An integrated toolkit for assessment of hydrology-dependent outcomes in the Murray-Darling Basin: HydroBOT

Galen Holt

Georgia Dwyer

David Robertson

Martin Job

Lara Palmer

Rebecca E Lester

2024-07-05

Table of contents

# 1. Graphical abstract

(../images/GraphicalAbstract-01.png){#fig-tktomgmt}

# 2. Abstract

*WRITE EORDS HERE GEORGIA/GALEN*

FIND.REPLACE for consistent terminology broader ecological values OR environmental targets ?? ecological objective OR ecological objective flow OR water availability

historical base level

# 3. Introduction

Water management in large river systems typically targets a wide range of desired values. For example, in the Murray-Darling Basin (MDB), Australia, the objective of water management is to maintain a healthy working river that supports productive and resilient water-dependent industries, healthy and resilient ecosystems, and communities with access to sufficient and reliable water supplies (Murray-Darling Basin Authority 2011). These values, defined here as any social, economic, environmental or cultural asset or function of significance, importance, worth, or use, are not unique to Australia; water is managed to protect these values of communities worldwide (e.g. Stern and Sheikh 2019; Kaye-Blake et al. 2014) (**connor2014?**). When these values are the target of water management implies that each value depends on water in some way but, in many cases, these dependencies are not well-defined and management proceeds under the relatively simple assumption that, if water is provided, these values will be maintained.

Water managers typically have a limited range of levers at their disposal which primarily affect hydrology, e.g. flow releases from dams or inundation of wetlands. When non-hydrologic levers are available, such as water trading rules in the Murray-Darling Basin, their impact is still often assessed on how they alter hydrologic conditions– their impact on the spatio-temporal patterns of flow in the system. Further, hydrologic modelling is often better-integrated with management workflows than models of other values. In part, this is due to the availability of large-scale physical models that provide robust ability to model flows as they arise from natural drivers, such as rainfall, and capture infrastructure such as dams and diversions. While complex, these models have relatively high precision and relatively few outcomes compared to models of ecological or economic responses, for example. The impact of particular management actions (e.g. dam releases) on hydrology is well-understood and captured by the hydrograph, allowing accurate targeting of particular aspects of the flow regime by management actions (e.g. (**Loire2021?**)). While these proximate models of system state and impact of management actions are invaluable, they do not provide the necessary information to assess the status of the range of flow-dependent values.

Assessing values as they respond to flow and management actions requires models of these relationships. Such models exist, though their quality and extent vary widely, ranging from unstated mental models to highly-detailed population dynamics or economic models (Lester 2019). When these models are quantitative or computational, they are written in various languages by subject-matter experts, often focus on one or a limited suite of *GEORGIA* target values and return disparate outputs depending on their particular goals and approaches. These models are often designed with the goal of studying the dynamics of highly specific variables (e.g. fish population responses to water management regimes), not to produce the most useful analysis for management questions or to capture the breadth of target values. Moreover, these models cannot individually provide information about larger-scale values; single-species models are insufficient to assess the condition of all fish or the whole ecosystem, for example. Integrating these models into a holistic modelling approach is necessary to assess management-relevant outcomes of disparate, typically broad-scale, values under different hydrologic conditions and management actions. Integrating diverse response models gives each response model utility beyond its original purpose and identifies where new work is needed (e.g. where response models are limited or non-existent). The creation of an integrated toolkit provides the opportunity for robust decision-making (Harrison et al. 2023) and the ability to prioritize water planning across a range of values and identify conditions that achieve disproportionately large (or small) impacts, as well as the explicit consideration of synergies and trade-offs among different values.*These first few paras probably need more general cites in them GEORGIA*

Exemplifying these issues, the Murray-Darling Basin Authority (MDBA) is a federal body charged with water management in the MDB, Australia. The Murray-Darling Basin is centrally important to the economic, cultural, social and environmental well being of Australia. Within the MDB, 70% of water is utilized for agricultural purposes. The MDB contributes nearly 2% to gross domestic product from agriculture and tourism, is home to 2.2 million people, including more than 40 Aboriginal Nations, and contains approximately 400,000 water-dependent ecosystems (Hart et al. 2021). While balancing this range of values is difficult in any highly-utilized basin, the Murray-Darling is unusually dry and variable compared to other systems of similar importance (Hart et al. 2021). Modern management of the MDB arises from the Commonwealth (**wateract2007?**), which established the Murray-Darling Basin Authority with obligations to develop and implement a Basin Plan governing water resource management (Hart et al. 2021) *I’m citing Hart a lot because it’s a big overview, but are there better cites? Something internal?*.

The MDBA has obligations to manage water to maintain a healthy working river and the state of the MDB is required to be assessed annually, with a major review of, and updates to, the Basin Plan made at legislated intervals (e.g. 15 years between development and first review, Hart et al. 2021). Beyond these requirements, the MDBA has an increasing focus on assessing system response to climate change and possible adaptation actions that may be taken by the MDBA itself or other stakeholders in the MDB (**neave2015?**). In addition to the MDBA, Basin states and other stakeholders (e.g. the Commonwealth Environmental Water Holder) have additional responsibilities for water management in the MDB (Hart et al. 2021) *Is there a better cite for who has control over what?*. Taken together, these management goals require a holistic modelling approach that integrates across values and is adaptable to these various management needs while also simplifying repeated, ongoing use. These needs are not unique to the MDBA and other water management groups, researchers and stakeholders are likely to have similar complex and competing needs.

The MDBA has access to robust hydrologic models describing flows in the system and responses to current management practices, and these models are under continual use, development, and improvement ( *What’s the best cite? David, Martin?*). Despite the need for ongoing assessment and forward planning across a range of values, models of the responses to those hydrologic conditions are patchy, only represent some values, and have been developed and used in a more ad-hoc approach rather than integrated with each other or the hydrologic models. In this paper, we describe an integrated modelling ‘toolkit’ to address these water management needs in the Murray-Darling Basin, Australia, hereafter referred to as *HydroBOT* (Hydrology-dependent Basin Outcomes Toolkit).

The toolkit described here greatly improves the capacity of the relevant manager (here, the MDBA) to assess outcomes across a range of values, provides the structure to adapt and include additional values as additional component models become available, and addresses a number of issues with current assessment practices and modelling approaches. By providing a single, consistent interface to a range of response models, we avoid the need to manually run models separately and we abstract their different interfaces, languages, and idiosyncrasies. Moreover, the toolkit design is highly modular, built to allow the integration of new response models with limited additional updates to HydroBOT itself. Continuing with this consistency, HydroBOT provides a standard set of synthesis approaches and functions that target management-relevant analysis and interpretation of outcomes from these disparate models with a common approach and design language. Because the need for assessing outcomes is nearly universal in water management, and not limited to a particular scenario or project, HydroBOT is designed with strong scenario-comparison capabilities but is agnostic to what those scenarios represent. For a similar reason, the standardised synthesis and outputs are highly flexible, allowing the user to choose the most relevant outputs for different management needs.

One common issue with modelling in general, and particularly integrated models spanning several tools, is that such tools can become a black box where the relationships modelled are opaque. This can yield mistrust by the public and other stakeholders influenced by the model, but also mistrust and misunderstanding by users and developers of the model. To avoid these issues, HydroBOT has been continuously co-designed with the relevant management agency, the MDBA, and an emphasis is put on public code, reproducibility, self-documentation and production of outputs that describe the model actions (e.g. causal relationships).

To demonstrate HydroBOT, we develop a set of example scenarios capturing some qualitative aspects of one intended use: the assessment and comparison of outcomes under different climate scenarios and different adaptations to those changes. These axes represent changes to the system due to processes over which managers have no control (i.e. a changing climate), and management actions which would typically be targeted interventions. We use this demonstration to illustrate the problem space and show how the toolkit can aid in assessing potential system change and prioritising management actions to mitigate impacts.

# 4. Methods

The purpose of HydroBOT is to assess how hydrographs representing various climate and adaptation scenarios affect multiple sets of values, which may span many scales and disciplines. We seek to make these assessments in a consistent and repeatable way despite differences in the response models. HydroBOT then has the capability to synthesize a wide range of outputs from the response models into results that compare scenarios and are digestible and useful for management decision making (**?@fig-tktomgmt**). A co-design process including scientists, software developers, and managers was developed to ensure the toolkit achieved its goal of producing scientifically robust results that are also management-relevant.

