHI	X				
PB					X
ITI		X			
SEEA				X	
EPI			X		

3.2.6.8 Dealing with missing Data

Missing Data: The ability to tolerate missing data is essential in a global index as these are inevitable in such wide-ranging datasets. Possibly the most suitable approach is that of the FHI, to simply ignore missing values and adjust the calculation of indicators scores (i.e., through adjusted weighting) accordingly. The greatest benefit of the FHI approach is that it shows data gaps where monitoring can be improved.

Another possible solution is to use **hot-deck imputation**, as used by (Srebotnjak *et al.*, 2012), where missing data is estimated through the use of results from an area with similar conditions (ecological & human). Another is the use of **predictive models**, which use more easily available proxies (e.g., river morphology/ hydrology, land uses, development statistics etc) to model ecological responses at scale. With extensive testing and empirical research, there is no reason why such modelling methods cannot be accurately employed. Geometric means, standard Euclidean distances, or simple averaging (IECA) can also be used.

3.2.7 Approaches to reporting of results

Reporting is the means of communicating assessment results to interested parties. It is a crucial step in the adaptive management process as effective communication of ecological conditions, threats, and trends, as relevant, is vital to determine what actions are needed (inform policy), assess policy effectiveness, and adjust monitoring programmes accordingly. It is also vital to support stakeholder engagement and public participation. Reporting is also the step most relevant to ensuring transparency, especially regarding clear reporting of methods, knowledge gaps and missing data.

Thus, it is central to the framework's informativeness. The approach to reporting also has important implications for the scalability of the framework by enabling effective and consistent communication of results across different scales of assessment.

The first consideration to reporting is the scale at which information is required. This is usually determined by the scale of information most useful to the management objectives and at which management decisions are made. However, some indicators may also require more detailed reporting to allow more informed management decisions to be made. Most regional frameworks regard the watershed-scale as the most appropriate for reporting ecological conditions as it is the scale at which most ecological processes occur and at which management actions to mitigate human impacts are taken. This is thus the preferred scale for the WFD, NARS, REMP, IECA, SRA and FHI. However, to improve the scalability of the framework, the IECA and FBEHF recommend a hierarchical or tiered approach to reporting to allow for reporting at different levels and scales of assessment (See Figure 3.8). This involves creating reports at different levels of detail. At the one end are the more detailed reports, from the publication of raw data to detailed assessments of individual indicators or components. These are highly beneficial for the management of certain sectors of water resources (e.g., management of pollution sources etc.). At the other end are more synthesised reports involving integration of scores from the site to global scales, which provide a more generalised overview of overall conditions at varying scales. From a management point of view, the synthesised reports thus indicate where problems may lie and the more detailed reports, the information required for actions to be taken. Most modern frameworks also include web-based reporting to allow public access to results.

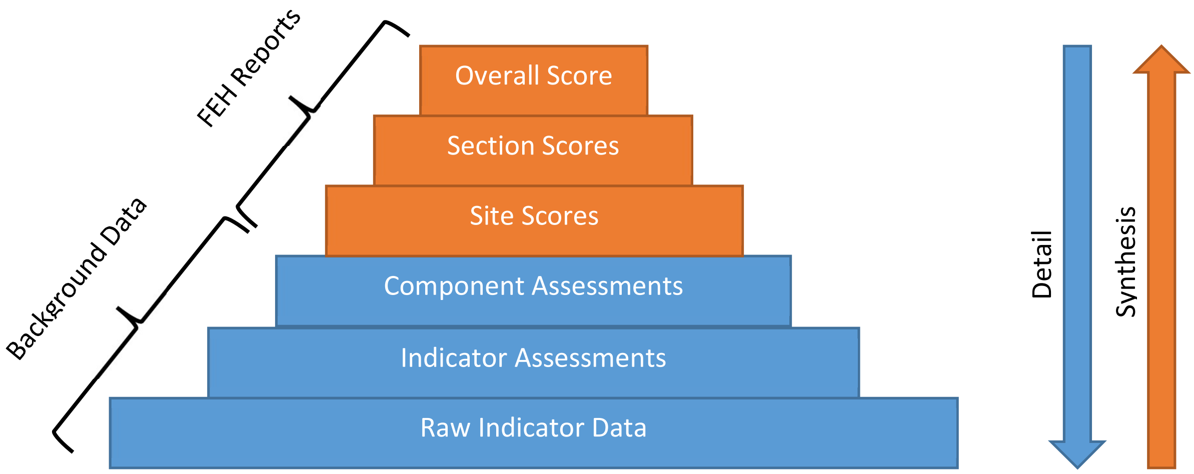


Figure 3.8: Hypothetical example of a tiered system of reporting supporting different levels of detail and synthesis (based on Fig 9 in (Clapcott et al.*, 2018)).*

The other important consideration for reporting is the means of graphic representation of results, for which there are several variations. Many use a colour spectrum to depict conditions or risk, usually some variation on the scale blue – green – orange – yellow – red (representing natural to poor conditions or low to high risk), which varies according to whether the scores are continuous or categorised and, if categorised, the number of categories used. The earlier frameworks (e.g., WFD, REMP) preferred to indicate score for indicators, components or and overall RH condition separately with coloured boxes or other graphics using the chosen colour scheme (see Figure 3.7a, b & e). However, some of the more recent frameworks prefer methods that allow the representation of indicator, component, and overall RH scores within a single graphic, which more closely corresponds to the actual ecosystem structure, which is composed of distinct components that all contribute to the ecosystem as a whole. Circular diagrams (Figure 3.9a & b) are one of the most commonly used integrated ways for reporting assessment results and simply involve a circle (representative of the ecosystem) divided into segments representing separate components. They are thus particularly beneficial in that they show components (and sometimes indicators) independently, along with the overall score at the centre of the diagram. The colour of the segments represents the conditions scores, whilst their length corresponds to the weighting of the component/ indicator involved, graphically representing their relative importance. They also easily account for the absence of results for certain indicators/ components (by leaving them 'grey'). They are thus very useful for reporting an overall summary of assessment results at all scales and are used by the FBEHF, FHI and EPI. A variation of the circular diagram that further divides segments into layers representing levels of risk to the ecological state based on triggers or threshold values can also be used to depict measures of threat or risk, such as used by the PB (Figure 3.9c). Radar diagrams (Figure 3.9d) are also suggested by the FEBEHF as a suitable means of representing conditions or threat. However, they are not able to represent the relative importance (weighting) of different components, nor can they show the different levels of assessment (indicators – components – overall scores) in a single diagram, as with circular diagrams. Maps are also highly effective at communicating the spatial distribution of results (e.g., Figure 3.9e). However, it is only possible to map single indices at a time, including indicator, component, and overall RH scores. Given the different strengths associated with each means of reporting results, most frameworks use a combination involving some kind of diagram or graph to show the relative conditions of different components and maps to show how these are spatially distributed.

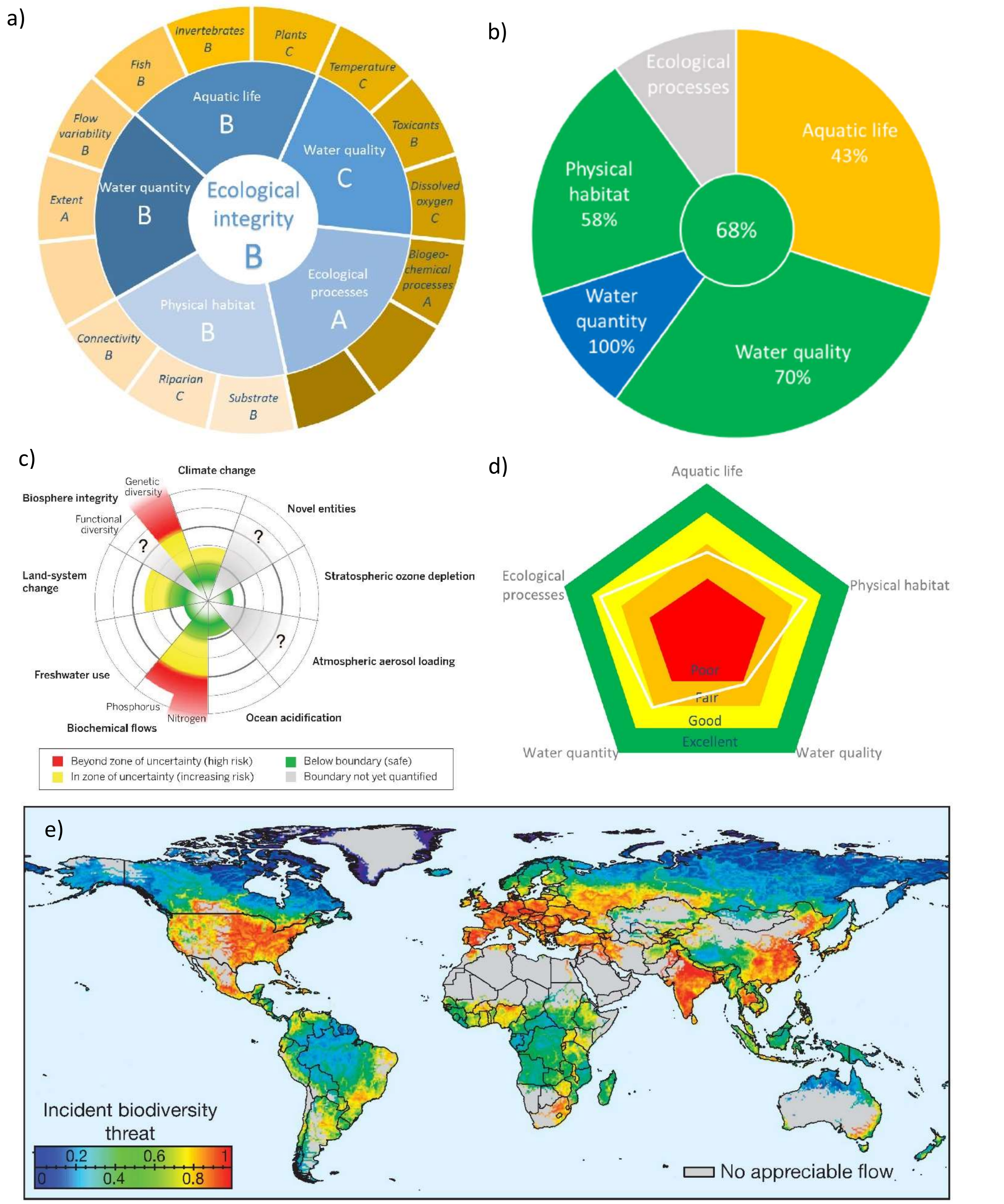


Figure 3.9: Approaches to reporting results by different frameworks including a) a circular diagram with different levels for the overall, component and indicator scores, b) a circular diagram with segment length representative of component weights, c) a variation of the circular diagram showing levels of risk (trigger values or thresholds) to causing changes in the ecological state per component, d) a radar diagram indicating conditions for each component at each angle, and e) a map of score values (in this case the Incident Threat to River Biodiversity).

3.3 THE DATA AND SCALE PERSPECTIVE TO RIVER HEALTH

Ecosystem health analysis requires a process of decision making that involves data acquisition and processing. In the endeavor of building global river health indices, the availability and quality of data constitutes a major concern. On one hand, thanks to the advance in data science and technology, some data that can be used to support ecosystem health assessment has already been readily available, which makes river health assessment at a global scale possible. On the other hand, limited data availability or data scarcity is often listed as a main barrier in ecosystem assessment or modeling (Pandeya et al., 2016). This problem may become particularly prominent in global ecosystem health assessment. During discussions at the workshops on river health at the global scale held on October 6-7, 2021 (see Section 4 GLOBAL WORKSHOP ON RIVER HEALTH), when asked what the main challenges are to constructing indices to characterize global river ecosystem health, many participants cited scarcity in input data as a major limitation. Given the importance of the data issue, we revisit the issue here from a data science perspective, which may shed additional insight into where we currently stand and where we could go.

3.3.1 Advances in data science and technology for RH monitoring

An important advance in data science in recent years has been the acceptance of the concept of big data by scientific communities, including in the disciplines of ecology and ecosystem health (Dafforn et al., 2015). According to the popular "4v" definition of big data, big data is characterized by large volume, large variety, high velocity and high veracity (Mayer-Schönberger and Cukier, 2013).

Behind the emergence of the concept of big data is the enhanced capacity for data acquisition and production. Three main types of data which could be used to support the RH assessment are 1) in-situ monitoring data, 2) Earth observation or remote sensed data and 3) modelled data.

Ecosystem monitoring data is conventionally collected in-situ, which is costly and labor intensive, especially at a global level. The invention and application of monitoring devices with continuous and automatic data collection and transmission functionalities greatly lowers the cost of monitoring and increase the sampling density and frequency. Some recent development in automation on river flow, water quality and bio-habitat are reported or review in Chapin (2015), Duffy and Regan (2017), Kawanisi et al. (2018), LeGrand et al. (2020).

Remote sensing refers to "*the process of detecting and monitoring the physical characteristics of an area by measuring its reflected and emitted radiation at a distance (typically from satellite or aircraft*)" (USGS, n.d.). Remote sensing has a long history of being applied to Earth observation (Kansakar and Hossain, 2016), and the list of environmental variables which can be monitored has been greatly expanded in recent years. Two notable examples use remote sensing techniques to estimate or monitor river stream flow and water quality (Bjerklie et al., 2003; Gholizadeh et al., 2016; Sichangi, et al., 2018; Kebede et al., 2020; Topp et al., 2020).

There has also been a rapid development in Earth system modeling. Models can be roughly classified into data driven models and physical-based models. Data driven models tend to work as "black boxes" and are constructed from raw data without resorting to prior knowledge of physical behaviors of system (Solomatine et al., 2008). The conventional tools that are commonly used in Earth system data-driven modeling include a range of parametric models, e.g., regression, time series and spatial interpolation models (Smith et al., 2003; Hipel and McLeod, 2005; Lin et al., 2011), while in the last decade the application of machine learning approaches, such as deep neural network, has gained momentum (Shen, 2018). As a contrast to the data-driven models, the physically based models incorporate the prior knowledge of processes governing the behaviors of the study systems. The advances in physically based Earth system models are reflected by the increased model complexity and improved spatial and temporal resolution. Take physically based global hydrological model as an example, more detailed groundwater modules have been developed (de Graaf et al., 2017; Reinecke et al., 2019) while global hydrological models in early years typically ran on a 0.5 degree latitude/longitude grid and at a monthly time step (Alcamo et al., 2003; Hanasaki et al., 2008), while nowadays the state-of-the-art-models have a spatial resolution of 5 arc minutes and daily temporal resolution (Sutanudjaja et al., 2018).

3.3.2 Spatially heterogeneous uncertainty and varied availability across variables- data challenge in global river health assessment

As noted, in-situ data, remote sensing data and modelled data all play important roles contributing to RH assessment. It is noted that while the site-based nature of in-situ data may hamper its use in large-scale ecosystem assessment, it should be borne in mind that the in-situ data are often considered superior to remotely sensed and modelled data in terms of data uncertainty and therefore should form an underlying component in the data hierarchy of ecosystem observations as shown in Figure 3.10. The acquisition and generation of remotely sensed and modelled data rely on the use of in-situ data.

Specifically, as hinted by its name, the data driven modeling approach is an outgrowth of increased availability of in-situ data. The performance of data driven models critically depends on the size and quality of training data (Millard and Richardson, 2015; Soranno et al., 2020). In physically-based modeling, the large-scale physically based earth system models often contain parameters whose values are difficult to determine via direct measurement and have to be estimated through calibration. In model calibration, model parameters are tuned to enhance agreement between model output and data representing historical states and behaviors of the study system, which traditionally consists of in-situ observations. There is rich literature on the methods and techniques for model calibration (Beven and Binley, 1992; Duan et al., 1994; Kennedy et al., 2001; Gupta et al., 2006). Likewise, the efficacy of the calibration hinges on the information content of the in-situ observation data set.

