Australian Government

BIOREGIONAL ASSESSMENTS

PROVIDING SCIENTIFIC WATER RESOURCE
INFORMATION ASSOCIATED WITH COAL
SEAM GAS AND LARGE COAL MINES

# Receptor impact modelling

## Submethodology M08 from the Bioregional Assessment Technical Programme

2018

A scientific collaboration between the Department of the Environment and Energy,
Bureau of Meteorology, CSIRO and Geoscience Australia

## The Bioregional Assessment Programme

The Bioregional Assessment Programme is a transparent and accessible programme of baseline assessments that increase the available science for decision making associated with coal seam gas and large coal mines. A bioregional assessment is a scientific analysis of the ecology, hydrology, geology and hydrogeology of a bioregion with explicit assessment of the potential impacts of coal seam gas and large coal mining development on water resources. This Programme draws on the best available scientific information and knowledge from many sources, including government, industry and regional communities, to produce bioregional assessments that are independent, scientifically robust, and relevant and meaningful at a regional scale.

The Programme is funded by the Australian Government Department of the Environment and Energy. The Department of the Environment and Energy, Bureau of Meteorology, CSIRO and Geoscience Australia are collaborating to undertake bioregional assessments. For more information, visit http://www.bioregionalassessments.gov.au.

## Department of the Environment and Energy

The Office of Water Science, within the Australian Government Department of the Environment and Energy, is strengthening the regulation of coal seam gas and large coal mining development by ensuring that future decisions are informed by substantially improved science and independent expert advice about the potential water related impacts of those developments. For more information, visit https://www.environment.gov.au/water/coal-and-coal-seam-gas/office-of-water-science.

## Bureau of Meteorology

The Bureau of Meteorology is Australia’s national weather, climate and water agency. Under the *Water Act 2007*, the Bureau is responsible for compiling and disseminating Australia's water information. The Bureau is committed to increasing access to water information to support informed decision making about the management of water resources. For more information, visit http://www.bom.gov.au/water/.

## CSIRO

Australia is founding its future on science and innovation. Its national science agency, CSIRO, is a powerhouse of ideas, technologies and skills for building prosperity, growth, health and sustainability. It serves governments, industries, business and communities across the nation. For more information, visit http://www.csiro.au.

## Geoscience Australia

Geoscience Australia is Australia’s national geoscience agency and exists to apply geoscience to Australia’s most important challenges. Geoscience Australia provides geoscientific advice and information to the Australian Government to support current priorities. These include contributing to responsible resource development; cleaner and low emission energy technologies; community safety; and improving marine planning and protection. The outcome of Geoscience Australia’s work is an enhanced potential for the Australian community to obtain economic, social and environmental benefits through the application of first class research and information. For more information, visit http://www.ga.gov.au.

**ISBN-PDF 978-1-925315-68-4**

## Citation

Hosack GR, Ickowicz A, Hayes KR, Dambacher JM, Barry SA and Henderson BL (2018) Receptor impact modelling. Submethodology M08 from the Bioregional Assessment Technical Programme. Department of the Environment and Energy, Bureau of Meteorology, CSIRO and Geoscience Australia, Australia. http://data.bioregionalassessments.gov.au/submethodology/M08.

Authorship is listed in relative order of contribution.

## Copyright

## Disclaimer

## Cover photograph

Wards River, NSW, 10 December, 2013

Credit: Heinz Buettikofer, CSIRO

Australian Government
Department of the Environment and Energy
Bureau of Meteorology
Geoscience Australia

CSIRO

v20180427

# Executive summary

This submethodology describes the process for assessing impact on, and risk to ecological assets in bioregional assessments (BA) due to coal resource development.

The potential impact on and risk to ecosystems and water-dependent ecological assets due to coal resource development is quantitatively assessed using receptor impact models. They translate predicted changes in hydrology at specified points in time into a distribution of ecological outcomes that may arise from those changes. A receptor impact model predicts the response of a receptor impact variable (an ecological indicator, such as annual mean percent canopy cover of woody riparian vegetation), to one or more hydrological response variables (for example, *dmax*, maximum groundwater drawdown due to additional coal resource development). Receptor impact models in a bioregion or subregion are developed at landscape class level, and any ecological implications for a water-dependent asset subsequently consider predictions of receptor impact variables for landscape classes that intersect that asset. For BA purposes, a landscape class is defined as an ecosystem with characteristics that are expected to respond similarly to changes in groundwater and/or surface water due to coal resource development.

Economic water-dependent assets are confined to groundwater and surface water management zones or areas and comprise specific water access entitlements or rights and other water supply features or infrastructures. The scope of BA precludes detailed analysis of economic impacts, and is limited to the assessment of changes to the availability and reliability of the relevant groundwater or surface water. Potential impacts to economic assets are tied more directly to hydrological response variables and specific management thresholds (e.g. cease-to-pump rules), than for ecological assets. As a result, receptor impact models are not considered for economic assets.

Water-dependent sociocultural assets are classified into ‘Heritage site’, ‘Indigenous site’ or ‘Recreation area’ classes. Receptor impact modelling is not considered directly for sociocultural assets and the impact and risk is limited to characterising the hydrological change that may be experienced by those assets. However, a receptor impact model for ecological assets may also play an important role in assessing any change for those sociocultural assets that are ecological in character.