Conceptualisation of HydroBOT archetecture (blue arrows) and benefits to management (yellow arrows). Scenarios, represented by hydrographs, reflect observed flows or modeled scenarios (produced via hydrological modelling outside HydroBOT). These hydrographs are fed in a consistent way to various response models (see [Section 4.5.1](#sec-modules)) via the HydroBOT Controller (see [Section 4.3.1](#sec-controller)). Additional inputs of spatial data (see [Section 4.4](#sec-spatial_data)) and causal networks (see [Section 4.5](#sec-causal_networks)) provide grouping information for scalling via the HydroBOT Aggregator (see [Section 4.5.2](#sec-aggregator)). the HydroBOT Comparer synthesizes the results into management-relevant outputs that aid decision-making (see [Section 4.5.3](#sec-comparer)). This decision-making process is also supported by feedback from stakeholders, other expert advice and the social and political landscape. Water planning and policy decisions can then guide the development of scenarios to assess potential adaptaion options, which in turn can be run through the HydroBOT archetecture.{height = 30 width=30}

## 4.1 Co-design and scoping for management relevance

Avoiding HydroBOT becoming a ‘black box’ builds trust, interpretability and usability of outputs created with the toolkit. To ensure HydroBOT meets management needs and is trusted by management users, a toolkit development team was created consisting of primary toolkit developers, hydrologic modelers, and MDBA staff. Collaboration among this team ensured decisions about toolkit construction reflected the needs of the end-user. As implementation proceeded, the collaboration provided a mechanism by which goals could be adjusted or the implications of decisions discussed and understood. By having this window into the toolkit development, the managers obtained a much more granular view of how it operates, and its capabilities and limitations.

Key to the collaborative development was identification of the response models to include. These models need to provide information about values relevant to water management decisions as those values respond to hydrology. Early in the development process, a wide scan was taken to identify candidate response models. This scan considered models being used or developed within the MDBA itself, other government agencies (both federal and state), and externally. This scan found that there were a number of responses models available, but many were outdated or highly manual, and their use was patchy both within and across management agencies (**holt2022?**). Ultimately, only one was suitably modern and well-developed to include in HydroBOT, the Environmental Water Response (EWR) Tool. Other in-development modules were identified as candidates to include during the life of the toolkit, including economic and social models. Further, this scan identified values that are mandated management targets for which there were no available or in-development response models, highlighting areas needing future work.

## 4.2 Toolkit overview

The key goals of the HydroBOT are to provide a flexible platform for comparing outcomes between different possible scenarios. Scenario comparison is essential for use in both evaluating past actions or conditions and planning for the future, whether to better understand potential shifts in the system from external forces (e.g. climate change) or to assess potential management actions. However it can be difficult to collapse a multitude of scenarios into a format that is easily digested by decision-makers. Further, this process often needs to be supported by feedback from stakeholders, other expert advice and the social and political landscape. Accordingly, a flexible architecture with distinct components that can be run separately (modular) or together can assist with the iterative nature of decisions-making for Water planning and policy.

### 4.2.1 Architecture

Taken holistically, the architecture of the toolkit comprises five primary components and the links between them (**?@fig-architecture**). The data flow through the toolkit is represented by the bold arrows. Input data (hydrographs) is ingested by the Controller, which packages and runs Response Models. Results from the Response Models are then processed by the Aggregator, and the analysis and final outputs are prepared by the Comparer. The two components outside this data flow are the Causal networks and Spatial units, which are used within the Aggregator and Comparer to define groupings for aggregation and for visual representation of space and the complex inter-relationships defined in the response models. Each of these components is described in more detail below.

The architecture of the toolkit emphasizes modularity. Each of the components can save its outputs, allowing users to run either the whole toolkit or re-run needed components only to update analysis. For example, if a user wished to change how they aggregate the results of the modules, they could re-run the aggregation step and the subsequent comparison step to match. This ability to adjust intermediate steps allows rapid iteration of results to address a given management question, or adaptation of preexisting results to new questions. This modularity also allows rapid, iterative development of the toolkit itself. Any of the components of the toolkit can be updated without affecting others, and so obtaining new results from updated components (e.g. new aggregation capability) simply requires re-running the toolkit from that point forward.

As each stage of the toolkit runs, it produces yaml and json metadata files including the run settings and other attached information relating to the scenarios to enhance repeatability, ensure correctness and increase comprehension. These metadata include all parameters for the run, allowing exactly repeatable analyses to be conducted by using the metadata files as parameter files for subsequent runs. Even if a run is started in the middle, such as to re-run aggregation in a different way, the metadata for that run will capture the metadata for the previous steps, ensuring that outputs at every stage are tagged with full provenance information about the run that created them.

### 4.2.2 Implementation

The toolkit is available as an R package *GIT HUB LINK*, which provides a suite of functions representing the steps in the architecture. These functions are designed to be general, allowing users much flexibility in how they run the toolkit for a particular set of analyses while retaining a consistent structure and outputs. Because translating from hydrologic scenarios to various responses is a general problem in water management, we expect the ways in which the toolkit is used will be highly variable. Thus, by providing a general structure in the toolkit, users can target the particular questions and particular analyses needed for a given question. For example, in some cases we might want to look at the responses of different components of fish life cycles for a small subset of locations in the MDB. In another situation, we might want a big-picture view of how climate might shift long-term goals for the environment as a whole. Although the initial impetus for creating the toolkit was to assess climate scenarios, its use in practice can be far more general. Any hydrograph can be assessed, and any set of scenarios represented by hydrographs can be compared, provided a response model exists.

The toolkit is designed for analysis of management questions in the Murray-Darling Basin, and so along with the functions to perform the analyses, it also provides a standard set of spatial data comprising the MDB itself, gauges within the MDB (at which hydrographs may be available), and various management units (Sustainable Diversion Limit units, Resource Plan Areas, and catchment boundaries for major subcatchments). Some example hydrograph data are also included for testing and demonstration. The toolkit provides a clean causal network for included modules, described in more detail in [Section 4.5](#sec-causal_networks). Analogous elements can be constructed for other spatial units or responses to extend functionality within or outside the Murray-Darling Basin.

As a result of the scan of available response models, toolkit development proceeded with the EWR tool as a single response model, but with an architecture designed to allow modular integration of additional response models as they become available. The scan accentuated a critical feature of this modularity; the toolkit must be able to incorporate and standardise models written in various languages and with a wide range of input needs and outputs. The toolkit achieves this by wrapping those other tools so as to make their differences as hidden from the user as possible. In the case of Python modules, the toolkit uses the {reticulate} R package (Ushey, Allaire, and Tang 2023) in combination with small amounts of internal Python code to call Python modules directly. The Python code internal to the toolkit performs limited cleaning and translation to prevent passing large objects between languages and ensure that any idiosyncrasies in module inputs and outputs are handled consistently. Python dependencies (and Python itself) are automatically installed on first use of Controller functions that call Python modules unless they already exist on the user’s system. This approach provides essentially invisible Python for most users, while providing flexibility for the user to provide their own Python environment if desired. Modules in other languages are not yet available, but the key requirement is that they be available in a format that is scriptable. In that case, the toolkit will provide small setup and cleanup functions as with the Python modules and wrapper functions to call these modules.

The modularity of the toolkit means it can be run stepwise, with the user calling the relevant functions at each step (**?@fig-architecture**). The outputs of each step can be saved or returned directly to an interactive session or both. Typically sufficient selected outputs would be saved for reproducibility and speed unless the project is small enough to retain a comprehensive set in memory or re-run quickly in a notebook. There are also wrapper functions provided that allow running the entire toolkit from Controller through Aggregator, which are extremely useful for large runs or remote runs on batching services. One particularly useful wrapper provides the ability to run from a yaml config file providing function arguments. This function allows the use of a default file and a ‘modified run’ file, making it ideal for holding many parameters constant at a default for a particular set of analyses and only changing some on a per-run basis in a smaller file. It also takes command-line or R list arguments, making it a flexible solution to run the toolkit from the command line, in scripting contexts, or Quarto notebooks. The metadata saved at each step in the process is a yaml file with parameters that are a superset of those needed to run the toolkit. Thus, an exactly identical run can be produced by running the toolkit using an output metadata file as an input parameter file. These metadata files also include the git hash, further allowing reproducibility in the face of code changes.

In practice, the toolkit functions are primarily run in one of two ways. Interactive investigation of relatively small sets of hydrographs can be done in notebooks (typically Quarto; Allaire et al. (2022)) or simple R scripts. Larger investigations typically would be run on remote computers as part of batching systems, whether HPC, Azure, AWS, or other, with the outputs at the end of the aggregation (and potentially each step) returned. The Comparer step would usually not be run as part of this larger batching, but considered interactively for two reasons. First, through the aggregation step, all operations can proceed in parallel over scenarios, and in some cases parallelizing over gauges is possible. The Comparer necessarily looks across scenarios, and so breaks this parallelisation. Conducting comparisons interactively is typically not excessively CPU- or memory-intensive, provided the Aggregator step has been well thought through. If there are enough data to require high processing or memory, it is unlikely to be simplified enough to make interpretable figures. Second, the primary goal of the Comparer is to produce usable, interpretable outputs. Arriving at a set of meaningful outputs is an iterative process that rarely will be known *a priori*, and so working through this step interactively is necessary. If, however, the toolkit is being used for ongoing monitoring of the same analyses (e.g. a ‘dashboard’), then the first iteration may be done interactively, with subsequent uses incorporating those settings written into a script to auto-generate the same figures.

## 4.3 Toolkit components

Here we describe the specific implementation of each component of the toolkit illustrated in **?@fig-architecture**. An overview of this specific implementation is given in [Appendix 1](#tbl-components), with descriptions given here.