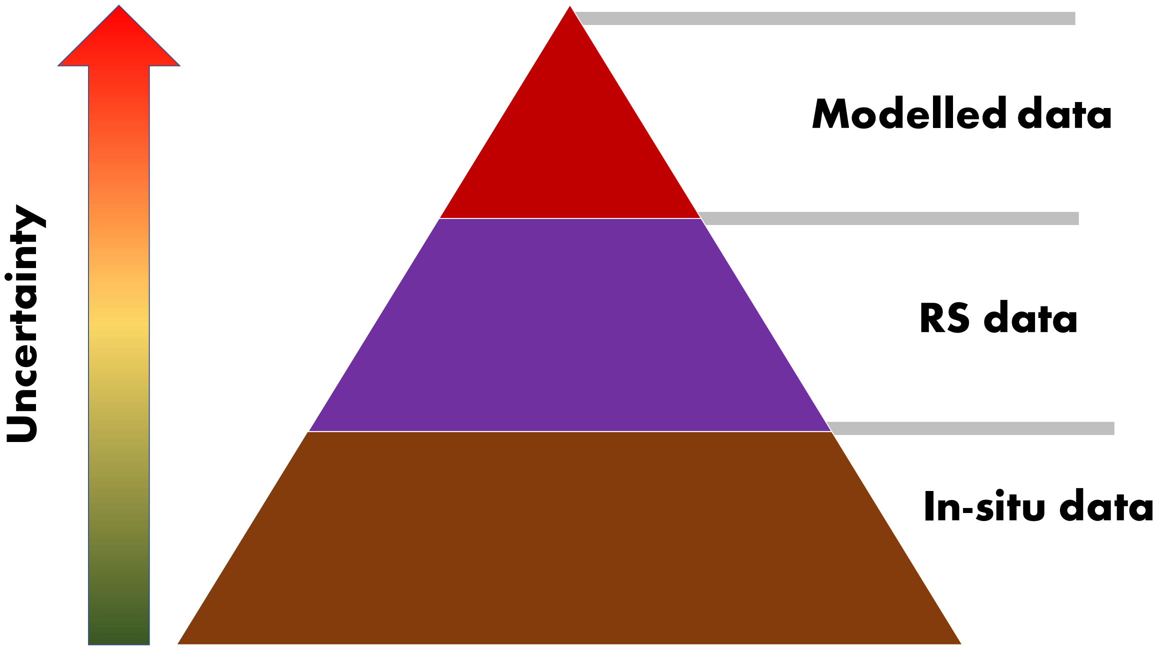


Figure 3.10 Main types and levels of uncertainty of data for ecosystem health assessment

It is clear that the availability and quality of in-situ data varies by region or country, or is generally better in developed countries and less favorable in the developing world. An illustration of this situation is provided in Figure 3.11, which shows the distribution of hydrostations and the length of time series of river discharge data at each hydrostation, to be found in the in global river discharge database compiled by Global Runoff Data Centre (GRDC) - a data set that is widely used for model calibration in global hydrological model development. We can expect that the heterogeneity of in-situ data availability is translated into uncertainty of the Earth system modeling, or the modelled data are subject to greater uncertainty in in-situ data scarce area.

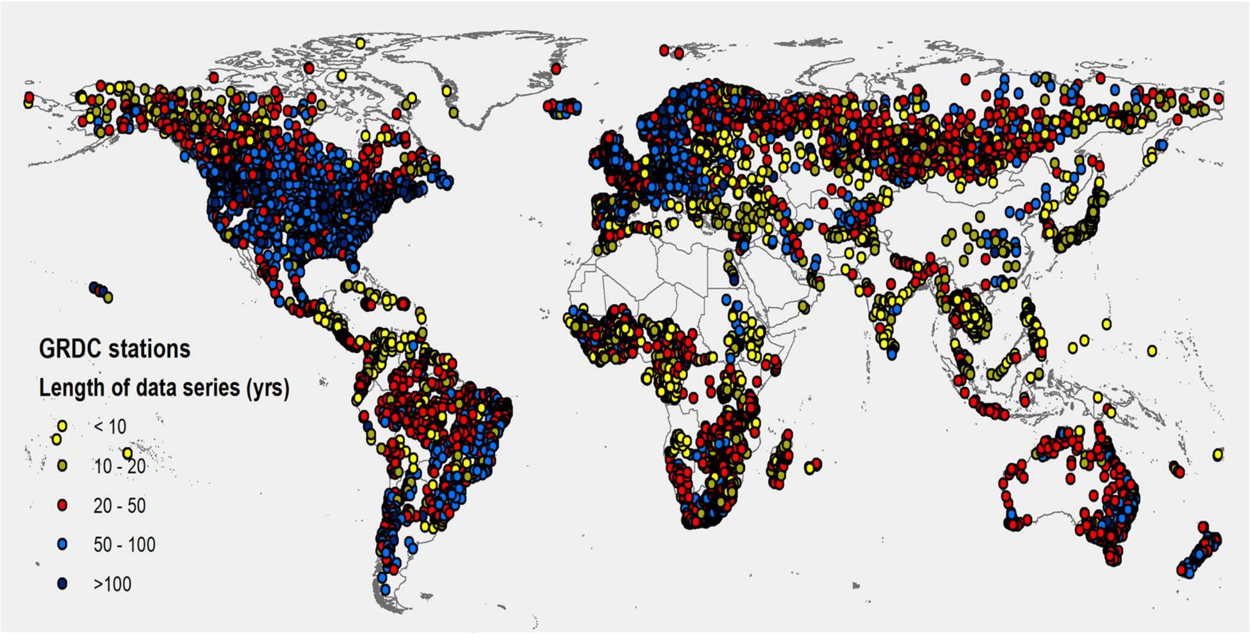


Figure 3.11 GRDC stations and length of river discharge data (data source: Global Runoff Data Centre)

When it comes to remote sensing data, it is becoming common practice to substitute remote sensing data for the in-situ data in various types of analyses such as using the remoted sensing data as input and calibration data in Earth system modeling (Stisen and Sandholt, 2010; Musie et al., 2019; Huang et al., 2020; Oliveira et al., 2021). However, in remote sensing data production it takes effort to convert the raw data collected by remote sensors to estimates of variables of interest. The process can be viewed as a modeling process subject to uncertainty and requires in-situ measurements for the estimation method development, calibration and validation (Niro et al., 2021). Thus, the observation made above on heterogeneity in uncertainty distribution in global modelled data is also applicable to the global remote sensing data products.

In addition to the spatial heterogeneity of the uncertainty distribution, there is also varied availability of data across variables. Still take river discharge data and water quality data as example, a collection of global modeling tools for hydrological/river discharge simulation and water quality simulation are listed in Table 3.11. As evident from the table, development in global water quality modeling area lags behind the advances in global hydrological/river discharge modeling area, and lack of in-situ water quality data for model evaluation is one of major barriers that lead to the slow progress in water quality modeling (Strokal et al., 2019).

Table 3.11 Selected global hydrological and water quality models

Hydrological/river discharge	Water quality
• Community Water Model (CWatM) (Burek et al., 2020) • GWAVA (Global Water Availability Assessment model) (Meigh et al., 1999) • H08 (Hanasaki et al., 2008) • ISBA-TRIP (Interactions between Soil, Biosphere, and Atmosphere – Total Runoff Integrating Pathways) (Oki and Sud, 1998; Boone et al., 1999) • Macro-PDM (Macro Probability Distribution Model) (Gosling and Arnell, 2011) • MPI-HM (Max Planck Institute – Hydrology Model) (Hagemann, S. and Dümenil, 1997) or HydroPy (Stacke and Hagemann, 2021) • PCR-GLOBWB (PCRaster GLOBal Water Balance model) (Sutanudjaja et al., 2018) • WASMOD-M (Water And Snow balance MODeling system) (Widén-Nilsson et al., 2007)	• Global NEWS (The Global Nutrient Export from Watersheds model) (Beusen et al., 2005) • WorldQual (UNEP, 2016)

- WaterGAP (Water – Global Analysis and Prognosis model) (Alcamo et al., 2003)
- WBM (Water Balance Model) (Fekete et al., 2002)

3.3.3 Addressing the data challenge in global ecosystem health assessment

The data uncertainty undoubtedly poses challenges for global ecosystem health assessment. In the long run, addressing the data challenge requires continued effort to enhance capacity of data acquisition, including developing new data acquisition technologies. The last decade has seen the emergence of such technologies, such as eDNA (Section 5.9.5.3). It is also important to establish platforms to facilitate data exchange and sharing (Michener, 2015; LaDeau et al., 2017).

In the short term, care needs to be taken to inspect the uncertainty in relevant data sets. The ecosystem health assessment typically involves weighting multiple environmental variables in a multi-criteria decision framework, and the scientific considerations behind the construction of the weighting scheme are discussed in Section 5.9.6.7. At the same time, from a perspective of data analysis the input data uncertainty also has implications for the weighting process. It would be desirable to introduce uncertainty/sensitivity analysis techniques (Ezbakhe and Perez-Foguet, 2018; Pelissari et al., 2021) to quantify the uncertainty in future endeavor of developing global ecosystem health indices, and this constitutes an interesting topic inviting future research.

3.4 Successful global indices

There are several indices that are applied at a global level that have relevance to development of the global RH framework, either just by providing an example of a successful index, or they are indices that could form part of the RH framework.

3.4.1.1 The IUCN Red List Index

The IUCN Red List (IUCN, 2021) is a comprehensive assessment of the extinction risk of the word's species. Species are categorised according to their extinction risk as Extinct, Extinct in the Wild, Critically Endangered, Endangered, Vulnerable, Near Threatened, Least Concern or Data Deficient. Assessments have been carried out for over 100,000 species. However, these have focused on only a handful of groups, namely birds, mammals, amphibians, cycads, and reef-forming corals, leaving a major knowledge gap concerning other groups, particularly plants and invertebrates. Therefore, the IUCN and the Red List Partnership are actively working to broaden the taxonomic coverage of the Red List Index (RLI), especially for marine and freshwater ecosystems and have set out an ambitious strategy to achieve this (IUCN, 2017). This includes representative sampling of some of the most species rich groups, with RLIs now available for freshwater crabs, crayfish, lobsters, fish, and reptiles and underway for butterflies, dung beetles, freshwater molluscs, gymnosperms and monocotyledon plants. However, since the RLI is based only on species extinction, it does not provide an indication of overall ecological state. It also suffers a lag effect, with extinction following the impacts, so is a poor management tool if the objective is to conserve or protect ecosystems. However, it does provide a sobering reminder of the cost of ecological degradation.

3.4.1.2 Biodiversity Intactness Index (BII)

The **Biodiversity Intactness Index** (BII) measures the loss of functional diversity as the change in population abundance because of human impacts across a wide range of taxa and functional groups at a biome or ecosystem level using pre-industrial era abundance as a reference point. The index typically ranges from 100% (abundances across all functional groups at preindustrial levels) to lower values that reflect the extent and degree of human modifications to populations of plants and animals. The boundary is preliminarily proposed at 90 % BII, although the zone of uncertainty from 30 – 90% reflects the large gap in our understanding of the links between biodiversity intactness and Earth-system functioning. The BII has presently only been applied to Southern Africa but observations are that decreases in BII adequately capture increasing levels of ecosystem degradation (defined as land where the land-cover type has not changed but there is a persistent loss of productivity). They also estimated the *mean species abundance of original species* (MSA) at 84 % globally as an approximation of aggregated human impacts on the terrestrial biosphere but have not yet disaggregated this by functional groups or considered aquatic ecosystems. They write that in the long-term, the concept of biome integrity – the functioning and persistence of individual biomes – offers a promising and robust approach (Newbold *et al.*, 2016, 2019; Martin, Green and Balmford, 2019), (Scholes and Biggs, 2005).

3.4.1.3 Living Planet Index (LPI)

The Living Planet Index (LPI) is a measure of the state of the world's biological diversity based on population trends of vertebrate species from terrestrial, freshwater and marine habitats (WWF, 2020b). It has been adopted by the Convention of Biological Diversity (CBD) as an indicator of progress towards its 2011-2020 target to 'take effective and urgent action to halt the loss of biodiversity'. It is based on the population trends of 4807 vertebrate species (mammals, birds, reptiles, amphibians, and fish) and 27695 populations from set monitoring sites around the world. This includes 3,741 monitored populations of 944 species in the Freshwater Living Planet Index (WWF, 2020a). The sources of data for the calculation of the index are wildlife population datasets gathered from almost 4,000 sources. The majority of these are publicly available and are found in scientific literature or in online repositories of wildlife census data. Unfortunately, data are missing for some species or places, reducing the robustness of the index. However, advances in wildlife monitoring, including eDNA, acoustic monitoring, camera traps and use of drones etc. are making monitoring of vertebrates easier, more widespread and more accurate, which will likely result in the LPI becoming more useful over time. It is designed at the continental/ ocean-basin scale with results given separately for each continent and for terrestrial, marine and freshwater biomes separately. They also provide an analysis of the main threats facing each region (as the percentage of populations affected), divided into habitat loss and degradation, species overexploitation, invasive species and disease, pollution, and climate change. This provides a better understanding of the drivers of biodiversity loss in each region.

The results of the LPI 2020 for freshwater biodiversity shows an average decline of 84% (range: -89% to -77%), equivalent to 4% per year since 1970 with 1 in 3 species threatened with extinction. Most of the declines are seen in amphibians, freshwater reptiles, and fishes; and are recorded across all regions, but particularly Latin America and the Caribbean. This decline in freshwater biodiversity is significantly higher than the average for all biomes. The decline is also greatest among ‘megafauna’, i.e., species with body mass > 30kg, including large fish, crocodilians, river dolphins, otters, beavers, and hippos, which are subject to severe anthropogenic threats, including over-exploitation of larger fish.

The main advantage of the LPI is that it is one of the few available macro-scale biodiversity indices showing the state of biological diversity globally. It is based on readily accessible datasets and the consideration of population changes (as a %) is also more accurate and representative of actual ecological conditions than only considering the conservation (threatened) status of species, such as in the Red List Index. Such an approach for monitoring is thus especially useful for larger scale assessments. Unfortunately, it is not easily down-scalable, as the smaller the area considered, the fewer data sources (i.e. species & populations) are relevant. Given the sporadic distribution of sites, areas with more sources of information, (N America, Europe, S & E Africa, India, Japan, Australia and New Zealand) may be able to create similar regionally applicable indices, whilst areas with fewer data sources (Central South America, Africa and Asia) may struggle in this regard. However, as mentioned above, the capacity of wildlife population monitoring is improving rapidly, improving the relevance of the index. But, even still, it is not suitable for small- to medium- spatial scales.

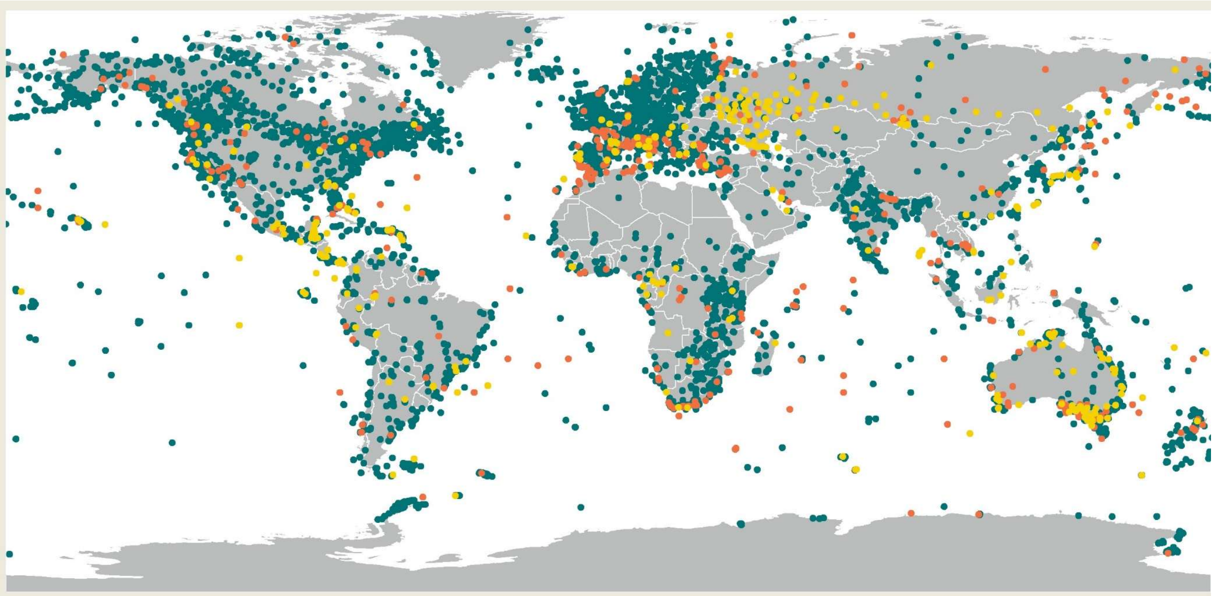


Figure 3.12: Map showing the location of species populations utilised for the Living Planet Index (LPI). Populations considered in the previous LPI are shown in green, whilst new species (yellow) or populations (orange) added since the last LPI are also shown.