### 4.3.1 Controller

The ‘Controller’ component of the toolkit is the interface between the externally generated input data (scenarios), the chosen response model, and other external and internal components of the toolkit (**?@fig-architecture**). This component initiates the downstream processing steps according to user-defined settings for a particular run. It includes arguments for locating the input data and the response model(s) to use along with any necessary parameters for those models. The Controller can also control later components of the toolkit, allowing the full toolkit to be run at once. These include defining aggregation steps as discussed in [Section 4.5.2](#sec-aggregator) and analysis of the results with the Comparer ( [Section 4.5.3](#sec-comparer) ). The Controller determines whether and where outputs are returned at each step. Having control over the full toolkit process enables large batched runs using parameter files to specify the control arguments in yaml files. This core functionality of the Controller is delivered via simple functions that apply to the input data for each scenario and can be looped over scenarios in parallel. The controller can be accessed by the user by using Quarto notebooks to work interactively, R scripts, or via the command line, depending on the use case.

## 4.4 Spatial data

*GEORGIA TEXT FOR SOMEWHERE ELSE?* Causal networks are linked closely to each Response model, and while both are defined externally to the toolkit they require significant work to integrate into a compatible, modular toolkit component. Spatial units are typically more general, typically polygons of management interest, and require only light changes to make compatible.

TEXT FOR SOMEWHERE ELSE? Not pictured, hydrographs and direct Response model output can directly utilise the Comparer functionality without aggregation. Each component of the HydroBOT toolkit is distict, allowing modular changes to be made without altering the function of other components.

## 4.5 Causal networks

Causal networks are models that describe the topology of dependence among many drivers and outcomes of different type Peeters et al. (2022))(**Martínez2019?**). The use of a causal network framing both at a high level for visualisation and communication with the MDBA and embedded in HydroBOT illustrate both the goals and the functioning of the toolkit, avoiding it becoming a black box. The relevant causal network for water management in the Murray-Darling Basin captures relationships between management, flow, and outcomes for a wide range of values, from ecological response to economic performance and human wellbeing. Thus, they include climatic and management drivers, but also include the causal relationships which form the basis of the response models (e.g. flow timeseries to hydrologic indicators, or potentially to the life history inherent in a population dynamics model). They incorporate any links connecting those outcomes to larger-scale outcomes (e.g. from hydrologic indicators to ecological response). Thus, causal networks define an overarching model from initial inputs (e.g. rainfall or flows or adaptations) through to all values of interest, with each link defining a response model or component of a response model. The specifications of the models underlying each link are highly variable. These models range from detailed physical models linking runoff to flow, to simpler ecological models of environmental water requirements, to simple averages, to leaving the model unspecified where information about the relationships is unavailable. Assessing the quality of knowledge around each link provides a powerful assessment of knowledge deficiencies and uncertainty in responses.

Causal networks themselves, i.e. the structure of the links (relationships) and nodes (state variables) can be derived from many sources, including empirical studies defining the existence of causal relationships and expert opinion. HydroBOT provides a causal network for included modules, where available, to describe how their outputs arise from hydrology and how they relate to various levels of management-relevant outcomes. The causal networks enable: 1) visual representation of the complex inter-relationships between scenario inputs and outcomes across a range of objectives and 2) assessment of outcomes aggregated along the value dimension (see [Section 4.5.2](#sec-aggregator)). The former aids transparency, elucidating the intentions and causal relationships behind the response models and is a useful device for communication alongside other final outputs. The latter allows outcomes to be quantified for individual (or sets of) objectives (e.g. fish breeding), values (e.g. native fish), or at the overarching levels of environmental, cultural, social, or economic values. This quantification provides a powerful assessment tool and the ability to identify synergies and trade-offs across interrelated values.

### 4.5.1 Response models

The impacts of climate and adaptation options on social, cultural, environmental, or economic values are estimated based on causal relationships between drivers (e.g. hydrology) and responses (e.g. the state of values). The response models may exist in many different forms, ranging from binary achievement of hydrologic indicators to fully quantitative responses. These tools are expected to be sourced from existing or in-development models developed by subject-matter experts and, as such, will be written in different languages and will target different outcomes. HydroBOT then provides a unified interface and ongoing analysis and modelling of the response model results.

#### 4.5.1.1 EWR (Environmental Water Requirements)

The EWR tool, which forms the core of this demonstration, is one such response model, written in Python and in use by the MDBA internally, the state of New South Wales for water planning and other interested stakeholders. It models the response of ecological values of the Murray-Darling Basin founded on hydrologic indicators paired with causal relationships to ecological values. It holds databases of the environmental water requirements (EWRs; the indicators) and the volume, frequency, timing and duration of flows or inundation required to meet those indicators ([Section 9.2](#sec-ewr-table)). These indicators were developed based on hypothesized relationships to the ecological objectives of the MDB, which protect or enhance environmental assets and ecological functions that are valued based on ecological significance (**sheldon2024?**). The EWR tool itself only provides an assessment of the hydrologic indicators. The EWR tool assesses whether spatially explicit flow timeseries data meets each EWR (hydrologic indicator) at each gauge (illustrated in [Figure 1](#fig-ewr-example)). The precise definitions for each EWR differ at each gauge, due to the unique hydrology and channel morphology. For example, a small fresh for 10 days is required every year, ideally between October and April to meet indicator EWR SF1 (small fresh 1), but the flow volume defined as a ‘small fresh’ differs between gauges. For HydroBOT to model environmental outcomes, we connect the EWR tool results to a specific causal network defining these links.

|  |
| --- |
| Figure 1: The magnitude of required EWR are illustrated at two example gauges (412002 and 419001) on the historical base level (scenario E1) hydrographs. This does not illustrate the requred frequency, duration or timing or EWR; see [Section 9.2](#sec-ewr-table) for a table. |

*GALEN should we still be saying that we developed causal networks?* The EWR tool results are binary responses of hydrologic indicators and so, to link these to expected ecological responses, we developed causal networks from the relationships implied in the Long-Term Watering Plans in setting the EWRs (**DPIE2020abc?**)). These plans describe how the hydrologic indicators are expected to influence both proximate and larger-scale ecological outcomes. We extracted these causal relationships from the Murray-Darling Basin Long Term Watering Plans (LTWPs), which were developed based on the best available information from water managers, ecologists, scientific publications, and analysis of gauged and modelled flows (e.g. Planning and Environment (2020), Lobegeiger and Prior (2022)). This provided a dense network of links across a range of ecological outcomes at various scales, from the hydrologic indicators to components of the life cycle of single species to whole-community outcomes at 10- or 20-year target dates. For example, for the SF1 indicator described above might contribute to multiple ecological objectives pertaining to native fish (ecological objectives NF1-9; e.g. NF1 = No loss of native fish species), native vegetation (NV1), and ecological functions (EF1-5). These ecological objectives are a defined to support the completion of all elements of a life cycle of an organism or group of organisms (taxonomic or spatial) to achieve a goal state, condition or characteristic of an environmental asset or ecological function” *[LTWP2020]*. Outcomes for ecological objectives are then linked in the causal network to five ecological values (native fish, native vegetation, waterbirds, other species, and priority ecological functions) and are associated with long-term targets (5, 10, and 20 year) of the LTWP’s management strategies. The chain of dependencies from EWRs to ecological objectives to long-term targets are captured in the causal networks as links, with nodes defining the ecological objectives or long-term targets (**?@fig-causal-example**). This structure not only provides a visual definition of the links in the LTWP, it also enables assessment of outcomes in direct equivalence to the LTWP’s management strategies.

|  |
| --- |
| Figure 2: a) HydroBOT incorporates causal networks that describe the ecological objectives for a system. In the current example these causal networks are extracted from the Murray-Darling Basin Long Term Watering Plans (LTWPs), which sets environmental watering requirements (EWRs), ecological objectives, and long-term targets for key water-dependent plants, waterbirds, fish and ecological functions. From left to right, the columns here represent EWR codes, ecological objectives, specific goals, ecological values, and 5-year management targets. The network shown here is for a single gauge (412002) and Planning Unit, which is the scale at which EWR codes and ecological objectives are defined. The other levels are defined at larger spatial scales, but only those that apply to the EWRs present at this gauge are shown here. b) Subset of the causal network for gauge 419001 relevant to a selection of *lengthsampleenvobj* ecological objectives. The data are at the gauge scale for the first and second columns (EWRs and ecological objectives) and at the SDL unit scale for the third (ecological values). Thus, the final set of nodes contain information from other gauges that are not shown here for clarity. Colors represent the log of the relative change in condition from the baseline for the halving (panel a, scenario A) and doubling (panel b, scenario I) ‘climate’ scenarios. Large reductions in condition are dark purple, large improvements are dark green, and no change is white. See Appendix 3 *sec-causalnetwork-versions* for more detailed causal network diagrams. |

|  |
| --- |
| Figure 3: a) HydroBOT incorporates causal networks that describe the ecological objectives for a system. In the current example these causal networks are extracted from the Murray-Darling Basin Long Term Watering Plans (LTWPs), which sets environmental watering requirements (EWRs), ecological objectives, and long-term targets for key water-dependent plants, waterbirds, fish and ecological functions. From left to right, the columns here represent EWR codes, ecological objectives, specific goals, ecological values, and 5-year management targets. The network shown here is for a single gauge (412002) and Planning Unit, which is the scale at which EWR codes and ecological objectives are defined. The other levels are defined at larger spatial scales, but only those that apply to the EWRs present at this gauge are shown here. b) Subset of the causal network for gauge 419001 relevant to a selection of *lengthsampleenvobj* ecological objectives. The data are at the gauge scale for the first and second columns (EWRs and ecological objectives) and at the SDL unit scale for the third (ecological values). Thus, the final set of nodes contain information from other gauges that are not shown here for clarity. Colors represent the log of the relative change in condition from the baseline for the halving (panel a, scenario A) and doubling (panel b, scenario I) ‘climate’ scenarios. Large reductions in condition are dark purple, large improvements are dark green, and no change is white. See Appendix 3 *sec-causalnetwork-versions* for more detailed causal network diagrams. |