3.4.1.4 Water Footprint

The Water Footprint Assessment is an index that quantifies the water use for any process (e.g. growing a particular crop), product (textile, car, food etc) or region at various scales (e.g. city, river basin, province, country, or globe) (Hoekstra *et al.*, 2011). This is measured as m^3 of water relative to the relevant unit (m^2, tonne etc). The footprint consists of three components: green, blue and grey water, providing a comprehensive picture of water use by delineating the source of water consumed and volume required for assimilation of pollutants***. Green water*** is defined as water from precipitation that is stored in the root zone of the soils and used by plants. This mainly involves agriculture and forestry. ***Blue water*** is water sourced from surface and groundwater resources. This mainly concerns domestic and industrial water use and irrigation. ***Grey water*** is the amount of water required to assimilate pollutants to meet water standards. This considers both point and diffuse sources of pollution. For an area such as a country, the water footprint is usually calculated from two perspectives: production and consumption. The first, measuring the water used for production of goods and the second, the amount involved in the consumption of product by people living in the area. This is useful at providing information on local pressures put on water resources in a region and helps us understand for what purposes our limited freshwater resources are being consumed and polluted. It also distinguishes whether the water consumed is sourced internally or externally to the region in question and recognises that water use can be externalised by importing water-intensive products, which can thus be referred to as trade in 'virtual water'. This is a particularly useful tool for regions with water shortages to reduce pressure on local water resources. The utility of this framework for a global assessment of RH is perhaps limited but it does provide a useful way of calculating water use as a pressure indicator.

3.4.1.5 Connectivity Status Index (CSI)

The CSI (Connectivity Status Index) (Grill *et al.*, 2019) quantifies the degree to which an individual river reach (i.e. the short river segment between two tributaries) remains connected to its neighbouring reaches within the larger river network. River connectivity is defined in four dimensions: longitudinal (river channel), lateral (floodplains), vertical (groundwater & atmosphere) and temporal (intermittency). They defined five 'pressure factors' impacting these dimensions, including river fragmentation (longitudinal), flow regulation (lateral and temporal), sediment trapping (longitudinal, lateral and vertical), water consumption (lateral, vertical and temporal) and infrastructure development in riparian & floodplain areas (longitudinal and lateral). These are represented by six proxies (two for infrastructure) of variables available in global datasets and numerical model outputs (Table 3.12)

Table 3.12 Pressure factors and Indicators used in the CSI (Grill et al., 2019)

Pressure factor	Pressure indicator	Description	Connectivity aspect affected
River fragmentation	DOF	Degree of fragmentation	Longitudinal
Flow regulation	DOR	Degree of regulation	Lateral, temporal
Sediment trapping	SED	Sediment trapping index	Longitudinal, lateral, vertical
Water consumption	USE	Consumptive use	Longitudinal, lateral, vertical, temporal
Infrastructure development in riparian and floodplain areas	URB	Nightlight intensity in urban areas	Lateral

These data are obtained from earth observation data, including the high resolution HydroSHEDS database (a global hydrographic map of river networks), which includes estimates of naturalised discharges, and datasets on the distribution of dams and reservoirs. In (Grill *et al.*, 2019) the distribution of dams was based on the Global Reservoir and Dam Database (GRanD) (Lehner *et al.*, 2011), which was limited to large reservoirs. However, future versions of the index (Grill and Lehner, no date) will make use of the Global Dam Watch (GDW) database. The GDW will combine data on the distribution of large dams from GRanD with that of medium-sized dams from the GlObal geOreferenced Database of Dams (GOODD) (Mulligan et al., 2020), which includes dams visible on global remote sensing imagery and that can be confirmed against existing GIS databases i.e. (Messager *et al.*, 2016; Pekel *et al.*, 2016). The CSI is thus calculated for individual river reaches as the weighted average of the pressure factors for every river reach and ranges from 0 – 100 % with 100 % indicating a 'free flowing' or fully connected reach.

The greatest benefit of this index is that it is based on global data and can be applied at variable scales, from a single reach to global. Its sensitivity to small-scale changes, including year on year, and ability to provide larger-scale information make this highly suitable as a global indicator. Compared to past connectivity indices it is also unique in being multidimensional as many of these previous methods evaluate impacts on only a single dimension. For longitudinally based indices, this makes them highly susceptible to 'singularities', where a single change (even minor) can lead to all downstream reaches being indicated as more impacted than they really are.

The shortfall of this index is that the data on the distribution of dams is biased to the impacts pf larger dams. Although this will be greatly improved by the GDW, smaller dams, especially on smaller systems may still be overlooked, thus overestimating the health of smaller rivers, as was found in Grill & Lehner (no date). There is also a lack of empirical evidence relating CSI to aquatic ecosystem health, clearly a subject in urgent need for research. The weighting of pressure indicators is also relatively arbitrary (not empirically based), whilst the selection of pressure factors may be incomplete. Indeed, (Grill *et al.*, 2019) mention that modifications to temperature, changes in hydroheic flux and pollution are also relevant to connectivity but were left out due to the lack of suitable datasets at a global scale.

In terms of RH, although connectivity is certainly a central determinant factor of RH it does not indicate impacts from sources such as pollution and direct biotic impacts (fishing). Structural (riverbed & riparian) changes that may impact RH, however, are included in its calculation. Therefore, although the CSI is highly suited for inclusion in a global RH index, it should be weighted against other indices of water quality and biotic health for a comprehensive representation of RH.

3.4.1.6 Canadian Water Quality Index (WQI)

Water quality is key to river health, thus the several indices available need to be considered.

The Water Quality Index (WQI) of the Canadian Council of Ministers of the Environment (CCME) provides a standard procedure to report water quality information by integrating various complex parameters of water quality into a single value. It is highly robust and has been used extensively to indicate surface and groundwater quality both in Canada (Davies, 2006; Environment Canada, 2011) and elsewhere around the world, including but not limited to New Zealand (Unwin and Larned, 2013; Henkel, 2017), India (Sharma and Kansal, 2011),Turkey (Bilgin, 2018), Iraq (Al-Janabi, Abdul-Rahman Al-Kubaisi and Al-Obaidy, 2012), and Greece (Alexakis, 2020). The same formulae are also used for the determination of water quality for the protection of human health by the Mekong River Commission. It operates by assessing water quality relative to its desired state as defined by water quality objectives (i.e., local guidelines) and incorporates three elements: 1) Scope – the number of parameters not meeting water quality objectives, 2) frequency – the number of times the objectives are not met, and 3) amplitude – the extent to which objectives are not met. The index is calculated using the formulae below on scale of 0 -100 (100 being the objective) and classified as 'excellent' (95 – 100), 'good' (80 – 94), 'fair' (65 – 79), 'marginal' (45 – 64) or 'poor' (0 – 44), accordingly. The specific parameters, as well as objectives and time periods used can thus vary between regions.

$$WQI = 100 - \frac{\sqrt{F_1^2 + F_2^2 + F_3^2 \text{ etc.}}}{1.732}$$

Scope: $$F_1 = \frac{Number\ of\ Failed\ Variables}{Total\ Numebr\ of\ Variables}\ X\ 100$$

Frequency $$F_2 = \frac{Number\ of\ Failed\ Tests}{Total\ Number\ of\ Tests}\ X\ 100$$

Amplitude $$F_3 = \frac{nse}{0.01nse \quad .01}$$

Where, *nse* is the "normalised sum of excursions" and *excursions* are the number of times an individual variables exceed a given objective.

Where the test value must (1) not exceed the objective or (2) fall below the objective:

(1) $excursion_i = \frac{Failed\ Test\ Value_i}{Objective_j} - 1$ OR (2) $excursion_i = \frac{Objective_j}{Failed\ Test\ Value_i}$ - 1

$$nse = \frac{\sum_{i=1}^{n} excursions_i}{\#\ of\ tests}\ X\ 100$$

The greatest advantage of the CWQI is that it is highly flexible in the variables included and specific objectives so can be used to compare the water quality of locations with very different characters. It is also open to citizen science engagement, facilitating widespread reporting. However, weaknesses include the requirement for repeated sampling to get accurate information for the variables involved. The equal weighting of scope, frequency and amplitude is also questioned, as the importance of each is likely not equal, so a weighting method may be helpful to improve the accuracy of the index. Also, the focus purely on water quality overlooks sources of impact not linked to water quality, which could be ameliorated by integrating other variables e.g., biological.

The major lesson learned for a global RH assessment is, again, the need for standardisation of methods for integrating data that is applicable across different contexts. Indeed, the formulae above provide a potential means of integrating various indicators of RH, not just water quality (e.g., variables of biotic health, hydrology, geomorphology etc) into a single value that is sensitive to differences between locations. This approach thus deserves consideration for the RH framework.

3.4.1.7 The global Water Quality Index (WATQI)

The Water Quality Index (WATQI) was the first global effort at reporting and estimating water quality and formed part of the 2008 and 2010 EPIs (Hsu *et al.*, 2016), see above. It is based on the five most commonly reported water quality parameters, namely dissolved oxygen, electrical conductivity, pH value, and total nitrogen and phosphorus concentrations (Srebotnjak *et al.*, 2012) as reported in the UN Global Environmental Monitoring System (UN GEMS), the only globally available database of national-level water quality parameters (*United Nations Global Environment Monitoring System.*, no date). For each parameter, the condition is calculated from the proximity-to-target of the measured value relative to the target for 'good water quality' set by international or local guidelines (see section 2.2.1.8 above) and the poorest performing country on a scale of 0 to 100. However, the UN GEMS is a self-reported database. Therefore, controversy arose around gaps and biases in the data as New Zealand ranked second in the word in terms of water quality, which experts suggested was of biased sampling with more polluted water bodies being overlooked in the data provided to the UN GEMS database. As such, the WATQI was dropped in subsequent versions of the EPI (Hsu *et al.*, 2016). Nevertheless, (Srebotnjak *et al.*, 2012), recommended hot-deck imputation as a method to deal with missing data and check for bias. Hereby, the (missing or questioned) WATQI value is estimated by matching the conditions of the country in question to others with similar natural and socioeconomic conditions.

Building onto this WATQI method is the SDG 6.3.2 methodology **Proportion of bodies of water with good ambient water quality** which draws extensively on the GEMS/Water programme and relies on it for much of its data.

The greatest benefits of the WATQI are that it is applicable at multiple scales, standardised for global applicability and integrative of several of the most important parameters of water quality. However, shortfalls include that it is subject to gaps and biases in the data, especially since the data are self-reported and lacking for many countries. The generalisation of the index means that it may also overlook impacts to water quality caused by other parameters, such as the presence of certain toxins.

Lessons for a global RH assessment given that water quality is a very important component of freshwater ecosystem health, include that it is possible to integrate several parameters of water quality into a single value but that the sampling sites used for *in situ* data acquisition of a region must be representative of that region and contain a high rick of bias. The proximity-to-target approach is also relative to the country with the poorest water conditions and in unsuitable for evaluating the absolute condition of aspects of ecosystem health. The hot-deck imputation carried out in (Srebotnjak *et al.*, 2012) may also provide a solution to dealing with missing data.

3.4.1.8 SDG 6.6.1 on change in extent of water-related ecosystems

A key global indicator that is part of the SDGs is 6.6.1 **Change in the extent of water-related ecosystems over time** and is available at the UN SDG Indicators Repository (no date). The authors of this indicator have taken the liberty of interpreting "extent" to mean quantity, and thus the indicator includes components of spatial extent and water quality (of lakes and reservoirs) for its Level 1 data-drive, with both indicators based on earth observation data. A second level that is country optional includes discharge, water quality imported from SDG 6.3.2 and groundwater quantity, all of which are collected *in situ*. They promote access to the method through an online portal https://www.sdg661.app/home with a "geospatial platform (that) allows you to explore data at national, sub-national and basin levels to better understand and quantify the state of freshwater ecosystems".

The target of SDG 6.6.1 is **By 2020 protect and restore water-related ecosystems, including mountains, forests, wetlands, rivers, aquifers and lakes.** A biologically orientated assessment of aquatic ecosystem health (or RH) would be ideal for this indicator, however they state in the method "*It is also essential to monitor how the health of water-related ecosystems is changing. However, health has not been included as a formal Sub-Indicator for Indicator 6.6.1 because monitoring ecosystem health is context specific, and the most appropriate methodology is based on local ecological conditions*". They instead suggest that the recommended sub-indicators are used in combination with locally derived biological indicators to inform the state of water-related ecosystem health, but no method for the integration of this data is offered. The first version of this method that was circulated in 2017 did contain a measure of locally derived biological data, with the requirement to normalise the data by reporting the deviation from reference conditions. This would have allowed a country to use a biological indicator of its choosing, provided the results could be normalised in this way. Unfortunately, this was rejected by UNEP based on uncertainty about the data that would come from such a heterogenous data collection.

This monitoring approach demonstrates the advantages of making use of global datasets that doe does not rely on country participation. However, it offers little to the RH framework because of the limited scope of sub-indicators that are used that would only be proxies of RH. However, a successfully developed RH framework in the future would find an automatic home in the 6.6.1 indicator.

3.4.1.9 Mekong River Commission WQI for the Protection of Aquatic Life

The MIF water quality index for the protection of aquatic life is calculated using the equation below (MRC, 2019a).

$$\mathrm{WQI} = \frac{\sum_{\mathrm{i}=1}^{\mathrm{n}} p_{\mathrm{i}}}{M} \; x \; 10$$

Where,

- "p_i" is the points scored on sample day i. For each parameter that meets its target one point is scored, otherwise the score is zero.
- "n" is the number of samples from the station in the year
- "M" is the maximum possible score for the measured parameters in the year

This index has been developed to allow the addition of more parameters should they become available.

3.4.1.10 AquaSat and future water quality remote sensing methods

AquaSat (Ross *et al.*, 2019) is the initial attempt to gather coincident data of satellite (Landsat 5, 7, and 8) reflectance values and *in situ* water quality measurements (Total Suspended Sediment (TSS), Dissolved Organic Carbon (DOC), Chlorophyll a, and Secchi Disk Depth (SDD)) from the US National Water Quality Portal (WQP) and Lake Multi-Scaled Geospatial and Temporal Database (LAGOS-NE) to build models to measure water quality remotely. Although the application of remote sensing for water quality assay is currently very limited, multiple current and future government and private satellite missions are aimed at providing continuous measurements with ideal spatial/ spectral resolutions ideal for inland water so developments in this field are expected to be rapid (Topp *et al.*, 2020). This is coupled with rapid developments in data processing methods, including empirical and semi-analytical models and as machine learning, to more accurately estimate parameters (Topp *et al.*, 2020).

The major benefits of methods to measure water quality remotely are that they are scale-independent (not limited to individual sites), overcome problems of missing data or sampling bias that are common with *in situ* measurements, and promise to be open access. A major shortfall, more so than with *in situ* measurements, continues to be the representativeness of the data gathered, especially in cases where the main stressors in a system (i.e., specific pollutants) are not detected by the satellites. The scale at which data are recorded is also restricted by the scale of data acquisition that is more suited to lakes and very large rivers and would likely exclude smaller rivers. Indications are that this may be possible down to 10m resolution.

Lessons for a global RH assessment include the enormous potential of remote-sensing data to provide continuous, scale-independent, and unbiased data.

4 GLOBAL WORKSHOP ON RIVER HEALTH

A workshop of selected specialists from around the world was held over two days (6th and 7th October 2021) with the purpose of presenting a status quo of river health monitoring, and then canvassing for ideas on the way forward.

Table 4.1 Agenda for the global workshop on river health 6th and 7th October 2021

Minutes	Content	Presenter	Org
5	Introduction and purpose of the meeting	Chris Dickens	IWMI
30	Review of the present state of knowledge	Jeremy Dickens	IWMI
10	Bending the curve of global freshwater biodiversity loss	David Tickner	WWF
15	SDG 6.6.1 water-related ecosystems	Stuart Crane	UNEP
5	Post 2020 CBD	Hazel Thornton	UNEP-WCMC
5	UN Decade of Ecological Restoration - FERM	Maria Nuutinen	FAO
5	Living Planet WWF	David Tickner	WWF
5	New Zealand Freshwater Biophysical Ecosystem Health Framework	Paul Franklin	NIWA New Zealand
5	eDNA	Mike Morris	Nature Metrics
10	Concepts for integration of data	Hua Xie	IFPRI
	TEA		
25	Discussion on above presentations		
60	Workshop on future possibilities 1. Indicators suitable for global reporting? 2. *In-situ* and/or remote sensing data for river health? 3. Holistic indices/dashboards etc. vs isolated "keystone" indicators? 4. Appropriate reporting at country/global scale – what kind of data is most useful?	Chris Dickens & Dave Tickner	IWMI WWF
	Future communications	Chris Dickens	IWMI
180	Closure		

4.1 Introduction

Healthy rivers are the veins of the living world. Ancient human civilizations, from those on the Euphrates and Tigris (6500 BC) to Malwathu Oya in Sri Lanka (300 BC), were dependent on rivers, signifying the pivotal role of rivers in sustaining human life. Sadly, most rivers in the world are now under stress from human developments, so that the benefits that they provided our ancestors are now under threat. Management of the health of rivers has thus become an essential part of sustainable development and is increasingly entrenched in national policy in many countries.

As human society grows and exerts an ever-increasing impact on natural resources including on rivers, the need to manage these resources at a global scale is becoming more important. The water in rivers by nature keeps moving downstream and crosses borders, and with the heavy pressures being exerted on them it makes sense to manage river health at a large scale, certainly at a basin scale, but also at a regional and global scale. Global sustainability programmes such as the Sustainable Development Goals and the Convention on Biological Diversity, as well as the UN Decade of Ecosystem Restoration, are all in need of data on river health at a regional and global level, but right now that does not exist.

In its original form, river health monitoring has generally been carried out *in-situ*, with assessments of habitat and biota carried out at river sites, which are then integrated to provide a picture of the health of an entire river basin or even a larger region. With the need to manage natural resources at scale, there is an urgent need to move beyond *in-situ* data to larger scale data for reporting at regional and global scale. But is this possible, and if so, what methods and indicators are most appropriate?

These were the questions discussed at the two recent workshops convened by IWMI, the CGIAR Research Program on Water Land & Ecosystems (WLE), and WWF, engaging key scientists and practitioners working on river health, across the world. In the words of Chris Dickens, the lead IWMI scientist on the project, the workshops aimed to "critically review the overall concept, provide insights and lessons from existing and emerging methods and chart a possible future for large-scale monitoring of river ecosystem health".

The workshops were conducted as two 3-hour sessions on Zoom catering to two different time zones. Attendees are shown in the Annexure below.

4.1.1 History of River Health Monitoring at the Regional and Global Scale

After the initial introduction by Chris Dickens, each workshop commenced with a review of existing regional and global scale river health monitoring frameworks/programs by Jeremy Dickens, University of Bonn ad IWMI Consultant. A brief summary of this review is presented below.

There are already some fledgling river health programmes at a regional scale, however, most of them are either dependent on *in-situ* data which are not directly amenable to upscaling, require intensive and costly measurement programs, or appear limited in their ability to comprehensively indicate river health. Examples of regional scale frameworks include the EU Water Framework Directive (WFD), USA EPA National Aquatic Resource Survey, South African River Eco-Classification and Eco-status Report, Chinese River and Lake Health Index, Australian Integrated Ecosystem Condition Assessment (IECA) Framework and the New Zealand Freshwater Biophysical Ecosystem Health Framework. While each one has its unique advantages and disadvantages these frameworks collectively point to key attributes a global framework should possess., i.e., standardized protocols; clear definition of river health; multiple indicators on the biology, water quality, hydrology, physical processes etc.; standardized methods of data aggregation, harmonization and integration, and scale independence.

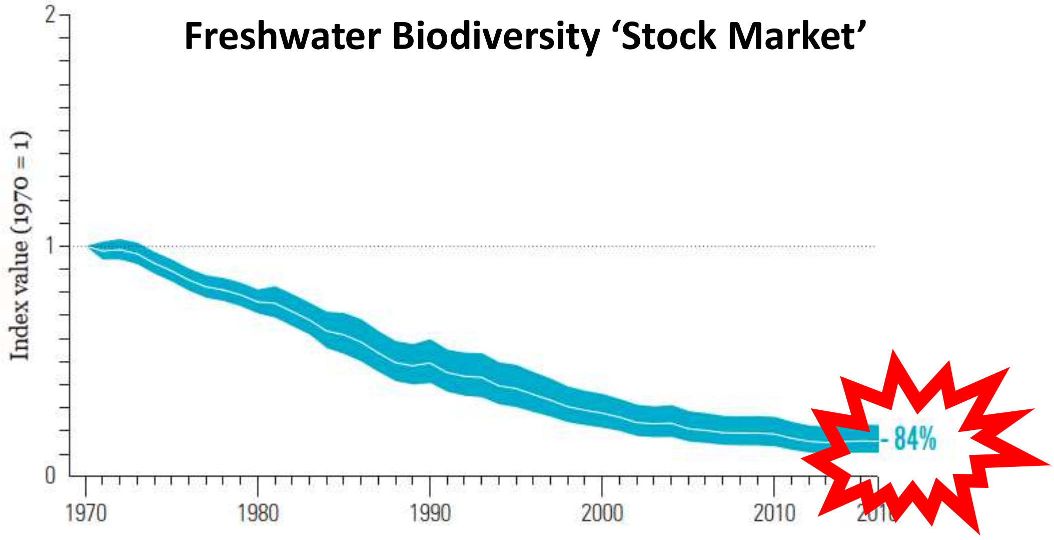
The demand for global scale data is apparent from the likes of the UN Sustainable Development Goal (SDG) indicator 6.6.1 (change in extent of water-related ecosystems over time) and the Convention on Biological Diversity (CBD) Post 2020 Global Biodiversity Framework. Some key indices and concepts like the River Connectivity Status Index, the Living Planet Index (LPI), Planetary Boundaries and the Incident Biodiversity Threat Index are all applicable at large scale although they only tangentially point to river health. The question remains, however, whether these frameworks and indices can provide a vehicle to holistically capture river health at the global scale?

The review of existing frameworks was followed by a set of presentations that elaborated on river health monitoring initiatives that are currently happening and how they relate to global scale monitoring of river health and future needs. The final presentation focused on aspects of global scale data and modelling in relation to river health. The same set of presentations were repeated in each session. The presentations in the order of appearance, and a brief summary of their content are given below.

4.2 List of Presentations

1. Review of Freshwater Health Assessment Frameworks – Jeremy Dickens, University of Bonn
2. Bending the Curve of Global Freshwater Biodiversity – David Tickner, WWF, UK
3. SDG Indicator 6.6.1 – Stuart Crane, UNEP (UNEP)
4. Indicators and the Post-2020 Global Biodiversity Framework - Hazel Thornton, Biodiversity Indicators Partnership, UNEP World Conservation Monitoring Centre (WCMC)
5. River Monitoring within the UN Decade on Ecosystem Restoration - Maria Nuutinen, Food and Agriculture Organization (FAO) of the UN
6. Living Planet Index (LPI) - David Tickner, WWF, UK
7. New Zealand Ecosystem Health Framework - Paul Franklin, National Institute of Water and Atmospheric Research, New Zealand
8. Mapping the World's Biodiversity using Environmental DNA (eDNA) - Mike Morris, NatureMetrics
9. Advances in Data and Modelling for Aquatic Ecosystem Health Assessment – Hua Xie, International Food Policy Research Institute (IFPRI)

According to David Tickner, the lead scientist on the project from WWF, the Living Planet Index (LPI) "measures average percentage change in size of discrete vertebrate populations and is in fact a stock market for wildlife." The Living Planet Report 2020 (https://livingplanet.panda.org/en-us/) shows that average abundance of 20,811 populations representing 4392 species monitored across the globe declined by 68% from its 1970 value and that freshwater biodiversity in particular, declined by 84% from its 1970 value (biodiversity is only one aspect of river health). This report triggered the high-level motivation to the CBD of an emergency recovery plan for freshwater biodiversity, which sought to bend back the declining freshwater biodiversity curve through interventions targeting conservation and drivers of biodiversity loss. Although very communicable and creating traction in policy circles, the LPI is only an abundance metric, but not a species richness, extinction threat or distribution metric. It likely suffers from sampling bias (geographic, habitat type, taxonomic and temporal), and may not be representative of actual trends across the globe. Therefore, to better track the implementation of the emergency recovery plan globally, there is an urgent need for robust, representative indicators of freshwater biodiversity at the global level.



Source: World Wildlife Fund and Zoological Society of London, Living Planet Report 2020 as appearing in the presentation by Dave Tickner, October 2021

The SDG indicator 6.6.1 measures changes in different types of freshwater ecosystems (lakes, rivers, wetlands and groundwater aquifers) over time and by country. The changes to freshwater ecosystems are measured in terms of spatial extent, water quality and water quantity using a mix of in country and earth observation data. These measurements can be tracked on the Freshwater Ecosystem Explorer (https://www.sdg661.app/) for each type of freshwater ecosystem. One of the main limitations in the ability of SDG indicator 6.6.1 to track the "health" of freshwater ecosystems is the non-inclusion of biological parameters. Future development of this indicator into a "traffic light" system to indicate the status of river basins worldwide (in terms of both ecosystems and vulnerability) would benefit from standardized measurements of river health, ideally including the state of the biology.

The Biodiversity Indicators Partnership (BIP) is a global initiative mandated by the CBD to coordinate the development and delivery of biodiversity indicators to support the Post 2020 Global Biodiversity Framework (https://www.bipindicators.net/). As part of this initiative, 70 plus global partners are working on 80 plus indicators to assess their suitability for the draft monitoring framework of the CBD. The UN Decade on Ecosystem Restoration is a similar initiative that has been tasked by the UN General Assembly to "prevent, halt and reverse the degradation of ecosystems worldwide". The Task Force on Monitoring of the UN Decade, consisting of 270 experts from 100 organizations, are responsible for developing a Framework for Ecosystem Restoration Monitoring (FERM), building on, and complementing, existing international, regional and national reporting processes, their goals, criteria and indicators. Draft river related indicators that are currently considered for monitoring and reporting include: SDG 6.3.2, 6.6.1, 15.1.2; CBD Post 2020 Biodiversity target 2.0.1, and Ramsar Convention on Wetlands targets 8.5, 8.6, 12.1 and 12.2. The (FERM) Geospatial platform (https://data.apps.fao.org/ferm/) is already available online. A framework for global river health assessment is likely to both support and build on important global initiatives like the BIP and the UN Decade that are currently taking shape.

4.3 The Way Forward

In each session, the presentations were followed by a discussion on the way forward on global river health monitoring, picking up from some of the presentations. The discussion was moderated by David Tickner. Given below is a brief summary of the discussion and potential next steps.

One of the key questions raised at the workshops was "What purpose would a global river health scorecard/index/dashboard serve and who would be the key audience for it?" The answers to this question ranged from "simple scientific inquiry, highlighting river health issues to (national and international) financial and political decision makers, to guiding international policy frameworks." In order to achieve these objectives and attract the intended audience, a global river health monitoring framework should deliver results that are scalable, consistent, robust, informative, representative, flexible, replicable, transparent, comprehensive, communicable and affordable, i.e., ideally the indicators, data, and methods of data aggregation, harmonization and integration should be vertically scalable from basin to regional and global levels (and vice versa) and horizontally transferable from region to region.

The New Zealand Freshwater Biophysical Ecosystem Health Framework (https://environment.govt.nz/assets/Publications/Files/freshwater-ecosystem-health-framework.pdf), developed in 2018, assesses nearly all key biophysical attributes of river health and is hailed as a good blueprint for a global framework. In this framework ecological Integrity is represented by the components water quantity, habitat, water quality, aquatic life and ecological processes. For each component key indicators are identified. For each indicator metrics are identified to measure and quantify the components using a mix of *in-situ* and earth observation data. Assessment progresses from indicator scores to component scores and finally to an ecological integrity score. As mandated by the National Policy Statement for Freshwater Management 2020, each regional council is required to publish data on each component annually, and an ecosystem health scorecard at least every five years. A process has also been formulated to show how a national scorecard can be created using existing data. The question is, how can robust biological data of sufficient quality be acquired at the global level to replicate such a framework globally? Could Environmental DNA (eDNA) be the answer?

eDNA essentially comprises traces of DNA that organisms leave in the environment. eDNA has revolutionized biological sampling by enabling the identification of species that exist in vast areas quickly and cheaply. Sampling may even be carried out by local communities and school children. The eBioAtlas (https://ebioatlas.org/) is an initiative that uses eDNA from water samples to map the world's freshwater biodiversity and lay the foundation for a global biodiversity framework that may feed into the IUCN red list assessment and the CBD. The initiative is hoping to analyze an initial 30,000 globally scattered samples, building an interactive database which can link with other databases like the IUCN red list. eBioAtlas is also building capacity to upscale *in-situ* point samples to whole river basins using eDNA transport and decay models. Model outputs can be used to: detect biodiversity hotspots that can be targeted for conservation; infer ecological requirements for different types of species, and to identify threats. However, eDNA is not without its limitations: currently, eDNA cannot measure true abundance or counts of individuals but can only indicate relative abundance; Lack of standardization in measurements creates difficulties in communicating between different data sets.

The real "breakthrough" for a global framework may come from a combination of earth observation data, eDNA, crowdsourcing/citizen science and mathematical models that have predictive power to fill gaps in *in-situ* data. Examples of predictive mathematical models include the eDITH model which combines earth observation data and eDNA; and the Biodiversity Intactness Index which predicts changes in biodiversity on terrestrial systems as an outcome of land use change and other pressures. Once the "high-level framework" is in place, appropriate indicators, weights and even targets, checks to avoid sampling bias, gap filling in monitoring networks, methods for the optimal use of existing data and the management and sharing of new data will need to be worked out. Lessons and best practices to support this exercise may be gained from the Arctic Freshwater Biodiversity Report: the output of a multinational freshwater biodiversity monitoring program across the 12 arctic council countries, that tracked the status of biodiversity change combining modelling with *in-situ* observations and linking to pressures and responses.

Another key topic that was discussed at length at the workshops was the importance of moving beyond biophysical indicators to establish links between humans and river health., i.e., should indicators on ecosystem services, socio-cultural elements and human induced threats and pressures be an integral part of global river health monitoring? While including them may make river health more accessible to the public, it may also dilute the ability to report on the ecological status. Other aspects that were highlighted include the need for validation of desktop indicators against true ecological integrity, and acknowledging that because of large data gaps some indicators may not paint a true picture of river health at the global level.

Feedback from the rich discussions indicated that the current project may potentially move forward on three integrated fronts around the questions: (a) Is it possible to formulate a dashboard or index of river health at the global scale and what would be the pathway to producing that dashboard or index? (b) Can this dashboard or index have separate modules on biophysical aspects, socio-cultural aspects and threats and pressures? (c) Where are the big gaps in monitoring networks that require interventions to get a better handle on river health?

IWMI together with the WWF are working to document the progress outlined above and hope to establish a framework that could be used by researchers across the globe to take this initiative forward. Feedback from specialists around the world would be welcome! Please write to Chris Dickens at c.dickens@cgiar.org or Nishadi Eriyagama at e.eriyagama@cgiar.org.

4.4 Annexure: List of attendees

Listed below are the attendees (excluding the project team) for each session:

<u>Americas/Canada meeting: October 06, 2021, GMT 14.30; California 06.30; New York 10.30</u>

1. Peter-John Meynell, International Centre for Environmental Management (ICEM)
2. Steven Cooke, Carlton University, Canada
3. Michael McClain, IHE Delft Institute for Water Education, the Netherlands
4. Jonathan Higgins, The Nature Conservancy (TNC)
5. Bernhard Lehner, McGill University, Canada
6. Rebecca Tharme, RiverFutures, UK
7. Derek Vollmer, Conservation International, USA
8. Ian Harrison, Conservation International, USA

9. Robin Abell, Conservation International, USA
10. Julian D. Olden, University of Washington, USA
11. Abigail Julia Lynch, United States Geological Survey (USGS)
12. Maria Nuutinen, Food and Agriculture Organization (FAO) of the UN
13. Linda Pistolesi, Centre for International Earth Science Information network (CIESIN)
14. Günther Grill, McGill University, Canada
15. LeRoy Poff, Colorado State University, USA
16. Hazel Thornton, Biodiversity Indicators Partnership, UNEP World Conservation Monitoring Centre (WCMC)
17. Martin Wolf, Yale University, USA
18. Chris McOwen, UNEP
19. Tara Moberg, The Nature Conservancy (TNC)
20. C. J. Kleynhans, Department of Water Affairs, South Africa (Retired)
21. MicheleThieme, World Wildlife Fund (WWF), USA.

<u>European/African/Eastern meeting: October 07, 2021, GMT 06:00; Colombo 11:30; Bangkok 13:00; Sydney 16:00</u>

1. Bruce Chessman, University of New South Wales, Australia
2. C. J. Kleynhans, Department of Water Affairs, South Africa (Retired)
3. Mike Morris, NatureMetrics
4. Paul Franklin, National Institute of Water and Atmospheric Research, New Zealand
5. Uthpala Pinto, New South Wales Department of Planning Industry and Environment, Australia
6. Christa Thirion, Department of Water and Sanitation, South Africa
7. Matthew McCartney, International Water management Institute (IWMI), Sri Lanka
8. Eren Turak, New South Wales Department of Planning Industry and Environment, Australia
9. Sonja Jahnig, Leibniz-Institute of Freshwater Ecology and Inland Fisheries (IGB), Germany
10. Maria Nuutinen, Food and Agriculture Organization (FAO) of the UN
11. Simon Linke, Griffith University, Australia
12. Ben Stuart-Koster, Australian Rivers Institute, Griffith University
13. Bonani Madikizela, Water Research Commission, South Africa
14. Conor Linstead, World Wildlife Fund (WWF), UK
15. Gordon O'Brien, University of Mpumalanga, South Africa
16. Kashif Shaad, Conservation International, USA
17. Mark Graham, Groundtruth, South Africa
18. Stuart Crane, UNEP
19. Kenneth Irvine, IHE Delft Institute for Water Education, the Netherlands
20. Thembi Masilela, Department of Water and Sanitation, South Africa
21. C. J. Kleynhans, Department of Water Affairs, South Africa (Retired)
22. MicheleThieme, World Wildlife Fund (WWF), USA.