#### 4.5.1.2 Module standardization

Each response model will have a distinct set of outputs, reflecting the captured responses and the structure of the model. When run within HydroBOT, these outputs are cleaned and processed into standard, expected formats for further toolkit processing, and metadata is saved. This enables HydroBOT to provide a consistent, unified home for disparate response models. The outputs can be saved to disk or retained in-memory for interactive use, depending on the user’s needs. The outcomes of the response models are then processed by the Aggregator to enable outputs to be viewed at relevant scales (in time, space or value dimensions; [Figure 4](#fig-aggregation-dims)).

### 4.5.2 Aggregator

Results from the response models are typically very granular in many dimensions because the best response models operate near the scale of the processes being modelled. In many cases, those processes (e.g. fish breeding, crop planting) individually occur at small spatial and temporal scales. Note that this is the scale of the process itself, but that may be replicated over much larger scales. For example, fish may breed across the MDB but each female breeds in only one location at a given time. Moreover, outcomes from response models are typically at small value scales as well, e.g. capturing portions of the life cycle of fish species, rather than an overall outcome for all fish, or representing planting of particular crops, rather than overall agricultural output. The consequence is that there are potentially thousands of different modeled outcomes across time, locations and values. This plethora of outcomes must then be aggregated to extract meaning at larger scales and condense the information for digestibility in management decision making.

Aggregation condenses that information to scales that are useful for interpretation and planning. Depending on the use, the desired scale(s) may range from local, short-term responses of fine-grained outcomes to large spatial scales over longer time periods for high-level outcomes such as environmental condition ([Figure 4](#fig-aggregation-dims)). Thus, HydroBOT must achieve a robust, consistent aggregation approach along three dimensions (time, space, and value), while maintaining the ability to define those aggregation steps flexibly to meet the needs of the specific analysis. For example, the EWR tool assesses EWRs at individual gauges, but outcomes may be desired within areas defined for management purposes (e.g. Sustainable Diversion Limit (SDL) units) or at the basin scale. Likewise, multiple EWRs are required to meet ecological objectives, of which many are required for each ecological value or long-term target (see Appendix 1 for glossary **?@sec-glossary**). The HydroBOT Aggregator enables scaling from the hydrology at each gauge to results at any of these scales. HydroBOT provides a standard set of spatial units for aggregation but can accept any user-supplied polygons as the spatial aggregation units ([Figure 4](#fig-aggregation-dims)). Aggregation along the value dimension follows the causal network, which is supplied by HydroBOT for included tools, though the user can specify other causal relationships if required.

|  |
| --- |
| Figure 4: Aggregation along multiple dimensions. HydroBOT provides flexible capability to aggregate along spatial, temporal and value scales. Users can control the sequence of steps along these axes, with the capability to switch between axes at different steps. |

A flexible approach to aggregation is needed as specific dimensions (time, space and value) and units within those are best combined in different ways, depending on the meaning of the data and the use of the final outputs. The choices made in these aggregation steps are critical to ensure the final values are scientifically justified and capture the intended meaning. The Aggregator defines a sequence of aggregation steps to take with the data, specifying the dimensions along which to aggregate and the aggregation functions to apply at each step. Each step can be along any dimension and multiple functions can be assigned at each step. The sequence of aggregation steps can interleave the dimensions, e.g. aggregate along the value dimension to an intermediate level, then aggregate in space, then value again. For example, following [Figure 4](#fig-aggregation-dims), the user might want to aggregate from hydrologic indicator (EWR) to ecological objective to specific goal (e.g. species) at each gauge, then aggregate those gauges into a SDL unit to assess performance of each species over a larger area, followed by aggregating to ecological values (broad group, e.g. native fish, waterbirds, native vegetation, ecosystem function).

The Aggregator allows the user to choose any aggregation function at each aggregation step, reflecting the need to account for both the processes being aggregated and the outputs needed for management decision making. These aggregation functions should be considered carefully, as they have different meaning both for the processes being modeled and the interpretation of the outcomes (**?@fig-aggregation\_types**). For example, in some situations the user might want to know the rate at which EWRs pass across some area, or perhaps the average value of an abundance measure. In this case, a mean would be appropriate. In other situations, however, a single failure may be disproportionately important (e.g. ‘no loss’ requirements), or perhaps a single passing value is sufficient (e.g. bird breeding can occur anywhere in a catchment). These might be captured with a minimum and maximum, respectively. To address this need for flexible aggregation, HydroBOT provides a standard set of functions (e.g. mean, max, min, and spatially-weighted mean), but the user can also specify any other aggregation function, including custom-written functions.

The complexity of the potential aggregation sequences and functions highlights the importance of tracking the provenance of the final values to understand their meaning. Thus, the Aggregator saves the sequences and functions applied at each stage to a metadata file that also can serve as the aggregation parameters for repeatable runs. Moreover, the output dataset retains the full sequence and functions alongside each value, ensuring that values are always paired with their provenance and meaning.

### 4.5.3 Comparer

The HydroBOT Comparer is designed primarily to make comparisons between scenarios (typically hydrographs reflecting climate or climate adaptation in the examples that follow), allowing assessment and visualization of their differences. This component also provides generalized capacity to produce plots and other outputs such as tables using a consistent approach, even when not directly comparing scenarios. Comparisons are essential to assess the outcomes of various scenarios, e.g. the behaviour of the system under different climate regimes or with different adaptation options. The functions within the Comparer can be divided into two main categories, those for analysis and those for plotting. Although in some instances presenting the absolute outcome values can be useful across scenarios, explicitly calculating comparisons (e.g. the absolute or relative difference between scenarios) provides distinct advantages. Difficulty in accurately simulating a complex system means that comparisons between scenario outcomes can be more useful and accurate because any bias in the baseline assumptions applies to all outcomes and the focus moves from the total level to the change between scenarios *RL: Can use the same ref you cite in MER once you find one. GALEN*. The best method for comparing will vary depending on the quantities being compared and the intended use of the comparison, so several common default options (differences, relative change) have been developed along with the flexibility for the user to define alternatives.

The comparison functions provide the ability to choose a baseline level for comparison, which may be one of the scenarios, but also may be a reference dataset or a scalar value. For example, we might want to compare a set of outcomes for climate scenarios to a ‘no change’ climate scenario, to historical observations, or to a mean value. Output values are calculated relative to the defined baseline using either default functions for the absolute or relative difference, or any other user-supplied function. These functions can be applied to any dataframe, but a major advantage of HydroBOT is including them within the plotting functions. This allows the generation of comparison plots from raw Aggregator outputs without the need for subsequent data calculations by the user and avoids potential errors that arise from repeated or forgotten data transformations in an analysis workflow.

The plotting functions in the Comparer provide capability to present and visualize comparisons using standardized procedures for all outputs within a project. Different purposes require different sets of outputs; for instance, maps are particularly useful for visualizing geographic patterns, tables and graphs typically provide more precise ability to assess impacts on values, and timeseries plots are useful for visualizing climate trajectories. While potential visualizations and comparisons vary widely depending on the intended use, the Comparer standardizes data cleaning and processing for each plot, as well as aesthetics and plotting approaches. This standardization ensures consistency across plot types and through the project, ensuring values plotted are robust and interpretable. Key to this standardization is the internal data cleaning, which allows the raw outputs of the Aggregator and arguments for the comparisons to be provided to the plotting functions. By standardizing data cleaning within the Comparer, we avoid losing information or performing unsupported data manipulations and so ensure the quality and meaning of the outputs.

## 4.6 Demonstration scenarios

We demonstrate toolkit functionality for ecological values by using the existing EWR (Environmental Water Requirements) tool for the response model (see [Section 4.5.1](#sec-modules) for a description of the EWR module and its associated causal network). Our input data consists of hypothetical flow timeseries generated from historical hydrographs at 46 gauges. These gauges fall within the Lachlan, Macquarie–Castlereagh, and Namoi SDL units of the Murray-Darling Basin, Australia ([Figure 5](#fig-sdl-comparison)A). These SDL units have detailed LTWPs, and so well-specified EWRs and causal networks.