5 A FRAMEWORK FOR RIVER HEALTH MONITORING AND REPORTING

Based on all the above information and taking into consideration the discussions held with experts from around the world during the workshop, the following essential components of a framework for monitoring will need to be outlaid. This important next step will be reported in 2022.

5.1.1 Key attributes

These are the key attributes that the framework should reflect, all of which are necessary to make the framework effective.

- **Consistency** - understanding of what constitutes ecosystem health and how to measure it
- **Representativeness** - includes measurement of a full range of the core components of ecosystem health
- **Robustness** - rigorous science with justified selection of components and indicator variables based on empirical evidence
- **Informativeness** - easily understood
- **Flexibility** - can be meaningfully applied across a wide range of waterbodies
- **Scalability** - application remains consistent across spatial scales
- **Feasibility** - not highly demanding on time, labour or money

5.1.2 Best approaches

Drawing from the many frameworks that are outlined above, this framework will select the most appropriate for inclusion in the design of a new framework.

5.1.3 Concepts for integration and reporting of data

Reporting the data is the essential and final requirement of any useful framework. Again, drawing on the experience reflected in the frameworks above, the best approach to integration and reporting of data will be used to define a pragmatic and appropriate approach for this new RH framework.

5.1.4 Possible framework

It is likely that the new framework will integrate global datasets and *in situ* data with a modelling approach.

6 CONCLUSIONS AND WAY FORWARD

This report has documented in some detail the *status quo* of RH and aquatic ecosystem health monitoring and reporting from around the world, with a focus on those existing frameworks that are suited in some way to upscaling and have potential to be part of a future global RH reporting framework. The purpose of this report was to establish the *status quo* and then to use this to recommend and then design a RH framework that would be appropriate for application at a global level. That remains to be done and will be the next step taken.

6.1 Lessons learned

Many lessons have been learned from the frameworks and indicators that have been evaluated in this report (see Table 6.1). Together these provide a valuable start to development of a global RH framework that would satisfy the needs for global sustainability reporting.

Table 6.1 Summary of the lessons learned from RH frameworks and indicators reviewed in this report.

Source	Lesson learned
WFD	The greatest lesson learned for a global RH assessment is that a standardised protocol is extremely useful for integrating data from various sources, especially where monitoring systems already exist. The WFD is highly flexible and robust and has attributes that would be highly beneficial to emulate. The standardised protocol is extremely useful if it is flexible and allows inclusion of different indicator variables.
NARS	A standardised protocol is essential, while standardisation of variables is key for widespread applicability. The probability-based sampling design is probably too rigid for applicability at multiple scales. NARS reports 'ecological conditions' relative only to biological conditions but does provide other indicators which have value in interpretation of biological conditions.
REMP	Ecological status can be determined by focusing on biotic response attributes, whilst ecological drivers, determined in parallel, are used for interpretation of changes.
RHI	Key is the standardised protocol for application across various contexts and spatial scales through use of key indicators. The inclusion of non-ecological factors, particularly governance, in the quantification of ecological health, is not suitable for global application.
NHRP	Use of a single metric to define RH limits its representativeness, suggesting that the integration of multiple indicators offers a more representative means of assessment. It also emphasizes the importance of appropriate reference site selection, as inappropriate and inconsistent selection can seriously bias the results. This is a framework that infers RH from biological conditions.
SRA	Usefulness of using modelled data for widespread scale-independent assessments, with useful perspectives on reference condition and how this could be influenced by climate change.
IEACA	Due to its hierarchical nature that builds on existing frameworks and allows for the selection of metrics/ indicators, it is highly flexible and can be applied to any aquatic ecosystem and for different management needs at multiple scales. It is also highly consistent enabling comparability across jurisdictions due to the clear definition of terms and setting of common themes and standard methods for assessment and reporting of conditions. Most useful is the detailed guidance on

	how to carry out the relevant groundwork, select appropriate indicators, and aggregate, harmonise, and integrate scores for reporting at varying scales.
FBEHF	This is beneficial to the purposes of this study due to its similar objective, to provide a widely applicable RH assessment framework. The key performance attributes, methods and core components of an integrated assessment approach and definition of ecological integrity are useful to a global RH assessment.
MIF	The framework demonstrates the utility of a hierarchical framework to provide varying levels of detail according to the objectives (decision making (strategic), assessment or monitoring) and the importance of properly integrating indicators within the given framework for representativeness.
SDG 6	The SDGs provide the potential to consider the inter-relatedness between ecosystem health, water quality, sanitation, water stress, and water governance, all of which are dependent on each other, however there is no integrated SDG 6 indicator.
CDB post-2020	The primary value of this framework is that it has similar objectives to this project and is an existing global effort. Although the emphasis is on biodiversity, by definition this would include the state of the ecosystem. The CBD process is ongoing and should be resolved in 2022 at which time this should provide a useful contribution. Many of its component indicators should be of value (although mostly already included here while many are also not ready for implementation).
GEO BON	A collection of global indicators integrating biodiversity observations, remote sensing data, and models to address important gaps in our understanding of biodiversity change. A useful concept however of limited value here as they focus on terrestrial ecosystems.
FHI	Truly integrative, including aspects of ecology, water use and governance that enables the discovery of the driving causes of poorly performing indicators, and thus highlights areas for improvement. It is a highly interactive system, suitable for IWRM monitoring and for testing future scenarios. In terms of a global RH assessment, it is advantageous in its adaptability to existing datasets, especially large-scale spatial data.
PB	Novel and useful approach to setting boundaries (similar to targets/ benchmarks/ standards) for ecosystems that includes application of the precautionary principle. The consideration of each indicator independently (instead of integration into a single index) is valuable, as each is fundamentally different by nature with its own thresholds, potential impacts and management or mitigation options.
ITI	It is possible to use drivers of ecological health to measure threats across spatial scales up to the global level. However, the lack of available data restricted the application of this to the biodiversity threat index, making it less representative. Nevertheless, it may even be possible to use the Incident Threat Indices themselves as driver indices, alongside other 'ecological state' indicators, in the calculation of an overall index of RH for application at the global level.
SEEA	Great flexibility as it can be used to assess the condition of virtually any ecological aspects in virtually any context, at any scale. However, the lack of identification of the core components of ecosystem health or RH means that it lacks consistency at the global level as the indicators chosen can be vastly different between countries, making comparison problematic. It is also highly dependent on data availability.
EPI	This is an example of a broad-scope dashboard indicator, that is little value for RH reporting as water resources are represented by mainly wastewater.

	Biodiversity is dominated by terrestrial indicators with no direct measure of freshwater ecosystems.
ERP	Clearly identifies the main issues threatening freshwater biodiversity, and hence ecosystems, and the actions that must be implemented to reverse the decline. Based on driver (not state) indicators that are vital to understand the links between human actions and ecological integrity, especially to inform management actions to 'bend the curve' of freshwater biodiversity loss.
RLI	The Red List provides a measure only of species extinction risk and an indication of overall ecological state. It suffers from being based on what is a very short list of species, with aquatic species only recently added, and thus does not represent the ecosystem as a whole. It also suffers a lag effect, with extinction following the impacts.
BII	Measures the loss of functional diversity as the change in population abundance due to human impacts using pre-industrial era abundance as a reference point. The concept of biome integrity – the functioning and persistence of individual biomes – offers a promising and robust approach. At present does not include aquatic ecosystems although this is achieved through GLOBIO but in a limited way in that it makes use of publication data only.
LPI	One of the few available macro-scale biodiversity indices showing the state of biological diversity globally using population changes (as a %) which is representative of actual ecological conditions. Unfortunately, it is not easily down-scalable, as the smaller the area considered, the fewer data sources (i.e. species & populations) are relevant. Sporadic distribution of sites means some areas better served than others (mostly underdeveloped).
WF	Limited value for RH reporting but it does provide a useful way of calculating water use as a pressure indicator.
CSI	Based on global data and can be applied at variable scales, from a single reach to global. Is sensitivity to small-scale changes, including year on year, and has the ability to provide larger-scale information. Although connectivity is a central determinant factor of RH it does not indicate impacts from sources such as pollution and direct biotic impacts (fishing), thus should be weighted against other indices of water quality and biotic health for a comprehensive representation of RH. This is highly suitable as a global indicator.
WQI (Canada)	The major lesson is the need for standardisation of methods for integrating data that is applicable across different contexts. The formulae provided are a potential means of integrating various indicators of RH, not just water quality (e.g., variables of biotic health, hydrology, geomorphology etc) into a single value that is sensitive to differences between locations. This approach thus deserves consideration for the RH framework.
WATQI	Applicable at multiple scales, standardised for global applicability and integrative of several of the most important parameters of water quality. It shows that it is possible to integrate several parameters of water quality into a single value. The proximity-to-target approach is also relative to the country and is unsuitable for evaluating the absolute condition of RH.
SDG 6.6.1	This is a perfect home for RH data but presently lacks a direct measure of RH. The present approach shows the value of using global dataset that do not rely on country participation. However, it offers little to the RH framework because of the limited scope of sub-indicators that are used that would only be proxies of RH.
Aquasat	Showing the major benefits of methods to measure water quality remotely as these are scale-independent (not limited to individual sites), overcome problems

	of missing data or sampling bias that are common with *in situ* measurements, and promise to be open access. Lessons for a global RH assessment include the enormous potential of remote-sensing data to provide continuous, scale-independent, and unbiased data.

6.2 Definition of RH

The key issue that needs to be resolved right at the outset of the design of a RH framework, is just what to include in the definition of river health (RH). It can be expected that different definitions will be appropriate for different uses.

Several of the frameworks reported have included human values in their definitions, which has the advantage of promoting the concept of RH into society. This has been done, at the first level, by including consideration of the ecosystem services produced by the river, healthy or otherwise. Secondly, the value of the river to society, the impacts on society, the role of governance etc., these are included in some of the frameworks. While there is merit in these approaches, they potentially dilute the urgent need to reflect the state of the very resource itself, which must be based on the biophysical character of the ecosystem alone. If there is no knowledge of the state of the resource itself, then understanding the role of this resource in society is meaningless. It is proposed that these are two different considerations, first the state of the river ecosystem itself, and secondly the relationship to society. For this RH framework, it is proposed that it is the ecosystem alone that is of relevance, and that this indicator can then be used to provide a second layer contribution to other indicators that include the human perspective. This approach is however open to debate and could change for a future framework, but it is an important issue that will need to be resolved before a final framework can be recommended or adopted.

Thus, for this report and as a recommendation to the RH framework to follow, where it is the health of the river ecosystem alone, as an ecological entity, that is being promoted, the following definition is proposed:

Recommended definition of River Health (Karr and Dudley, 1981; NRHP and SRA):

The definition of River Health is ***"the ability of the river ecosystem to support and maintain key ecological processes and a community of organisms with a species composition, diversity and functional organisation as comparable as possible to that of natural habitats within a region"***.

Application of this definition has implications for many of the frameworks that have been reviewed in this report, where the inclusion of the biological has been confined to a measure of biodiversity, especially if this is limited to endangered species. Meeting this definition would require that much more is known about the biota besides a measure of the biodiversity, and that the community structure and functions of the species are evaluated, including how this has deviated from a natural condition.

6.3 Key attributes of a RH Framework

This document summarises the work done to initiate a global framework for river health monitoring and reporting. Based on these existing initiatives, we have considered the key attributes of a successful framework.

The key attributes (Clapcott *et al.*, 2018) considered here include:

- **Consistency** - understanding of what constitutes ecosystem health and how to measure it
- **Representativeness** - includes measurement of a full range of the core components of ecosystem health
- **Robustness** - rigorous science with justified selection of components and indicator variables based on empirical evidence
- **Informativeness** - easily understood
- **Flexibility** - can be meaningfully applied across a wide range of waterbodies
- **Scalability** - application remains consistent across spatial scales
- **Feasibility** - not highly demanding on time, labour or money

Also considered were the key characteristics of existing successful frameworks as shown below. Many of these form part of the key attributes above, but here are framed from the way that they have been used:

- **Policy driven purpose** - A clear purpose is the foundational element of any framework as it influences decisions to all subsequent aspects of the framework development, from the definition of terms, choice of data acquisition methods, to processing and reporting.
- **Clear and consistent definitions** – the essential definition of RH considering all of those shown in Table 3.2 is adapted from that of the NRHP and SRA; "***The ability of the river ecosystem to support and maintain key ecological processes and a community of organisms with a species composition, diversity and functional organisation as comparable as possible to that of natural habitats within a region"***. However, any new program would need to be clear on which definition is being embraced.
- **Using conceptual models to direct the program** – with an overarching model such as the DPSIR, having a conceptual model helps to ensure that the program is fit for purpose and also that the results are not mis-represented.
- **Clear consideration of the key components** of a RH framework – there is clear separation of some frameworks that are based on ecological components and others that include socio-economic aspects. These would have different purposes.
- **RH indicator types** – some frameworks monitor the ecosystem directly while others make use of proxies where there is a risk that interpretation will be misleading.
- **Processing of data** - data processing usually involves three steps. 1) The aggregation of raw data to the appropriate scale for each metric. 2) Data are standardised to a common scale, to ensure consistency and flexibility. This often involves comparison to reference data. 3) The integration (or combination) of data at the indicator, component, or overall ecological condition levels for reporting.
- **Reporting of results** – reporting needs to consider the scale of the report, and also how to best communicate the data that is being used. Many formats for such reports have been produced, with perhaps the most useful being circulate pie-charts that provide an integrated score but also allow component scores to aid interpretation.

This report provides a baseline to formulate a global RH framework and will be followed up with further steps. Supported by IWMI and funded by the WWF, more conceptual thinking will be done to draft a paper that proposes a way forward for a global RH framework. It is hoped that this will provide a springboard for a renewed research effort to develop a global framework that can be adopted at a global level, feeding into global reports such as the SDGs, the CBD and many others.

7 REFERENCES

Alcamo, J., Döll, P., Henrichs, T., Kaspar, F., Lehner, B., Rösch, T. and Siebert, S., (2003) Development and testing of the WaterGAP 2 global model of water use and availability. Hydrological Sciences Journal, 48(3), pp.317-337.

Alexakis, D. E. (2020) 'Meta-Evaluation of water quality indices. Application into groundwater resources', Water (Switzerland), 12(7). doi: 10.3390/w12071890.

Al-Janabi, Z. Z., Abdul-Rahman Al-Kubaisi and Al-Obaidy, A.-H. M. J. (2012) 'Assessment of water quality of Tigris River by using Water Quality Index (WQI)', Journal of Al-Nahrain University, 15(1), pp. 119–126. doi: 10.31788/RJC.2020.1315344.

Álvarez-Cabria, M. et al. (2017) 'Modelling macroinvertebrate and fish biotic indices: From reaches to entire river networks', Science of the Total Environment. Elsevier B.V., 577(September 2017), pp. 308–318. doi: 10.1016/j.scitotenv.2016.10.186.

Arthington, A.H., Bhaduri, A., Bunn, S.E., Jackson, S.E., Tharme, R.E., Tickner, D., Young, B., Acreman, M., Baker, N., Capon, S., Horne, A.C., Kendy, E., McClain, M.E., Poff, N.L., Richter, B.D., Ward, S. (2018). The Brisbane declaration and global action agenda on environmental flows (2018). Frontiers in Environmental Science, 6. https://doi.org/10.3389/fenvs.2018.00045.

Baker, S. C. and Sharp, H. F. (1998) 'Evaluation of the recovery of a polluted urban stream using the ephemeroptera-plecoptera-trichoptera index', Journal of Freshwater Ecology, 13(2), pp. 229–234. doi: 10.1080/02705060.1998.9663611.

Beecher, H.A. Caldwell, B.A. DeMond, B. Seiler, D. Boessow, S.N. (2011) An Empirical Assessment of PHABSIM Using Long-Term Monitoring of Coho Salmon Smolt Production in Bingham Creek, Washington. North American Journal of Fisheries Management 30(6) https://doi.org/10.1577/M10-020.1

Bejarano, M.D., Sordo-Ward, A., Alonso, C., Nilsson, C. (2017). Characterizing effects of hydropower plants on sub-daily flow regimes. Journal of Hydrology, 550, pp. 186–200. https://doi.org/10.1016/j.jhydrol.2017.04.023.