We develop a set of simple scenarios that capture two sorts of changes that may be commonly represented in management analyses. First, we consider scaled flow throughout the period of the hydrograph, representing overall increases or decreases in flow as might occur from large-scale climate patterns. Second, we consider short-duration additions to flow, representing periodic pulses of change as might happen from targeted interventions. Each scenario characterizes all water in the system including natural inflows, extraction, and release of environmental water, yielding a complete hydrograph ([Section 10.1](#sec-scenarios) shows a selected subset). These do not reflect realistic future scenarios but provide an avenue to test and illustrate the capabilities of HydroBOT.

To scaled flow, we apply nine flow multipliers, ranging from 0.5 to 2.0, to the historical hydrographs ([Table 2](#tbl-scenarios)). We refer to these as ‘climate’ scenarios, reflecting a common representation where entire hydrographs might shift to represent climate change, though these scenarios are not derived from climate models and do not represent hypothesized climate change. To achieve pulsed change for each of the ‘climate’ scenarios, four flow additions were applied including 1) no addition (baseline), 2) addition of 250 ML/d, 3) addition of 6500 ML/d, and 4) addition of 12000 ML/d ([Table 2](#tbl-scenarios)). These additional flows were added throughout the period of September to December. We refer to these scenarios as ‘climate adaptations’ because management options are often available in the form of altering water availability for short time periods through mechanisms like water releases, though the options here do not represent proposed actions.

For simplicity for this demonstration, we use arithmetic means for all aggregations except the very first, though we typically present results only after this first step. The first step gives achievement of EWR indicators from timing-dependent sub-indicators. In that case, we consider that, if an EWR is achieved at any time, it is achieved and so use a maximum for that aggregation. Though we use means throughout, their meaning changes depending on the level of aggregation. For simplicity, we assume the value of each outcome represents some ‘condition’, and the ‘condition’ at a particular level can be assessed as the mean of the conditions contributing to it. For example, the condition of ecological objectives might simply be the proportion of EWRs that contribute to each objective that are achieved. Then, the condition of ecological values, e.g. native fish, might be the mean condition over all the NF1..n objectives, despite those objectives depending on different numbers and sets of EWRs. This captures the idea that native fish condition improves when the life-cycle components captured by those objectives are met, whether it takes 1 or 10 EWRs to meet an objective.

# 5. Results and discussion

Here, we present example results produced using HydroBOT, the demonstration scenarios (described above in [Section 10.1](#sec-scenarios)) and the EWR tool for the response model. These results demonstrate the flexibility and robustness of the toolkit, which give it the capability to produce rigorous but interpretable and management-relevant results in a reproducible way (all code is available at https://github.com/MDBAuth/toolkit-demo-paper).

Comparison of the direct outputs from the response model (in this example, pass/fail of EWRs) without any aggregation along the value dimension gives an overview of hydrologic responses at different locations (e.g. gauges) or areas (e.g. SDL units) to different scenarios. Such an assessment can identify areas that may be particularly vulnerable or which potential ‘adaptation options’ might be most impactful across many outcomes. In [Figure 5](#fig-sdl-comparison), we present the EWR outcomes for three climate scenarios (A, E and I; 0.5x, no change and 2x the historical base level, respectively) and three adaptation options (1, 2 and 3; additions of 0, 250 and 6500 ML/d, respectively).

This example aggregates the EWR outputs directly to the SDL unit, and shows that they are likely to be affected by the changes in flow resulting from these changes in ‘climate’ and application of these ‘adaptation options’. However, the scenarios affect each SDL unit differently. The Lachlan has the highest proportion of EWRs achieved under all scenarios and has relatively consistent increases in EWR success with the adaptation options (though the options themselves are not even; 0 to 250 to 6500 ML/d). The Macquarie–Castlereagh is most sensitive to small increases in additional water, with large jumps between ‘adaptation’ options 1 and 2. In all situations, the ‘climate’ scenarios have less of an impact than the ‘adaptation’ scenarios, though neither is reflective of expected change in these dimensions. These outcomes are not yet directly linked to ecological values (priority ecological functions and environmental assets) and so each hydrologic indicator (EWR) at each gauge is represented with equal importance, whether it is required for all ecological objectives or just one.

|  |
| --- |
| Figure 5: (a) Map of the Murray-Darling basin, in eastern Australia. (b) Gauge locations within three SDL units (Macquarie–Castlereagh, Lachlan, and Namoi). Scenario comparison for different Sustainable Diversion Limit areas (a) emphasizing quantititive differences and (b) spatial patterns. The mean proportion of environmental watering requirements (EWRs) that are achieved under each scenario for each SDL unit are shown. Climate scenarios and adaptation options as in [Table 2](#tbl-scenarios). |

Response models may not themselves provide responses to all values of interest, and so we use the causal networks to link these outputs to additional objectives. In the EWR example, this linking is essential to address ecological outcomes and investigate how changes in flow affect values of interest. Thus, we can compare how native fish, native vegetation, waterbirds, other species, and ecological function targets are likely to respond to our hypothetical set of climates and adaptation scenarios. We are also able to compare the planning unit areas and the scenarios themselves. While there are many ways to display this information, [Figure 6](#fig-obj_in_groups) consolidates a large amount of information across scenarios, planning units, and multiple scales of objective. Each set of colours (and row of panels) represents outcomes at the larger ecological value level. Within those, the different shades represent single ecological objectives (e.g. WB1; Maintain the number and type of waterbird species), with the heights of these shades being the proportion of the constituent hydrologic requirements that are met (EWRs). Because the overall bar heights are then a sum of these proportions, we provide a dashed line as reference that shows the situation where all EWRs are met for each ecological objective. This reference is necessary because the number of EWRs contributing to each ecological objective, and the number of ecological objectives vary within each larger ecological value group and spatial unit.

This view provides a useful overview of the impact of different scenarios on the ecological values across space, while also retaining the ability to assess individual ecological objectives. For example, WB1 (Maintain the number and type of waterbird species) and WB2 (Increase total waterbird abundance across all functional groups) are unlikely to be achieved in the Macquarie–Castlereagh under any scenarios, while they are met in the baseline and 2x climate scenarios (2 and 3) in both the Lachlan and Namoi. At the larger ecological value scale, we can see that the effect of halving water (moving from the baseline (2) scenario to the 1 scenario) has a disproportionately greater impact on native fish than doubling water (the 3 scenario) in the Lachlan, while these shifts are more proportional in the Namoi.

|  |
| --- |
| Figure 6: Proportion of EWRs achieved for broader ecological values, including: priority ecological Function (EF), native fish (NF), native vegetation (NV), other species (OS) and waterbirds (WB) (Colours & rows). Columns are scaled to illustrate the number of ecological objectives that contribute to each broader ecological value, and shades of colours illustrate the proportion of EWRs that contribute to each ecological objective. I.e. if all EWR contributing to ecological objectives for each ecological values the columns would reach the dashed horizonal lines. This view provides an assessment of the performance of individual ecological objectives, as well as how those contribute to the overall sensitivity of the broader groups. Not all EWR are applicable in all locations and there are not equal number of EWR in each category. There are no other species objectives in the Namoi*GALEN/GEORGIA CHECK THIS ONCE RERUN THORUGH EWR TOOL*. This illustration includes three climate scenarios: A (0.5x), E (historical base level/no change), and I (2x); and three adaptation options: 1 (no adaptation), 2 (+250 ML/d), and 3 (+6500 ML/d) |

|  |
| --- |
| Figure 7: Ecological objectives aggregated to SDL unit spatial scale, illustrated here for three examples (Priority ecological Function 3, native fish 1, and waterbirds 4). Polygon colours indicate the proportion of EWR passed for each objective in each SDL unit. Aggregation as in [Figure 11](#fig-spatial-scaling). Includes three climate scenarios: A (0.5x), E (historical base level/no change), and I (2x); and three adaptation options: 1 (no adaptation), 2 (+250 ML/d), and 3 (+6500 ML/d), though not all combinations [see Appendix 12.1.1](#sec-map-versions). |

While we can see which scenarios are better or worse than others with bar graphs ([Figure 6](#fig-obj_in_groups)) or maps ([Figure 7](#fig-gauge-to-sdl-map)), they can make it difficult to accurately assess relative differences that may illuminate disproportionate impacts and thresholds for when adaption options could be most effectively adopted. In general, this requires quantitative x-axes which would typically be referenced to some baseline. Here, we can represent both our ‘climate’ and ‘adaptation’ scenarios on a single axis by quantifying the difference in mean flow from the baseline condition (historical hydrograph, E1) using the baselining functionality provided by HydroBOT (see [Section 10.2](#sec-baselining)). Plotting the proportion of EWRs achieved for each of the native fish ecological objectives shows how that EWR achievement is related to changes in the overall levels of flow. One striking feature of [Figure 8](#fig-difference-baseline) is that the lines do not smoothly increase; EWR achievement is not simply a direct relationship to flow volumes. Instead, the timing matters. There are particularly steep slopes between the ‘1’ (baseline) and ‘2’ (+250 ML/d) adaptations. The application of water at those particular times is thus disproportionately impactful, yielding greater outcomes even when they have less water than the baseline adaptations (‘1s’) for climate scenarios ‘A’ and ‘E’. Such analyses can identify highly efficacious points to add water to the system.