Beusen, A.H.W., Dekkers, A.L.M., Bouwman, A.F., Ludwig, W. and Harrison, J., (2005) Estimation of global river transport of sediments and associated particulate C, N, and P. Global Biogeochemical Cycles, 19(4).

Beven, K. and Binley, A., (1992) The future of distributed models: model calibration and uncertainty prediction. Hydrological processes, 6(3), pp.279-298.

Bezerra, M. O. et al. (2021) 'Operationalizing Integrated Water Resource Management in Latin America: Insights from Application of the Freshwater Health Index', Environmental Management. Springer US. doi: 10.1007/s00267-021-01446-1.

Bilgin, A. (2018) 'Evaluation of surface water quality by using Canadian Council of Ministers of the Environment Water Quality Index (CCME WQI) method and discriminant analysis method: a case study Coruh River Basin', Environmental Monitoring and Assessment. Environmental Monitoring and Assessment, 190(9). doi: 10.1007/s10661-018-6927-5.

Bjerklie, D.M., Dingman, S.L., Vorosmarty, C.J., Bolster, C.H. and Congalton, R.G., (2003). Evaluating the potential for measuring river discharge from space. Journal of Hydrology, 278(1-4), pp.17-38.

Black, A.R., Rowan, J.S., Duck, R.W., Bragg, O.M., Clelland, B.E. (2005). DHRAM: a method for classifying river flow regime alterations for the EC Water Framework Directive. Aquatic Conservation: Marine and Freshwater Ecosystems, 15, pp. 427–446. https://doi.org/10.1002/aqc.707.

BMWP (1978) Final report: assessment and presentation of the biological quality of rivers in Great Britain. Unpublished report, Biological Monitoring Working Party, Department of the Environment.

Boone, A., Calvet, J.C. and Noilhan, J., (1999) Inclusion of a third soil layer in a land surface scheme using the force–restore method. Journal of Applied Meteorology, 38(11), pp.1611-1630.

Borja, A. et al. (2016) 'Overview of integrative assessment of marine systems: The ecosystem approach in practice', Frontiers in Marine Science, 3(MAR), pp. 1–20. doi: 10.3389/fmars.2016.00020.

Boulton, A.J. (1999) An overview of river health assessment: philosophies, practice, problems and prognosis Freshwater Biology 41(2) 469-479

Bovee, K.D. (1982) *A guide to stream habitat analysis using the instream flow incremental methodology*. Western Energy and Land Use Team, Office of Biological Services, Fish and Wildlife Service, US Department of the Interior

Bunn, S. Arthington, A. (2002) Basic principles and ecological consequences of altered flow regimes for aquatic biodiversity. Environ Manag 30:492–50.

Bunn, S. E. et al. (2010) 'Integration of science and monitoring of river ecosystem health to guide investments in catchment protection and rehabilitation', Freshwater Biology, 55(SUPPL. 1), pp. 223–240. doi: 10.1111/j.1365-2427.2009.02375.x.

Burek, P., Satoh, Y., Kahil, T., Tang, T., Greve, P., Smilovic, M., Guillaumot, L., Zhao, F. and Wada, Y., (2020) Development of the Community Water Model (CWatM v1. 04)–a high-resolution hydrological model for global and regional assessment of integrated water resources management. Geoscientific Model Development, 13(7), pp.3267-3298.

Burkhard, B., Kroll, F., Nedkov, S. and Müller, F. (2012) "Mapping ecosystem service supply, demand and budgets," *Ecological Indicators*, 21, pp. 17–29.

Burkhard, B. and Maes, J. (2017) "Mapping ecosystem services," *Advanced books*, 1, p. e12837.

Burkhard, B., Müller, F. and Lill, A. (2008) "Ecosystem Health Indicators," in *Ecological Indicators, vol. [2] of Encyclopedia of Ecology, 5 vols.*, pp. 1132–1138.

Burnett, M.J. O'Brien, G.C. Jacobs, F.J. Jewitt, G. Downs, C.T. (2021) Fish telemetry in African inland waters and its use in management: a review. Rev Fish Biol Fisheries https://doi.org/10.1007/s11160-021-09650-2

Camargo, J. A. (1993) 'Macrobenthic surveys as a valuable tool for assessing freshwater quality in the Iberian Peninsula', Environmental Monitoring and Assessment, 24(1), pp. 71–90. doi: 10.1007/BF00568800.

Carraro, L. Stauffer, J.B. Altermat, F. (2020) How to design optimal eDNA sampling strategies for biomonitoring in river networks. *Environmental DNA, DOI: 10.1002/edn3.137*

CBD Convention on Biological Diversity Post 2020 Biodiversity Framework (no date). Available at: https://www.cbd.int/post2020

CCME (2002) Canadian Water Quality Index. Available at: https://www.gov.nl.ca/ecc/waterres/quality/background/cwqi/.

CEC (2000) 'Directive 2000/60/EC of the European Parliament and of the Council of 23 October 2000 establishing a framework for Community action in the field of water policy', Official Journal of the European Communities, L327, p. 72.

Chapin, T.P., (2015) High-frequency, long-duration water sampling in acid mine drainage studies: A short review of current methods and recent advances in automated water samplers. Applied Geochemistry, 59, pp.118-124.

Chessman, B. C. (1995) 'Rapid assessment of rivers using macroinvertebrates: A procedure based on habitat-specific sampling , family level identification and a biotic index', Australian Journal of Ecology, 20, pp. 122–129.

Chessman, B. C. (2003) SIGNAL 2 – A scoring system for macro- invertebrates ('Water Bugs') in Australian rivers. Monitoring River Heath Initiative Technical Report 31. Canberra, Commonwealth of Australia.

Chessman, B. C. (2021) 'What's wrong with the Australian River Assessment System (AUSRIVAS)?', Marine and Freshwater Research, pp. 1110–1117. Available at: https://doi.org/10.1071/MF20361.

Chutter, F. M. (1972) 'An empirical biotic index of the quality of water in South African streams and rivers', Water Research, 6(1964), pp. 19–30.

Clapcott, J. E. et al. (2018) 'Freshwater biophysical ecosystem health framework', Cawthron Report No. 3194, (3194), p. 89. Available at: https://www.mfe.govt.nz/sites/default/files/media/Fresh water/freshwater-ecosystem-health-framework.pdf.

Costanza, R. *et al.* (1997) "The value of the world's ecosystem services and natural capital," *Nature*, pp. 253–260.

Cox, B., Oeding, S. and Taffs, K. (2019) "A comparison of macroinvertebrate-based indices for biological assessment of river health: A case example from the sub-tropical Richmond River Catchment in northeast New South Wales, Australia," *Ecological Indicators*, 106, p. 105479.

Dafforn, K.A., Johnston, E.L., Ferguson, A., Humphrey, C.L., Monk, W., Nichols, S.J., Simpson, S.L., Tulbure, M.G. and Baird, D.J., (2015) Big data opportunities and challenges for assessing multiple stressors across scales in aquatic ecosystems. Marine and Freshwater Research, 67(4), pp.393-413.

Dallas, H. Dickens, C. Hill, L. Kleynhans, N. Louw, D. Taylor, T. Thirion, C. (2008) National Aquatic Ecosystem Health Monitoring Programme (NAEHMP): River Health Programme (RHP) Implementation Manual. Department of Water Affairs and Forestry, Report N/0000/00/REQ/0308, Pretoria, South Africa, 130pp ISBN No. 978-0-621-383343-0.

Dallas, H. F. et al. (2018) 'Zambian Invertebrate Scoring System (ZISS): A macroinvertebrate-based biotic index for rapid bioassessment of southern tropical African river systems', African Journal of Aquatic Science, 43(4), pp. 325–344. doi: 10.2989/16085914.2018.1517081.

Davies, J. M. (2006) ‘Application and tests of the Canadian water quality index for assessing changes in water quality in lakes and rivers of central North America’, Lake and Reservoir Management, 22(4), pp. 308–320. doi: 10.1080/07438140609354365.

Davies, P. E. (2000) ‘Development of a national river bioassessment system (AUSRIVAS) in Australia’, in Assessing the biological quality of freshwaters: RIVPACS and other techniques. Cumbria, UK: Freshwater Biological Association, pp. 113–124.

Davies, P. E. et al. (2010) ‘The sustainable rivers audit: Assessing river ecosystem health in the murraydarling basin, Australia’, Marine and Freshwater Research, 61(7), pp. 764–777. doi: 10.1071/MF09043.

de Graaf, I.E., van Beek, R.L., Gleeson, T., Moosdorf, N., Schmitz, O., Sutanudjaja, E.H. and Bierkens, M.F., (2017) A global-scale two-layer transient groundwater model: Development and application to groundwater depletion. Advances in water Resources, 102, pp.53-67.

Department of the Environment and Energy (2017) ‘Module 5: Integrated ecosystem condition assessment’, in Aquatic Ecosystems Toolkit. Canberra, Commonwealth of Australia: Australian Government Department of the Environment and Energy, p. 123 pp. Available at: https://www.environment.gov.au/system/files/resources/ed4f9ae3-599d-46b6-b73a-aa15188fbd9c/files/ae-toolkit-module-5-integrated-ecosystem-condition-assessment.pdf.

Department of Water Affairs (2013) ‘Government Notice No. 665: Revision of General Authorisations in Terms of Section 39 of the National Water Act, 1998 (Act No. 36 of 1998) (The Act)’, Government Gazette, (36820), pp. 3–31.

Department of Water Affairs and Forestry (2008) Methods for determining the Water Quality component of the Ecological Reserve. Second Draft. Pretoria, South Africa.

Department of Water and Sanitation (2014) A Desktop Assessment of the Present Ecological State, Ecological Importance and Ecological Sensitivity per Sub Quaternary Reaches for Secondary Catchments in South Africa. Secondary: [W5 (for example)]. Compiled by RQIS-RDM: https://www.dwa.gov.za/iwqs/rhp/eco/peseismodel.aspx accessed on 3/12/2021.

Development Research Centre (2019) Technical outline for comprehensive implementation of the assessment of the River (Lake) chief system. Beijing, China: Report Completed in Collaboration with Hohai University and North China University of Water Resources and Electric Power; Ministry of Water Resources.

Dickens, C. and Graham, P. M. (2002) ‘The South African Scoring System (SASS) Version 5 Rapid Bioassessment Method for Rivers’, African Journal of Aquatic Science, 27(1), pp. 1–10. doi: 10.2989/16085914.2002.9626569.

Dickens, C. and McCartney, M. (2021) Water-Related Ecosystems. W. Leal Filho et al. (eds.), Clean Water and Sanitation, Encyclopedia of the UN Sustainable Development Goals, https://doi.org/10.1007/978-3-319-70061-8_100-1.

Döll, P., Kaspar, F. and Lehner, B. (2003) ‘A global hydrological model for deriving water availability indicators: Model tuning and validation’, Journal of Hydrology, 270(1–2), pp. 105–134. Available at: http://dx.doi.org/10.1016/S0022-1694(02)00283-4.

Duan, Q., Sorooshian, S. and Gupta, V.K., (1994) Optimal use of the SCE-UA global optimization method for calibrating watershed models. Journal of hydrology, 158(3-4), pp.265-284.

Duffy, G. and Regan, F., (2017) Recent developments in sensing methods for eutrophying nutrients with a focus on automation for environmental applications. Analyst, 142(23), pp.4355-4372.

Dynesius, M., Nilsson, C. 1(994). Fragmentation and flow regulation of river systems in the northern third of the world. Science, 266, pp. 753–762. https://doi.org/10.1126/science.266.5186.753

Environment Canada (2011) Water quality status and trends of nutrients in major drainage areas of Canada - technical Summary. Ottawa: Water Science and Technology Directorate.

eWater, DSEWPC and LWA (no date) AUSRIVAS Australian River Assessment System. Available at: https://ausrivas.ewater.org.au.

Ezbakhe, F. and Perez-Foguet, A., (2018) Multi-criteria decision analysis under uncertainty: two approaches to incorporating data uncertainty into water, sanitation and hygiene planning. Water resources management, 32(15), pp.5169-5182

Fairweather, P.G. (1999) State of environment indicators of 'river health': exploring the metaphor Freshwater Biology 41(2)

Falkenmark, M. (1997) "Meeting water requirements of an expanding world population," *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 352(1356), pp. 929–936.

FBM and DNRM (2001) Australia-Wide Assessment of River Health: Queensland AusRivAS Sampling and Processing Manual, Monitoring River Heath Initiative Technical Report no 12. Canberra: Environment Australia. Available at: http://155.187.2.69/water/publications/environmental/rivers/nrhp/manual-vic/pubs/manual-vic.pdf%0Ahttps://catalogue.nla.gov.au/Record/3095551.

Fekete, B.M., Vörösmarty, C.J. and Grabs, W., (2002) High-resolution fields of global runoff combining observed river discharge and simulated water balances. Global Biogeochemical Cycles, 16(3), pp.15-1.

Feio, M. J. et al. (2021) 'The biological assessment and rehabilitation of theworld's rivers: An overview', Water (Switzerland), 13(3), pp. 1–46. doi: 10.3390/w13030371.

Gao, Y. *et al.* (2009) "Development of representative indicators of hydrologic alteration," *Journal of Hydrology*, 374(1–2), pp. 136–147.

GEO BON (2015) *GEO BON - The Group on Earth Observations Biodiversity Observation Network*. Available at: https://geobon.org/.

Gholizadeh, M.H., Melesse, A.M. and Reddi, L., (2016) A comprehensive review on water quality parameters estimation using remote sensing techniques. Sensors, 16(8), p.1298.

Gippel, C.J., Stewardson, M.J. (1998) Use of wetted perimeter in defining minimum environmental flows. Regulated Rivers: Research and Management, 14, 53–67.

Gippel, C.J., Marsh, N., Grice, T. (2012). Flow Health - Software to assess the deviation of river flows from reference and to design a monthly environmental flow regime. In: Technical Manual and User Guide, Version 2.0. ACEDP Australia-China Environment Development Partnership, River Health and Environmental Flow in China. International WaterCentre, Brisbane, Fluvial Systems Pty Ltd, Stockton, and Yorb Pty Ltd, Brisbane.

Gosling, S.N. and Arnell, N.W., (2011) Simulating current global river runoff with a global hydrological model: model revisions, validation, and sensitivity analysis. Hydrological Processes, 25(7), pp.1129-1145.

Grill, G., Ouellet, D.C., Fluet, C.E., Sindorf, N., Lehner, B. (2014). Development of new indicators to evaluate river fragmentation and flow regulation at large scales: a case study for the Mekong River Basin. Ecological Indicators, 45, pp. 148–159. https://doi.org/10.1016/j.ecolind.2014.03.026.

Grill, G. et al. (2019) 'Mapping the world's free-flowing rivers', Nature. Springer US, 569(7755), pp. 215–221. doi: 10.1038/s41586-019-1111-9.

Grill, G. and Lehner, B. (no date) 'DRAFT. A standardized index to monitor trends in global and regional river connectivity. Version 5.' Confluvio Consulting Inc.

Gupta, H.V., Beven, K.J. and Wagener, T., (2006) Model calibration and uncertainty estimation. Encyclopaedia of hydrological sciences.

Hagemann, S. and Dümenil, L., (1997) A parametrization of the lateral waterflow for the global scale. Climate dynamics, 14(1), pp.17-31.

Halse, S. et al. (2002) Australia-Wide Assessment of River Health: Western Australian AusRivAS Sampling and Processing Manual, Monitoring River Heath Initiative Technical Report no 18. Canberra.

Hammond, A.L. (1995) *Environmental indicators: a systematic approach to measuring and reporting on environmental policy performance in the context of sustainable development*. World Resources Institute, Washington, DC.

Hanasaki, N., Kanae, S., Oki, T., Masuda, K., Motoya, K., Shirakawa, N., Shen, Y. and Tanaka, K., (2008) An integrated model for the assessment of global water resources–Part 1: Model description and input meteorological forcing. Hydrology and Earth System Sciences, 12(4), pp.1007-1025.

Hardie, S. A., Bobbi, C. J. and Uytendaal, A. R. (2018) 'Evaluation and redirection of a long-term, broad-scale river health monitoring program in Tasmania, Australia', in Proceedings of the 9th Australian Stream Management Conference. Hobart, Tasmania, pp. 317–324.

Haskell, B.D., Norton, B.G. and Costanza, R. (1992) "What is ecosystem health and why should we worry about it," *Ecosystem health: New goals for environmental management*, pp. 3–19.

Hatfield, T. Bruce, J. (2000) Predicting Salmonid Habitat–Flow Relationships for Streams from Western North America. North American Journal of Fisheries Management Volume 20 (4).

Henkel, S. (2017) 'Caseys Creek – Baseline Water and Sediment Quality', (July).

Hering, D. et al. (2010) 'The European Water Framework Directive at the age of 10: A critical review of the achievements with recommendations for the future', Science of the Total Environment. Elsevier B.V., 408(19), pp. 4007–4019. doi: 10.1016/j.scitotenv.2010.05.031.

Hilsenhoff, W. L. (1987) 'An improved biotic index of organic stream pollution', Great Lakes Entomologist, 20(1), pp. 31–40.

Hilsenhoff, W. L. et al. (1988) 'Rapid field assessment of organic pollution with a family-level biotic index', Journal of the North American Benthological Society, 7(1), pp. 65–68.

Hipel, K. W., and A. I. McLeod, (2005) Time Series Modeling of Water Resources and Environmental Systems. Developments in Water Science, Vol. 45, Elsevier, 1013 pp.

Hoekstra, A. Y. et al. (2011) The Water Footprint Assessment Manual. Setting the Global Standard, Social and Environmental Accountability Journal. doi: 10.1080/0969160x.2011.593864.

Hsu, A. et al. (2016) 2016 Environmental Performance Index. New Haven, CT: Yale University. Available at: www.epi.yale.edu.

Hsu, A., Johnson, L. and Lloyd, A. (2013) Measuring Progress: A Practical Guide From the Developers of the Environmental Performance Index (EPI). New Haven, CT: Yale Center for Environment & Policy. Available at: http://epi.yale.edu/content/measuring-progress-practical-guide-developers-environmental-performance-index-epi.

Huang, Q., Qin, G., Zhang, Y., Tang, Q., Liu, C., Xia, J., Chiew, F.H. and Post, D., (2020) Using remote sensing data-based hydrological model calibrations for predicting runoff in ungauged or poorly gauged catchments. Water Resources Research, 56(8), p.e2020WR028205.

ISPRA. (2011). IDRAIM – Stream Hydromorphological Evaluation, Analysis and Monitoring System Guidebook for the Evaluation of Stream Morphological Conditions by the Morphological Quality Index (IQM). (Rome, Italy).

IUCN (2017) The IUCN Red List of Threatened Species Strategic Plan 2017-2020. Prepared by the IUCN Red List Committee.

IUCN (2021) 'The IUCN Red List Index of Threatened Species. Version 2021-2.' Available at: https://www.iucnredlist.org.

IPBES (2018) Summary for policymakers of the regional assessment report on biodiversity and ecosystem services for Africa of the Intergovernmental Sciencepolicy Platform on Biodiversity and Ecosystem Services. In: Archer E, Dziba LE, Mulongoy KJ, Maoela MA, Walters M, Biggs R, Cormier-Salem M-C, DeClerck F, Diaw MC, Dunham AE, Failler P, Gordon C, Harhash KA, Kasisi R, Kizito F, Nyingi WD, Oguge N, Osman- Elasha B, Stringer LC, Tito de Morais L, Assogbadjo A, Egoh BN, Halmy MW, Heubach K, Mensah A, Pereira L, Sitas N (eds). IPBES secretariat, Bonn, 49 p

Janse, J.H. Kuiper, J.J. Weijters, M.J. Westerbeek, E.P. Jeuken, M.H.J.L. Bakkenes, M. Alkemade, R. Mooij, W.M. Verhoeven, J.T.A. (2015) GLOBIO-Aquatic, a global model of human impact on the biodiversity of inland aquatic ecosystems, Environmental Science & Policy, Volume 48. Pages 99-114, ISSN 1462-9011, https://doi.org/10.1016/j.envsci.2014.12.007

Joubert, A. (2004) The ecostatus project and multicriteria decision analysis: Report and background. Cape Town.

Jumani, S., Deitch, M.J., Kaplan, D., Anderson, E.P., Krishnaswamy, J., Lecours, V., Whiles, M.R. (2020). River fragmentation and flow alteration metrics: a review of methods and directions for future research. Environmental Research Letters, 15, pp. 123009. https://doi.org/10.1088/1748-9326/ABCB37

Kaaya, L. T., Day, J. A. and Dallas, H. F. (2015) 'Tanzania River Scoring System (TARISS): a macroinvertebrate-based biotic index for rapid bioassessment of rivers', African Journal of Aquatic Science, 40(2), pp. 109–117. doi: 10.2989/16085914.2015.1051941.

Kansakar, P. and Hossain, F., (2016) A review of applications of satellite earth observation data for global societal benefit and stewardship of planet earth. Space Policy, 36, pp.46-54.

Karr J.R. (1996) Ecological integrity and ecological health are not the same. Engineering Within Ecological Constraints (ed. P.C. Schulze), pp. 97±109. National Academy Press, Washington, DC.

Karr, J.R. (1999) “Defining and measuring river health,” *Freshwater Biology*, 41, pp. 221–234.

Karr, J.R. and Dudley, D.R. (1981) Ecological Perspective on Water Quality Goals. Environmental Management, Vol. 5, No. 1, pp. 55-8.

Karthick, B., Taylor, J.C., Mahesh, M.K., Ramachandra T.V. (2010) Protocols for collection, preservation and enumeration of diatoms from aquatic habitats for water quality monitoring in India. Journal of Soil and Water Sciences. 3 (1): 25-60

Kawanisi, K., Al Sawaf, M.B. and Danial, M.M., (2018) Automated real-time streamflow acquisition in a mountainous river using acoustic tomography. Journal of Hydrologic Engineering, 23(2), p.04017059.

Kebede, M.G., Wang, L., Li, X. and Hu, Z., (2020) Remote sensing-based river discharge estimation for a small river flowing over the high mountain regions of the Tibetan Plateau. International Journal of Remote Sensing, 41(9), pp.3322-3345.

Kennedy, Marc C., and Anthony O'Haga (2001) "Bayesian calibration of computer models." Journal of the Royal Statistical Society: Series B (Statistical Methodology) 63, no. 3: 425-464.

King, J.M. Tharme, R.E. (1994) Assessment of The Instream Flow Incremental Methodology, And Initial Development Of Alternative Instream Flow Methodologies For South Africa WRC Report No 295/1/94 ISBN No 1 86845 056 2.

Kleynhans, C.J. (1996) A qualitative procedure for the assessment of the habitat integrity status of the Luvuvhu River (Limpopo system, South Africa). Journal of Aquatic Ecosystem Health 5: 41-54.

Kleynhans, C. J. (2008) ‘Module D: Fish Response Assessment Index (FRAI)’, in River EcoClassification: Manual for EcoStatus Determination (version 2). Pretoria, South Africa: Joint Water Research Commission and Department of Water Affairs and Forestry report. WRC Report no TT 330/08.

Kleynhans, C. J. and Louw, M. D. (2008) ‘Module A: EcoClassification and EcoStatus determination’, in River EcoClassification: Manual for EcoStatus Determination (version 2). Pretoria, South Africa: Joint Water Research Commission and Department of Water Affairs and Forestry report. WRC Report No. TT 329/08.

Kleynhans, C. J., Louw, M. D. and Graham, P. M. (2009a) ‘Module G: Index of Habitat Integrity Section 1: Technical Manual.’, in River EcoClassification: Manual for EcoStatus Determination (version 2). Pretoria, South Africa: Joint Water Research Commission and Department of Water Affairs and Forestry report. WRC Report no TT 337/09.

Kleynhans, C. J., Louw, M. D. and Graham, P. M. (2009b) ‘Module G: Index of Habitat Integrity Section 2: Model Photo Guide.’, in River EcoClassification: Manual for EcoStatus Determination (version 2). Pretoria, South Africa: Joint Water Research Commission and Department of Water Affairs and Forestry report. WRC Report no TT 338/09.

Kleynhans, C. J., Mackenzie, J. and Louw, D. A. (2008) ‘Module F: Riparian Vegetation Response Assessment Index (VEGRAI)’, in River EcoClassification: Manual for EcoStatus Determination (Version 2). Pretoria.

Kibler, K.M. (2017). The Hydraulic Size of a Dam: A Metric Indicating Potential Hydrologic Impact for Storage and Diversion. Faculty Scholarship and Creative Works, pp. 622. http://stars.library.ucf.edu/ucfscholar/622.

LaDeau, S.L., Han, B.A., Rosi-Marshall, E.J. and Weathers, K.C., (2017) The next decade of big data in ecosystem science. Ecosystems, 20(2), pp.274-283.

LeGrand, M.C., Luce, J.J., Metcalfe, R.A. and Buttle, J.M., (2020) Development of an inexpensive automated streamflow monitoring system. Hydrological Processes, 34(13), pp.3021-3023.

Lehner, B., Reidy Liermann, C., Revenga, C., Vorosmarty, C., Fekete, B., Crouzet, P., Doll, P., Endejan, M., Frenken, K., Magome, J., Nilsson, C., Robertson, J., Rodel, R., Sindorf, N., Wisser, D. (2011a). High resolution mapping of the world's reservoirs and dams for sustainable river-flow management. Frontiers in Ecology and the Environment, 9, pp. 494–502. https://doi.org/10.1890/100125.

Lehner, B. et al. (2011b) 'Global Reservoir and Dam Database, Version 1 (GRanDv1): Dams, Revision 01'. Palisades, NY: NASA Socioeconomic Data and Applications Center (SEDAC).

Lehner, B. and 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), pp. 2171–2186. doi: 10.1002/hyp.9740.

Li, L., Zheng, B. and Liu, L. (2010) 'Biomonitoring and bioindicators used for river ecosystems: Definitions, approaches and trends', Procedia Environmental Sciences, 2, pp. 1510–1524. doi: 10.1016/j.proenv.2010.10.164.

Lin, Y.P., Wang, C.L., Yu, H.H., Huang, C.W., Wang, Y.C., Chen, Y.W. and Wu, W.Y., (2011) Monitoring and estimating the flow conditions and fish presence probability under various flow conditions at reach scale using genetic algorithms and kriging methods. Ecological modelling, 222(3), pp.762-775.

Lundqvist, J, (1998) Avert looming hydrocide. Ambio 27: 428–433

Martin, P. A., Green, R. E. and Balmford, A. (2019) 'The biodiversity intactness index may underestimate losses', Nature Ecology and Evolution. Springer US, 3(6), pp. 862–863. doi: 10.1038/s41559-019-0895-1.

Mathews, R., Richter, B.D., (2007) Application of the indicators of hydrologic alteration software in environmental flow setting. J. Am. Water Resour. Assoc. https://doi.org/ 10.1111/j.1752-1688.2007.00099.x

Mayer-Schönberger, V. and Cukier, K., (2013) Big data: A revolution that will transform how we live, work, and think. Houghton Mifflin Harcourt.

MEA (2005) Millennium Ecosystem Assessment, 2005. *Ecosystems and Human Well-being: Synthesis*. Island Press, Washington, D.C.

Meigh, J.R., McKenzie, A.A. and Sene, K.J., (1999) A grid-based approach to water scarcity estimates for eastern and southern Africa. Water Resources Management, 13(2), pp.85-115.

Messager, M. L. et al. (2016) 'Estimating the volume and age of water stored in global lakes using a geo-statistical approach', Nature communications, 7(13603).

Meyer, J. L. (1997) 'Stream health: Incorporating the human dimension to advance stream ecology', Journal of the North American Benthological Society, 16(2), pp. 439–447. doi: 10.2307/1468029.

Michener, W.K., (2015) Ecological data sharing. Ecological informatics, 29, pp.33-44.

Millard, K. and Richardson, M., (2015) On the importance of training data sample selection in random forest image classification: A case study in peatland ecosystem mapping. Remote sensing, 7(7), pp.8489-8515.

Morgan, M. G. (2014) 'Use (and abuse) of expert elicitation in support of decision making for public policy', Proceedings of the National Academy of Sciences of the United States of America, 111(20), pp. 7176–7184. doi: 10.1073/pnas.1319946111.

MRC (2010) Biomonitoring Methods for the Lower Mekong. Vientiane, Lao PDR: Mekong River Commission.

MRC (2019a) 2017 Lower Mekong Regional Water Quality Monitoring Report. Vientiane, Lao PDR.

MRC (2019b) Report on the 2017 Biomonitoring Survey of the Lower Mekong River and Selected Tributaries. Vientiane, Lao PDR.

MRC (2019c) Summary, State of the Basin Report 2018. Vientiane, Lao PDR.

MRC (2021) The integrated water resources management–based Basin Development Strategy for the Lower Mekong Basin 2021–2030 and the MRC Strategic Plan 2021–2025, MRC Strategic Plan. Vientiane, Lao PDR.

Mulligan, M., van Soesbergen, A. and Sáenz, L. (2020) GOODD, a global dataset of more than 38,000 georeferenced dams. *Scientific Data*. 7. 31. 10.1038/s41597-020-0362-5.

Musie, M., Sen, S. and Srivastava, P., (2019) Comparison and evaluation of gridded precipitation datasets for streamflow simulation in data scarce watersheds of Ethiopia. Journal of Hydrology, 579, p.124168.

Musonge, P. L. S. et al. (2020) 'Rwenzori score (RS): A benthic macroinvertebrate index for biomonitoring rivers and streams in the Rwenzori Region, Uganda', Sustainability, 12(24), pp. 1–18. doi: 10.3390/su122410473.

Nel, L. l. and Driver, A. (2015) 'National River Ecosystem Accounts for South Africa', South African National Biodiversity Institute, pp. 1–82.

Newbold, T. et al. (2016) 'Has land use pushed terrestrial biodiversity beyond the planetary boundary? A global assessment', Science, 353(6296), pp. 288–291.

Newbold, T. et al. (2019) 'Reply to "The biodiversity intactness index may underestimate losses"', Nature Ecology and Evolution. Springer US, 3(6), pp. 864–865. doi: 10.1038/s41559-019-0896-0.

Nichols, S. J. and Dyer, F. J. (2013) 'Contribution of national bioassessment approaches for assessing ecological water security: An AUSRIVAS case study', Frontiers of Environmental Science and Engineering, 7(5), pp. 669–687. doi: 10.1007/s11783-013-0556-6.

Niemeijer, D. and de Groot, R. S. (2008) 'A conceptual framework for selecting environmental indicator sets', Ecological Indicators, 8(1), pp. 14–25. doi: 10.1016/j.ecolind.2006.11.012.

Nilsson, C., Reidy, C.A., Dynesius, M., Revenga, C. (2005). Fragmentation and flow regulation of the world's large river systems. Science, 308 (5720), pp. 405–408. https://doi.org/10.1126/science.1107887

Niro, F., Goryl, P., Dransfeld, S., Boccia, V., Gascon, F., Adams, J., Themann, B., Scifoni, S. and Doxani, G., (2021) European Space Agency (ESA) Calibration/Validation Strategy for Optical Land-Imaging Satellites and Pathway towards Interoperability. Remote Sensing, 13(15), p.3003.

O'Brien, A. et al. (2016) 'How is ecosystem health defined and measured? A critical review of freshwater and estuarine studies', Ecological Indicators. Elsevier Ltd, 69, pp. 722–729. doi: 10.1016/j.ecolind.2016.05.004.

O'Brien, G.C., Dickens, C., Hines, E., Wepener, V., Stassen, R., Quayle, L., Fouchy, K., Mackenzie, J., Graham, P.M., Landis, W.G. (2018). A regional-scale ecological risk framework for environmental flow evaluations. Hydrology and Earth System Sciences, 22, pp. 957–975. https://doi.org/10.5194/hess-22-957-2018

OECD (1993) OECD Core Set of Indicators for Environmental Performance Reviews. Environment Monographs N° 83. OCDE/GD(93)179

Oki, T. and Sud, Y.C., (1998) Design of Total Runoff Integrating Pathways (TRIP)—A global river channel network. Earth interactions, 2(1), pp.1-37.

Oliveira, A.M., Fleischmann, A.S. and Paiva, R.C.D., (2021) On the contribution of remote sensing-based calibration to model hydrological and hydraulic processes in tropical regions. Journal of Hydrology, 597, p.126184.

ONU - The Committee of Experts on Environmental-Economic Accounting (2021) 'System of Environmental-Economic Accounting — Ecosystem Accounting Final Draft', p. 362 pp. Available at: https://unstats.un.org/unsd/statcom/52nd-session/documents/BG-3f-SEEA-EA_Final_draft-E.pdf.

Palmer, R. W. and Taylor, E. D. (2004) 'The Namibian Scoring System (NASS) version 2 rapid bio-assessment method for rivers', African Journal of Aquatic Science, 29(2), pp. 229–234.

Pandeya, B., Buytaert, W., Zulkafli, Z., Karpouzoglou, T., Mao, F. and Hannah, D.M., (2016) A comparative analysis of ecosystem services valuation approaches for application at the local scale and in data scarce regions. Ecosystem Services, 22, pp.250-259.