|  |
| --- |
| Figure 8: Quantitative scenario comparison for native fish objectives for three SDL units. Scenarios defined here as the difference in mean flow (GL) from the baseline historical scenario (E1). A key message here is that the lines are not smooth – EWR achievement is not simply a function of the amount of water. Instead, the ‘adaptation’ options, particularly the ‘2’ scenarios of adding 250 ML/d of water, have disproportionately large impact (steep slopes). Shading indicates different ecological objectives within the native fish ecological value. This illustration includes three climate scenarios: A (0.5x), E (historical base level/no change), and I (2x) and three adaptation options: 1 (no adaptation), 2 (+250 ML/d), and 3 (+6500 ML/d). |

With a large number of scenarios, we can better characterise the way values respond along different axes of change. Here, we use smoothed fits to show trends in the performance of ecological objectives (and groups thereof) in response to multiplicative flow changes (‘climate’ scenarios) and pulsed flow additions (‘adaptations’) ([Figure 9](#fig-smooth-climate-adapt), [Figure 16](#fig-smooth-all)). We use baselining (see [Section 10.2](#sec-baselining)) to show the relative change on both the x (flow) and y (response) axes. Thus, any slope other than 1:1 represents a disproportionate shift in condition. For example, other species and native vegetation show steeper increases with increasing flow relative to baseline than decreasing flow, while native fish respond disproportionately to flow change across the range considered here. Moreover, because we use smoothed fits, we can identify thresholds or areas of flow change that are disproportionately more important (small changes in flow can yield large shifts in condition).

*RL: This is confusing without any discussion of what Fig 20 is. Do we need it here or can this para move to the supplement too?* *GH: Hmm. I seem to have made major changes to these figs but not the text. Fig 20 used to be here, and fig 13 used to be different. The idea was that panel C could be used to accentuate the main point of Fig 20, that the adaptations here have a way bigger impact than the climate. So, yes, we should move most of this to the appendix. If we keep C, we should keep a sentence about it. And maybe we should decide we don’t need C at all and cut it back off the figure.* In [Figure 16](#fig-smooth-all) the y-axis scales are much larger than the x, and so proportionality is difficult to assess. However, this figure makes clear the nonlinear nature of the response to flow shift, particularly for the no adaptation scenario (1), as well as changes in sensitivity between adaptation options. The response to adapatation options (differences between the lines in [Figure 16](#fig-smooth-all), slope of lines in [Figure 9](#fig-smooth-climate-adapt) panel b) tends to be much stronger than the response to climate (slopes of the lines in [Figure 16](#fig-smooth-all) and [Figure 9](#fig-smooth-climate-adapt) panel a), even though the shift in total water might be lower. Moreover, the adaptation options are not as sensitive to the climate scenarios; adaptation options other than ‘1’ (no adaptation) show less change with climate in [Figure 16](#fig-smooth-all). Thus, these ‘adaptations’ are increasing resilience to those holistic flow changes. Although these scenarios do not represent true adaptation options or climate scenarios, this shows that such changes to resilience are possible with targeted interventions, and HydroBOT provides the capacity to investigate them.

The responses seen in [Figure 16](#fig-smooth-all) and [Figure 9](#fig-smooth-climate-adapt) vary between SDL unit (columns) and ecological values, indicating different sensitivity to both ‘climate’ and ‘adaptation’ flow conditions across space and among ecological groupings. For example, although both the Lachlan and Macquarie-Castlereagh have no successful waterbird outcomes under baseline conditions, targeted flow conditions have a greater ability to increase waterbird success in the Lachlan.

|  |
| --- |
| Figure 9: (a) Smoothed fits of shifts in condition (proportion of EWR achieved relative to the scenario with no climate or adaptation changes [E1]) of broad ecological values to relative changes in water availability (also relative to E1). All climate scenarios (without adaptation; adaptation = 1) are included. (b) The full combination of scenarios (all climate and adaptations options) are shown for the example of native fish in the Namoi. This illustrates the larger impact of the ‘adaptation options’ used here compared to the ‘climate’ shifts. Fitted lines are loess smooths. Separate lines are fit to each ecological value. Columns are SDL units. Y-axes are restricted to better visualise the majoity of the data; thus, a small amount of data is not plotted, but is accounted for in the loess fits. The full set of SDL units and groups as in c shown in Appendix [Figure 16](#fig-smooth-all). |

Scenarios will often be defined along more than one axis. In the demonstration here, we define both a flow multiplication axis as a proxy for ‘climate’ shifts, and a flow addition as a proxy for ‘adaptations’. Moreover, it will often be the case that values will respond to several different aspects of the flow regime, for example the mean flow and the flow variance or the return interval of floods or low-flow periods. Heatmaps or contour surfaces provide a powerful tool for visualising these interacting responses ([Figure 10](#fig-contour), [Figure 17](#fig-heatmap)). By examining multiple driver axes, such plots can highlight important interactions and identify where thresholds occur and where change along one axis (e.g. adaptations) can mitigate change along the other (e.g. climate shifts). This ability to identify the axes that provide resilience or sensitivity to changes in others will be critically important for management uses of this toolkit targeting climate change adaptation. These assessments must also take spatial and value differences into account; the different ecological values and SDL units show very different patterns in how they respond to the interaction between the ‘climate’ and ‘adaptation’ scenarios ([Figure 10](#fig-contour)).

Perhaps most usefully, if scenarios are carefully developed to explore the range of potential change in the system, then the heatmap represents the response surface of values over this range. Developing the response surface can be iterative, with subsequent scenarios developed to increase resolution near areas of rapid change in outcome. This approach differs from the usual approach of assessing only a small set of scenarios targeting specific proposed actions or environmental conditions. Instead, by covering the range to generate a response surface, it can provide hypotheses about how proposed scenarios might perform and assess how sensitive the outcomes are to uncertainty in single scenarios. Specific proposed scenarios could of course be assessed as well and mapped onto the surface. For example, we might think of the flow multipliers here as defining the entire range of plausible climate shifts, and the addition scenarios as the entire potential range of additions. Then, specific proposed adaptations or climate sequences could be mapped onto these heatmaps at particular points. The reverse is also possible; identification of areas of high sensitivity in the heatmaps could be used to choose a set of potential adaptations designed to increase climate resilience. This idea extends to additional dimensions, which can then be collapsed in various ways (e.g. via principal component analysis, or by identifying dominant hydrometrics) to visualise with 2d heatmaps even when the response surface itself is best defined and studied in more dimensions.

|  |
| --- |
| Figure 10: Condition results visualised as a surface with the two scenario definitions on the axes. This approach allows visualising changes in outcome as the result of multiple axes on which scenarios might differ. These axes might be different aspects of scenario creation (as here), or they might be different axes describing the outcome of scenarios (e.g. two different hydrometrics such as mean flow and flow variance) |

Causal networks can be used to show outcomes in addition to the causal relationships themselves. This sort of visualisation allows identification of the holistic consequences of the different scenarios across the complex and interrelated sets of values, as well as identifying nodes that are unusually resilient or sensitive to change (**?@fig-network-subset**). For example, NF7 is little changed across the range of climate scenarios considered, while EF5 shows greater sensitivity. The network structure can also identify critical nodes that have outsize consequences for other nodes in the network. Such nodes are critically important for outcomes, and so should be investigated to determine whether they capture key aspects of the system or are instead an artifact of model structure (possibly resulting from the way the network was developed) that may bias results. If their sensitivity does in fact capture the system, these nodes may be ideal targets for management intervention. The edges also identify the dependence between values, where we can see that different environmental outcomes depend on not only different hydrologic indicators (EWRs), but also different numbers of those indicators.

# 6. Values across themes

radar plot?

# 7. Implications and conclusions

### 7.0.1 Outline

What we achieved:

* a toolkit that provides a consistent, scientifically robust and repeatable capacity to model responses to flow, with management-relevant outputs
* co-design and development between scientists and managers

What this provides:

* new ability to assess the wide suite of values the MDBA is mandated to manage for.
  + moving beyond hydrology
* Clear (or at least open) modelling to build trust
* ability to assess a range of scenarios, including climate and climate adaptation

Advantages:

* Identification of areas of resilience and sensitivity.
* Uncertainty, understanding parameter space.
* Modularity
* adapt best models for responses into unified model and target management relevance
* agnostic to scenario
* causal networks

What cool software things do we do?

* Modularity in architecture
* ability to wrap models in multiple languages
* flexibility- e.g. ability to aggregate differently
* consistent comparison, with data provenance
* metadata and use as parameters for repeatability

### 7.0.2 From elsewhere- incorporate

*lots of this can be the backbone of the big-picture points above*

Large-scale natural resource management requires the capacity to make decisions relating to multiple spatial, temporal, and value dimensions, and is most successful when multiple scales within those dimensions are considered (Moore 2021). HydroBOT is a framework to navigate those dimensions for water-dependent social, economic, environmental, and cultural values. Combining such disparate information in a standard and comparable manner can illuminate synergies and trade-offs, which could be critical for final judgement. Our example concerns the ecological values of the Murray-Darling Basin; however, the framework is equally applicable to social, economic, and cultural values and the toolkit itself has a modular design to incorporate such response models. Our framework drives a consistent approach to processing, scaling, analyses, and visualising outcomes, with the flexibility built-in for most end uses. Thus, HydroBOT provides a good avenue for informed decision-making for water management in the Murray Darling Basin.