Pekel, J.-F. et al. (2016) 'High-resolution mapping of global surface water and its long-term changes', Nature, 540(418).

Pelissari, R., Oliveira, M.C., Abackerli, A.J., Ben-Amor, S. and Assumpção, M.R.P., (2021) Techniques to model uncertain input data of multi-criteria decision-making problems: a literature review. International Transactions in Operational Research, 28(2), pp.523-559.

Poff, N.L., Allan, J.D., Bain, M.B., Karr, J.R., Prestegaard, K.L., Richter, B., Sparks, R., Stromberg, J., (1997) The natural flow regime: a paradigm for river conservation and restoration. BioScience 47, 769–784.

Poff, N.L., Richter, B.D., Arthington, A.H., Bunn, S.E., Naiman, R.J., Kendy, E., Acreman, M., Apse, C., Bledsoe, B.P., Freeman, M.C., Henriksen, J., Jacobson, R.B., Kennen, J.G., Merritt, D.M., O'Keeffe, J.H., Olden, J.D., Rogers, K., Tharme, R.E., Warner, A. (2010). The ecological limits of hydrologic alteration (ELOHA): a new framework for developing regional environmental flow standards. Freshw. Biol. 55, 147–170. https://doi.org/10.1111/j.1365-2427.2009.02204.x.

Poikane, S. et al. (2014) 'Intercalibration of aquatic ecological assessment methods in the European Union: Lessons learned and way forward', Environmental Science and Policy. Elsevier Ltd, 44, pp. 237–246. doi: 10.1016/j.envsci.2014.08.006.

Railsback, S.F. (2016) Why It Is Time to Put PHABSIM Out to Pasture, Fisheries, 41 :12, 720-725, DOI: 10.1080/03632415.2016.1245991

Reinecke, R., Foglia, L., Mehl, S., Trautmann, T., Cáceres, D. and Döll, P., (2019) Challenges in developing a global gradient-based groundwater model (G 3 M v1. 0) for the integration into a global hydrological model. Geoscientific Model Development, 12(6), pp.2401-2418.

Richter, B.D., Baumgartner, J.V., Powell, J., Braun, D.P. (1996). A method for assessing hydrologic alteration within ecosystems. Conservation Biology, 10 (4), 99. 1163–1174. https://doi.org/10.1046/j.1523-1739.1996.10041163.x.

Richter, B.D., Baumgartner, J. V., Wigington, R., Braun, D.P. (1997). How much water does a river need? Freshwater Biology, 37, pp. 231–249. https://doi.org/10.1046/J.1365-2427.1997.00153.X.

Ritchie, J. C., Zimba, P. V. and Everitt, J. H. (2003) 'Remote sensing techniques to assess water quality', Photogrammetric Engineering and Remote Sensing, 69(6), pp. 695–704. doi: 10.14358/PERS.69.6.695.

Robinson W, B. R. (2017) Briefing Paper 3 IECA Framework: Approaches to Aggregation, harmonisation and integration. Integrated Ecosystem Condition Assessment (IECA) Framework Phase 3 Report to the Department of the Environment and Energy.

Rockström J, Steffen W, Noone K, Persson A, Chapin FS, Lambin E, Lenton TM, Scheffer M, Folke C, Schellnhuber H, Nykvist B, De Wit CA, Hughes T, van der Leeuw S, Rodhe H, Sörlin S, Snyder PK, Costanza R, Svedin U, Falkenmark M, Karlberg L, Corell RW, Fabry VJ, Hansen J, Walker B, Liverman D, Richardson K, Crutzen P, Foley J (2009) Planetary boundaries: exploring the safe operating space for humanity. Ecol Soc 14(2):32. http://www. ecologyandsociety.org/vol14/iss2/art32.

Rodrigues, A. S. L. et al. (2014) 'Spatially explicit trends in the global conservation status of vertebrates', PLoS ONE, 9(11), pp. 1–17. doi: 10.1371/journal.pone.0113934.

Ross, M. R. V. et al. (2019) 'AquaSat: A Data Set to Enable Remote Sensing of Water Quality for Inland Waters', Water Resources Research, 55(11), pp. 10012–10025. doi: 10.1029/2019WR024883.

Rowntree, K. M. (2013) 'Module B: Geomorphology Driver Assessment Index (GAI).', in River EcoClassification: Manual for EcoStatus Determination (version 2). Pretoria, South Africa: Joint Water Research Commission and Department of Water Affairs and Forestry report. WRC Report No TT 551/13.

Russi D, ten Brink P, Farmer A, Badura T, Coates D, Förster J, Kumar R, Davidson N (2013) The economics of ecosystems and biodiversity for water and wetlands. European Environmental Policy (IEEP), London/Brussels; Gland: Ramsar Secretariat.

Saaty, T., (2005) The analytic hierarchy and analytic network processes for the measurement of intangible criteria and for decision-making. Multiple Criteria Decision Analysis: State of the Art Surveys, pp. 345–405.

Schofield, N.J. and Davies, P.E. (1996) "Measuring the health of our rivers," *Water*, 5(6), pp. 39–43.

Scholes, R. J. and Biggs, R. (2005) 'A biodiversity intactness index', Nature, 434(7029), pp. 45–49. doi: 10.1038/nature03289.

Scholes, R.J. *et al.* (2012) “Building a global observing system for biodiversity,” *Current Opinion in Environmental Sustainability*, pp. 139–146. Available at: https://doi.org/10.1016/j.cosust.2011.12.005.

Sengupta, A., Adams, S.K., Bledsoe, B.P., Stein, E.D., McCune, K.S., Mazor, R.D. (2018). Tools for managing hydrologic alteration on a regional scale: Estimating changes in flow characteristics at ungauged sites. Freshwater Biology, 63 (8), pp. 769–785. https://doi.org/10.1111/fwb.13074.

Sharma, D. and Kansal, A. (2011) ‘Water quality analysis of River Yamuna using water quality index in the national capital territory, India (2000–2009)’, Applied Water Science, 1(3–4), pp. 147–157. doi: 10.1007/s13201-011-0011-4.

Sheldon, F. et al. (2012). ‘Identifying the spatial scale of land use that most strongly influences overall river ecosystem health score’, Ecological Applications, 22(8), pp. 2188–2203. doi: 10.1890/11-1792.1.

Shen, C. (2018). A transdisciplinary review of deep learning research and its relevance for water resources scientists. Water Resources Research, 54(11), pp.8558-8593

Shiau, J.T., Wu, F.C. (2008). A Histogram Matching Approach for assessment of flow regime alteration: application to environmental flow optimization. River Research and Applications, 24, pp. 914–928. https://doi.org/10.1002/rra.1102.

Sichangi, A.W., Wang, L. and Hu, Z., (2018) Estimation of river discharge solely from remote-sensing derived data: An initial study over the Yangtze river. Remote Sensing, 10(9), p.1385.

Smeets, E. and Weterings, R. (1999) Environmental indicators: typology and overview, Technical Report No 25. Copenhagen.

Srebotnjak, T. et al. (2012) ‘A global Water Quality Index and hot-deck imputation of missing data’, Ecological Indicators. Elsevier Ltd, 17, pp. 108–119. doi: 10.1016/j.ecolind.2011.04.023.

Stacke, T. and Hagemann, S., (2021) HydroPy (v1. 0): A new global hydrology model written in Python. Geoscientific Model Development Discussions, pp.1-28.

Stark, J. D. (1985) ‘A macroinvertebrate community index of water quality for stony streams.’, Water & Soil miscellaneous publications, 87, p. 53 p.

Stark, J. D. (1998) ‘SQMCI: A biotic index for freshwater macroinvertebrate coded-abundance data’, New Zealand Journal of Marine and Freshwater Research, 32(1), pp. 55–66. doi: 10.1080/00288330.1998.9516805.

Steffen, W. et al. (2015) ‘Planetary boundaries: Guiding human development on a changing planet’, Science, 347(6223). doi: 10.1126/science.1259855.

Steinmetz, M. and Sundqvist, N. (2014) “Environmental impacts of small hydropower plants-a case study of Borås Energi och Miljö’s hydropower plants,” *Master of Science thesis. Chalmers University of Technology, Gothenburg, Sweden*.

Stisen, S. and Sandholt, I., (2010) Evaluation of remote-sensing-based rainfall products through predictive capability in hydrological runoff modelling. Hydrological Processes: An International Journal, 24(7), pp.879-891.

Smith, R.A., Alexander, R.B. and Schwarz, G.E., (2003) Natural background concentrations of nutrients in streams and rivers of the conterminous United States.

Solomatine, DP, Abrahart, R., See L. (2008) Data-driven modelling: concept, approaches, experiences. , In: Practical Hydroinformatics: Computational Intelligence and Technological Developments in Water Applications (Abrahart, See, Solomatine, eds), Springer-Verlag.

Soranno, P.A., Cheruvelil, K.S., Liu, B., Wang, Q., Tan, P.N., Zhou, J., King, K.B., McCullough, I.M., Stachelek, J., Bartley, M. and Filstrup, C.T., (2020) Ecological prediction at macroscales using big data: Does sampling design matter?. Ecological Applications, 30(6), p.e02123.

Spence, R. Hickley, P. (2020) The use of PHABSIM in the management of water resources and fisheries in England and Wales Ecological Engineering Volume 16, Issue 1, October 2000, Pages 153-158

Strokal, M., Spanier, J.E., Kroeze, C., Koelmans, A.A., Flörke, M., Franssen, W., Hofstra, N., Langan, S., Tang, T., van Vliet, M.T. and Wada, Y., (2019) Global multi-pollutant modelling of water quality: scientific challenges and future directions. Current opinion in environmental sustainability, 36, pp.116-125.

Sutanudjaja, E.H., Van Beek, R., Wanders, N., Wada, Y., Bosmans, J.H., Drost, N., Van Der Ent, R.J., De Graaf, I.E., Hoch, J.M., De Jong, K. and Karssenberg, D., (2018) PCR-GLOBWB 2: a 5 arcmin global hydrological and water resources model. Geoscientific Model Development, 11(6), pp.2429-2453

Tampo, L. et al. (2020) ‘A multimetric index for assessment of aquatic ecosystem health based on macroinvertebrates for the Zio river basin in Togo’, Limnologica. Elsevier, 83(July 2019), p. 125783. doi: 10.1016/j.limno.2020.125783.

Thirion, C. (2007) ‘Module E: Macroinvertebrate Response Assessment Index (MIRAI). WRC Report No TT 332/08’, in River EcoClassification: Manual for EcoStatus Determination (version 2). Pretoria, South Africa: Joint Water Research Commission and Department of Water Affairs and Forestry report.

Tickner, D. et al. (2020) ‘Bending the Curve of Global Freshwater Biodiversity Loss: An Emergency Recovery Plan’, BioScience, 70(4), pp. 330–342. doi: 10.1093/biosci/biaa002.

Topp, S.N., Pavelsky, T.M., Jensen, D., Simard, M. and Ross, M.R., (2020) Research trends in the use of remote sensing for inland water quality science: Moving towards multidisciplinary applications. Water, 12(1), p.169.

UNDESA (2012) System of Environmental-Economic Accounting: Accounting for Water, UNDESA. ISBN: 978-92-1-161554-8.

UNEP, (2016) A Snapshot of the world’s water quality: towards a global assessment, UNEP, Nairobi, Kenya, 162 pp

UNEP (2017) A framework for freshwater ecosystems management vol. 1 overview and country guide for implementation. UN Environment, Nairobi.

UNEP-WCMC (2016) The state of biodiversity in Africa: a mid-term review of progress towards the Aichi biodiversity targets. Cambridge

United Nations (2015) Transforming our World: The 2030 Agenda for Sustainable Development A/RES/70/1. doi: 10.1201/b20466-7.

United Nations (2018) ‘An Introduction to Ecosystem Accounting: Key concepts and Policy Applicatons’, in System of Environmental Economic Accounting. New York, USA: UN Statistics Division (UNSD), Environmental Economic Accounts Section, pp. 213–219. doi: 10.4324/9781315775302-19.

United Nations Global Environment Monitoring System. (no date). Available at: http://gemstat.org/.

United Nations Statistical Division SDG Indicators Repository (no date) Available at: https://unstats.un.org/sdgs/metadata/

UN Water (2017) ‘Intergraded Monitoring Guide Sustainable Development Goal 6 on Water and Sanitation: Targets and global indicators.’, Version 14(July 2017), p. 40. Available at: https://www.unwater.org/publications/sdg-6-targets-indicators/.

Unwin, M. and Larned, S. (2013) ‘Statistical models, indicators and trend analyses for reporting national-scale river water quality (NEMAR Phase 3). Prepared for Ministry for the Environment. NIWA Client Report No. CHC2013-033.’

USEPA (2006) Wadeable Streams Assessment: A Collaborative Survey of the Nation’s Streams, Office of Water, US Environmental Protection Agency, Washington DC. Washington DC: Office of Water, US Environmental Protection Agency.

USEPA (2020a) National Rivers and Streams Assessment 2013–2014: A Collaborative Survey. EPA 841-R-19-001. Washington ,DC.

USEPA (2020b) ‘National Rivers and Streams Assessment Survey Design: 2013–2014’, pp. 2013–2014.

USGS, n.d., What is remote sensing and what is it used for?, https://www.usgs.gov/faqs/what-remote-sensing-and-what-it-used?qt-news_science_products=0#qt-news_science_products, last access: 11/23/2021

Van de Bund, W. and Solimini, A. G. (2007) Ecological Quality Ratios for ecological quality assessment in inland and marine waters. Luxembourg: Institute for Environment and Sustainability, European Commission Directorate-General Jointy Research Centre. Available at: http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:Ecological+Quality+Ratios+for+ecological+quality+assessment+in+inland+and+marine+waters#8.

Vogel, R., Seiber, J., Archfield, S., Smith, M., Apse, C., Huber-Lee, A. (2007). Relations among storage, yield, and instream flow. Water Resources Research, 43. https://doi.org/10.1029/2006WR005226.

Vollmer, D. et al. (2018) ‘Integrating the social, hydrological and ecological dimensions of freshwater health: The Freshwater Health Index’, Science of the Total Environment. The Authors, 627, pp. 304–313. doi: 10.1016/j.scitotenv.2018.01.040.

Vongsombath, C. et al. (2009) Report on the 2006 biomonitoring survey of the lower Mekong River and selected tributaries, MRC Technical Paper. Vientiane, Lao PDR.

Vorosmarty, C.J., Sharma, K.P., Fekete, B.M., Copeland, A.H., Holden, J., Marble, J. (1997). The storage and aging of continental runoff in large reservoir systems of the world. Ambio, 26, pp. 210–219.

Vörösmarty, C. J. et al. (2010) ‘Global threats to human water security and river biodiversity’, Nature, 467(7315), pp. 555–561. doi: 10.1038/nature09440.

Walley, W. J. and Hawkes, H. A. (1996) ‘A computer-based reappraisal of the biological monitoring working party scores using data from the 1990 river quality survey of England and Wales’, Water Research, 30(9), pp. 2086–2094.

Walsh, C.J. (2006) “Biological indicators of stream health using macroinvertebrate assemblage composition: a comparison of sensitivity to an urban gradient,” *Marine and Freshwater Research*, 57(1), pp. 37–47.

Whittington, J. *et al.* (2001) *Development of a Framework for the Sustainable Rivers Audit: A report to the Murray Darling Basin Commission*.

Widén-Nilsson, E., Halldin, S. and Xu, C.Y., (2007) Global water-balance modelling with WASMOD-M: Parameter estimation and regionalisation. Journal of Hydrology, 340(1-2), pp.105-118.

Wilkinson, S.N.; Henderson, A.E.; Chen, Y. (2004) SedNet User Guide. Report to the Cooperative Research Centre for Catchment Hydrology. Canberra: CSIRO Land and Water; http://hdl.handle.net/102.100.100/184593?index=1

Williams, J. (2010) Comment on Gard (2009): Comparison of spawning habitat predictions of PHABSIM and River2D models. Intl. J. River Basin Management. 8. 117-119. 10.1080/15715121003714977.

WWAP (UN World Water Assessment Programme) (2018) The UN World Water Development report 2018: nature-based solutions. UNESCO, Paris

WWF (2020a) ‘A deep dive into freshwater’, in Living planet report 2020. Available at: http://www.ecoguinea.org/papers-development.html.

WWF (2020b) Living Planet Report 2020 - Bending the curve of biodiversity loss. Edited by R. E. A. Almond, G. M., and T. Petersen. Gland, Switzerland: WWF.

Xie, C. et al. (2020) ‘A nation-wide framework for evaluating freshwater health in China: Background, administration, and indicators’, Water (Switzerland), 12(9). doi: 10.3390/W12092596.