We expand upon the utility of an existing diver-indicator model by linking its indicators *EF: indicators same as values here I suppose? I think cos this paper is quite dense and tackles a lot of different types of variables and data visualisation techniques, it’s really important to use consistent terms.* to objectives. This increases the transparency in the causal relationships that underpin the model and builds understanding and trust in the outcomes.

concerning which adaptation options should be implemented under changing climate by explicitly modelling effects on both ecological and other values.

The development of HydroBOT is necessary to better understand the impacts of climate change and the different adaptation options in response to climate change, on water-dependent social, economic, environmental and cultural values. It incorporates new and existing information, knowledge and models to enable transparent, repeatable assessments of impacts and adaptation to future climates. In summary, the toolkit ingests scenarios (for example, climate or flow timeseries), feeds them to a response model (such as the EWR tool), reports outcomes and enables comparisons among scenarios. It uses the links in the response models for visualisation of the complex inter-relationships between water-dependent outcomes. This aids transparency and improves communication of the outputs.

*GEORGIA (priority ecological functions and environmental assets)*

# 8. References

Allaire, J. J., Charles Teague, Carlos Scheidegger, Yihui Xie, and Christophe Dervieux. 2022. *Quarto*. <https://doi.org/10.5281/zenodo.5960048>.

Harrison, Paula A., Kate Beauchamp, Joe Cooper, Ian Dickie, Alice Fitch, Richard Gooday, Michael Hollaway, et al. 2023. “An Adaptable Integrated Modelling Platform to Support Rapidly Evolving Agricultural and Environmental Policy.” *Environmental Modelling and Software* 169 (November): 105821. <https://doi.org/10.1016/j.envsoft.2023.105821>.

Hart, Barry T., Nick R. Bond, Neil Byron, Carmel A. Pollino, and Michael J. Stewardson. 2021. “Chapter 1 - Introduction to the Murraydarling Basin System, Australia.” In, edited by Barry T. Hart, Nick R. Bond, Neil Byron, Carmel A. Pollino, and Michael J. Stewardson, 1:1–17. Ecohydrology from Catchment to Coast. Elsevier. <https://doi.org/10.1016/B978-0-12-818152-2.00001-2>.

Kaye-Blake, Bill, Chris Schilling, Chris Nixon, and Killian Destremau. 2014. “Water Management in New Zealand : A Road Map for Understanding Water Value.” *NZIER Working Paper*, NZIER Working Paper, March. <https://ideas.repec.org//p/ris/nzierw/2014_001.html>.

Lester, Rebecca E. 2019. “Wise Use: Using Ecological Models to Understand and Manage Aquatic Ecosystems.” *Marine and Freshwater Research*. <https://doi.org/10.1071/MF18402>.

Lobegeiger, Jaye, and Andrea Prior. 2022. “Border Rivers and Moonie Long-Term Watering Plan.”

Moore, Scott. 2021. “Toward Effective River Basin Management (RBM): The Politics of Cooperation, Sustainability, and Collaboration in the Delaware River Basin.” *Journal of Environmental Management* 298 (November): 113421. <https://doi.org/10.1016/j.jenvman.2021.113421>.

Murray-Darling Basin Authority. 2011. “Delivering A Healthy Working Basin.” *Basin Plan*, November.

Peeters, Luk J. M., Kate L. Holland, Cameron Huddlestone-Holmes, and Andrew J. Boulton. 2022. “A Spatial Causal Network Approach for Multi-Stressor Risk Analysis and Mapping for Environmental Impact Assessments.” *Science of The Total Environment* 802 (January): 149845. <https://doi.org/10.1016/j.scitotenv.2021.149845>.

Planning, NSW Department of, and Environment. 2020. “Murrumbidgee Surface Water Resource Plan SW9 Water Resource Plan Area.”

Stern, Charles V, and Pervaze A Sheikh. 2019. “Management of the Colorado River: Water Allocations, Drought, and the Federal Role.”

Ushey, Kevin, J. J. Allaire, and Yuan Tang. 2023. *Reticulate: Interface to ’Python’*.

# 9. Appendix 1

## 9.1 Glossary

### 9.1.1 Component overview

The HydroBOT architecture comprises three major components, the Controller, Aggregator and Comparer, that receive data and information from several input sources, including input scenarios, hydrological modelling, spatial data, causal networks and response models ([Table 1](#tbl-components)). Here we detail some specific details about each of these components and other input sources.

Table 1: Components of HydroBOT architecture

| General Toolkit components | General component definitions | Specific components used in our example | Details of Specific components |
| --- | --- | --- | --- |
| Input data | Hydrologic data (timeseries). Typically representing multiple scenarios, e.g. climate and climate adaptations. May include other inputs as needed by response models. | Modified historical hydrographs to represent hypothetical climate change and adaptations | Daily flow rates for NUMBER GAUGES for 15 scenarios |
| Controller | Interface between input data, response model, and other toolkit components. Sets up run(s). | Sets up links to data and parameters for EWR tool | NA |
| Response models | A model that contains are set of causal relationships between a driver and responses, e.g. social, cultural, environmental, or economic values. | Specific response model = EWR tool | The EWR tool holds databases of the EWRs required to meet the environmental objectives of the basin, which protect or enhance environmental assets that are valued based on ecological significance. |
| Aggregator | Aggregates response model results to scales across the dimensions of time, space, and theme. | Response model sets the base scale for aggregation. EWR tool assesses hydrologic indicators (theme) at gauges (space) and internally aggregates over time. | The spatial dimension consists of gauges nested within planning units within the basin. |
| NA | NA | NA | The theme dimension consists of multiple EWRs (hydrologic indicators) that apply to environmental objectives (many-to-one), which apply to other levels in the causal network (e.g. Objectives, Targets, and Long-term targets). |
| Comparer | Compares scenarios (typically) or other groupings. Provides standard outputs including comparison methods, plots, and tables. | Comparison of environmental values at various theme scales for the example climate and adaptation scenarios | - |
| Causal networks | Describe causal relationships between values. | Long Term Water Plan (LTWP) | Plans required of Basin States by the Murray-Darling Basin Plan. Long term water plans give effect to the Basin-wide Environmental Watering Strategy relevant for each river system and will guide the management of water over the longer term. These plans will identify the environmental assets that are dependent on water for their persistence, and match that need to the water available to be managed for or delivered to them. The plan will set objectives, targets and watering requirements for key plants, waterbirds, fish and ecosystem functions. DPIE is responsible for the development of nine plans for river catchments across NSW, with objectives for five, 10 and 20-year timeframes. |

### 9.1.2 Aggregation

Aggregation functions Aggregation functions determine outcomes and reflect processes or values. Aggregating a set of outcomes yields a different outcome depending on the function used. Thus, function choice should be considered carefully and reflect the processes involved or management goals. HydroBOT provides default functions for the mean, compensating (e.g. max), limiting (min), and spatially-weighted versions of the same. It also provides the ability for the user to define any desired function, allowing for more complex situations.

#### 9.1.2.1 Spatial aggregation

Maps allow large-scale visualization of ecological objectives under various scenarios. These visualisations can be quite important for communications and quickly grasping how outcomes aggregate spatially ([Figure 11](#fig-spatial-scaling)) and spatial patterns in the data ([Figure 7](#fig-gauge-to-sdl-map)). Management targets are often defined spatially, and in the case of EWRs, they are defined in local Planning Units (which are constituents of the SDL units defined above), where outcomes may depend on hydrographs at one or more gauges. Representing the outcomes as maps can provide intuitive assessment of the condition of values across space and whether different spatial scales or spatial locations are responding differently to scenarios. Moreover, the ecology (in this example) or other processes might themselves be large-scale, and so capturing the condition over a large area is a better descriptor of the true outcome than assessing each specific location separately. For example, objective EF3 is “Provide movement and dispersal opportunities for water dependent biota to complete lifecycles and disperse into new habitats within catchments”, and so necessarily incorporates a spatial dimension. Particularly in these situations, the aggregation method should be considered carefully – for movement opportunities to succeed, perhaps the success of the SDL unit should be determined by the lowest value at a gauge if it represents a loss of connectivity. In contrast, for WB4: Increase opportunities for colonial waterbird breeding, it might be sufficient if a single site within the SDL unit provides those opportunities.

|  |
| --- |
| Figure 11: Environmental water requirements (EWRs) are defined by hydrographs at gauges (a), and apply to Planning Units (b). The outcomes at these scales of definition can then be aggregated to larger spatial areas, such as SDL units (c), which we do here with an area-weighted mean for the example EWR NF1 in the baseline climate and adaptation scenario (E1, simple historical hydrograph with no climate or adaptation changes). |

#### 9.1.2.2 Temporal aggregation

#### 9.1.2.3 Objective aggregations

## 9.2 Response models

#### 9.2.0.1 EWRs

*GEORGIA Need a table of their English definitions, the full EWR table isn’t helpful*

**?(caption)**

# 10. Appendix 2

## 10.1 Scenarios

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| (a) Climate   | Climate code | Flow multiplier | | --- | --- | | A | 0.50 | | B | 0.67 | | C | 0.80 | | D | 0.91 | | E | 1.00 | | F | 1.10 | | G | 1.20 | | H | 1.50 | | I | 2.00 | | (b) Adaptation   | Adaptation code | Flow addition (ML/d) | | --- | --- | | 1 | 0 | | 2 | 250 | | 3 | 6500 | | 4 | 12000 | |

Table 2: Demonstration scenarios are a factorial combination of ‘climate’ (scaled flow) and ‘adaptation’ (pulsed additions). Climate scenarios included in this demonstration were produced by applying a flow multiplier to historical flows. Adaptation options were applied to each climate scenario with additional flows added throughout the period of September to December. For ease of display in some figures, we provide alpha codes for the ‘climate’ changes and numeric codes for the ‘adaptations’. Mostly frequently, we present the ‘A’ (0.5 x historical flow), ‘E’ (1.0 x historical flow and ‘I’ (2.0 x historical flow) scenarios, with adaptions 1 (no flow addition), 2 (addition of 250 ML/d) and 3 (addition of 6500 ML/d).

Scale for y is already present.  
Adding another scale for y, which will replace the existing scale.

|  |
| --- |
| Figure 12: Hydrographs for two example gauges (represented by colour) with ‘climate’ scenarios on rows and ‘adaptation’ scenarios as columns. Scenario codes as in [Table 2](#tbl-scenarios). |

## 10.2 Baselined hydrographs

|  |
| --- |
| Figure 13: Change in flow relative to the baseline scenario. These are flat lines because the relativisation occurs at each timepoint.The baseline scenario is represented by the observed historical data with no climate change or adaptaion options applied (scenario E1) |

|  |
| --- |
| Figure 14: Change in flow volume compared to the baseline scenario. This comparison is done using the difference, and so represents the flow additions. The baseline scenario is represented by the observed historical data with no climate change or adaptaion options applied (scenario E1) |

# 11. Appendix 3

## 11.1 Additional causal netwok plots

#### 11.1.0.1

# 12. Appendix 4

## 12.1 Additional comparison plots

### 12.1.1 Map aggregation

The goal here is to show the values at the sdl scale, but also to show how the gauges aggregate to that scale. In the text, [Figure 7](#fig-gauge-to-sdl-map) has a few examples of SDL-scaled outcomes, while [Figure 11](#fig-spatial-scaling) has the scaling for one. [Figure 15](#fig-gauge-to-sdl-map-all) shows the SDL outcomes for all ecological objectives contributing the priority ecological function, along with the constituent gauges.

|  |
| --- |
| Figure 15: all panels version. |

The full set of SDL units and ecological values

|  |
| --- |
| Figure 16: Smoothed fits to assess change in performance across the ‘climate’ scenarios. Points are individual ecological objectives, fitted lines are loess smooths. Separate fits are done for each adaptation option, and so differences between lines of different colors represents the impact of those adaptations. Rows are ecological values groupings, columns are SDL units. Note the very different scale of the y-axis. |

### 12.1.2 Heatmaps

|  |
| --- |
| Figure 17: Condition results visualised as a heatmap with the two scenario definitions on the axes. This approach allows visualising changes in outcome as the result of multiple axes on which scenarios might differ. These axes might be different aspects of scenario creation (as here), or they might be different axes describing the outcome of scenarios (e.g. two different hydrometrics such as mean flow and flow variance) |

### 12.1.3 Causal networks

The causal networks shown in the text have been subset for readability, and baselined. Here, we have the ones they’re based on. *I’ve turned these off for word because diagrammer hates word and I don’t have the energy to do a workaround*.

|  |
| --- |
| Figure 18: Subset of the causal network for gauge 419001 for all ecological objectives. The data is at the gauge scale for the first and second columns (EWRs and ecological objectives) and at the SDL unit scale for the third (ecological values). Thus, the final set of nodes contain information from other gauges that are not shown here for clarity. Colors represent the condition for the halving (a), baseline (b), and doubling (c) ‘climate’ scenario. Low condition is dark blue, high is light yellow. |

The same set of nodes as in the text, but the actual condition values, rather than relativized.

|  |
| --- |
| Figure 19: Causal networks as in **?@fig-network-subset**, but with the raw condition values for the halved (a), baseline (b), and doubled (c) scenarios instead of relativized to the baseline. |

# 13. Journal info

## 13.1 Environmental Modelling & Software

• Author names and affiliations. Where the family name may be ambiguous (e.g., a double name), please indicate this clearly. Present the authors’ affiliation addresses (where the actual work was done) below the names. Indicate all affiliations with a lower-case superscript letter immediately after the author’s name and in front of the appropriate address. Provide the full postal address of each affiliation, including the country name and, if available, the e-mail address of each author.

• Corresponding author. Clearly indicate who will handle correspondence at all stages of refereeing and publication, also post-publication. Ensure that telephone numbers (with country and area code) are provided in addition to the e-mail address and the complete postal address. Contact details must be kept up to date by the corresponding author.

• Present/permanent address. If an author has moved since the work described in the article was done, or was visiting at the time, a ‘Present address’ (or ‘Permanent address’) may be indicated as a footnote to that author’s name. The address at which the author actually did the work must be retained as the main, affiliation address. Superscript Arabic numerals are used for such footnotes.

### 13.1.1 Authorship

Authorship should be limited to those who have made a significant contribution to the conception, design, execution, or interpretation of the reported study. All those who have made significant contributions should be listed as co-authors. Where there are others who have participated in certain substantive aspects of the research project, they should be acknowledged or listed as contributors. Acknowledgement of the contributions of authors is encouraged (see Acknowledgements section below). The corresponding author should ensure that all appropriate co-authors and no inappropriate co-authors are included on the paper, and that all co-authors have seen and approved the final version of the paper and have agreed to its submission for publication.

Title, Authors, Affiliations and Contact details

### 13.1.2 Abbreviations

Define abbreviations that are not standard in this field in a footnote to be placed on the first page of the article. Such abbreviations that are unavoidable in the abstract must be defined at their first mention there, as well as in the footnote. Ensure consistency of abbreviations throughout the article.

## 13.2 Separate files

### 13.2.1 Highlights

Highlights are optional yet highly encouraged for this journal, as they increase the discoverability of your article via search engines. They consist of a short collection of bullet points that capture the novel results of your research as well as new methods that were used during the study (if any). Please have a look at the examples here: example Highlights.

Highlights should be submitted in a separate editable file in the online submission system. Please use ‘Highlights’ in the file name and include 3 to 5 bullet points (maximum 85 characters, including spaces, per bullet point).

Highlights are mandatory for this journal. They consist of a short collection of bullet points that convey the core findings of the article and should be submitted in a separate file in the online submission system. Please use ‘Highlights’ in the file name and include 3 to 5 bullet points (maximum 85 characters, including spaces, per bullet point). See https://www.elsevier.com/highlights for examples.

### 13.2.2 Graphical abstract

A Graphical abstract is optional and should summarize the contents of the article in a concise, pictorial form designed to capture the attention of a wide readership online. Authors must provide images that clearly represent the work described in the article. Graphical abstracts should be submitted as a separate file in the online submission system. Image size: Please provide an image with a minimum of 531 × 1328 pixels (h × w) or proportionally more. The image should be readable at a size of 5 × 13 cm using a regular screen resolution of 96 dpi. Preferred file types: TIFF, EPS, PDF or MS Office files. See https://www.elsevier.com/graphicalabstracts for examples.

### 13.2.3 Abstract (not included in section numbering)

A concise and factual abstract is required, with a restriction of 150 words. The abstract should state briefly the purpose of the research, the principal results and major conclusions. An abstract is often presented separately from the article, so it must be able to stand alone. For this reason, References should be avoided, but if essential, then cite the author(s) and year(s). Also, non-standard or uncommon abbreviations should be avoided, but if essential they must be defined at their first mention in the abstract itself.

### 13.2.4 Keywords

Immediately after the abstract, provide a maximum of 6 keywords, using American spelling and avoiding general and plural terms and multiple concepts (avoid, for example, ‘and’, ‘of’). Be sparing with abbreviations: only abbreviations firmly established in the field may be eligible. These keywords will be used for indexing purposes.

### 13.2.5 Software and/or data availability

Most EMS papers should include a software/data availability section containing as much of the following information as possible: name of software or dataset, developer and contact information, year first available, hardware required, software required, availability and cost. Also for software: program language, program size; for data: form of repository (database, files, spreadsheet), size of archive, access form. Note that “Contact the author” is not acceptable for software or data access. Please use online data and software storage and retrieval systems such as GitHub, BitBucket, FigShare, HydroShare or others to make your data and software readily available. Links to commercial software and data access web sites are also acceptable.

When a software component is an essential part of the paper, authors should be prepared to make it available to reviewers during the review process. To preserve the anonymity of reviewers, the authors should make the software available for a download, protecting it if needed by a password that is communicated to the editors.