This submethodology is constructed to address the complexity of ecological systems, propagate and communicate the uncertainty, and synthesise the current state of knowledge. The specified time points of interest include a reference assessment year in 2012, a short-term future assessment year in 2042 and a long-term future assessment year in 2102. These are selected to represent the current state of the ecosystem, the ecosystem close to peak production and the ecosystem following coal resource development. The receptor impact model construction process depends crucially on the definition of a landscape class. Each bioregion or subregion has multiple landscape classes, and only those that may be potentially impacted from hydrological change due to additional coal resource development are considered for receptor impact modelling.

For a given landscape class, a qualitative mathematical modelling approach is used with independent ecological experts to identify the key ecosystem components and dependencies, and importantly, their link to groundwater and surface water variables. These hydrological variables are then translated into hydrological response variables that can be estimated by the groundwater and surface water hydrological models. The qualitative models are used to assess potential candidate receptor impact variables and hydrological response variables to consider within a receptor impact model.

A suite of hydrological scenarios based on plausible combinations of hydrological response variables is developed for each receptor impact model. These scenarios are presented to independent ecological experts as part of an expert elicitation process, and the distribution of receptor impact variable under each scenario elicited. A statistical model is constructed based on the experts' responses to estimate the relationship between the receptor impact variable and the relevant hydrological response variables, and to enable predictions to be made for combinations of hydrological response variables not considered in the elicitation. Within a landscape class, probabilistic predictions from the model can then be made at specific locations based on the modelled changes in those hydrological response variables at that location in different time periods. These predictions may be extended to regions, such as reaches of river, by applying the receptor impact models at different locations and using the changes in hydrological response variables at each of those locations. It is important to note that predictions of a receptor impact variable at a specific location is a prediction of the likely response across the landscape class for the change in hydrology at that location.

The receptor impact model for a landscape class therefore permits the prediction of how cumulative hydrological changes may change a receptor impact variable at key time points in the BA futures. This represents an essential component of the impact and risk analysis (product 3-4), and is a key vehicle by which impacts on and risks to water-dependent ecosystems and assets are assessed. Distributions of change for hydrological response variables and receptor impact variables are summarised in BAs by a limited series of percentiles (or quantiles), nominally 5 percent increments between the 5th and 95th percentiles, for different futures (baseline and coal resource development pathway (CRDP)) and the specified time frames of interest. For a specific water-dependent asset, predictions of receptor impact variables represent an important line of evidence, but need to be considered in conjunction with the qualitative mathematical models, the wider suite of hydrological changes, detailed information about the presence and condition of that asset, and other more local information.

# Contents

# Figures

# Tables

# Contributors to the Technical Programme

The following individuals have contributed to the Technical Programme, the part of the Bioregional Assessment Programme that undertakes bioregional assessments.

| Role or team | Contributor(s) |
|---|---|
| Assistant Secretary | Department of the Environment and Energy: Matthew Whitfort |
| Programme Director | Department of the Environment and Energy: John Higgins, Anthony Swirepik |
| Technical Programme Director | Bureau of Meteorology: Julie Burke |
| Projects Director | CSIRO: David Post |
| Principal Science Advisor | Department of the Environment and Energy: Peter Baker |
| Science Directors | CSIRO: Brent Henderson<br>Geoscience Australia: Steven Lewis |
| Integration | Bureau of Meteorology: Richard Mount (Integration Leader)<br>CSIRO: Becky Schmidt |
| Programme management | Bureau of Meteorology: Louise Minty<br>CSIRO: Paul Bertsch, Warwick McDonald<br>Geoscience Australia: Stuart Minchin |
| Project Leaders | CSIRO: Alexander Herr, Kate Holland, Tim McVicar, David Rassam<br>Geoscience Australia: Tim Evans<br>Bureau of Meteorology: Natasha Herron |
| Assets and receptors | Bureau of Meteorology: Richard Mount (Discipline Leader)<br>Department of the Environment and Energy: Glenn Johnstone, Wasantha Perera, Jin Wang |
| Bioregional Assessment Information Platform | Bureau of Meteorology: Lakshmi Devanathan (Team Leader), Derek Chen, Trevor Christie-Taylor, Melita Dahl, Angus MacAulay, Christine Price, Paul Sheahan, Kellie Stuart<br>CSIRO: Peter Fitch, Ashley Sommer<br>Geoscience Australia: Neal Evans |
| Communications | Bureau of Meteorology: Jessica York<br>CSIRO: Clare Brandon<br>Department of the Environment and Energy: John Higgins, Miriam McMillan, Milica Milanja<br>Geoscience Australia: Aliesha Lavers |
| Coordination | Bureau of Meteorology: Brendan Moran, Eliane Prideaux, Sarah van Rooyen<br>CSIRO: Ruth Palmer<br>Department of the Environment and Energy: Anisa Coric, Lucy Elliott, James Hill, Andrew Stacey, David Thomas, Emily Turner |
| Ecology | CSIRO: Anthony O'Grady (Discipline Leader), Caroline Bruce, Tanya Doody, Brendan Ebner, Craig MacFarlane, Patrick Mitchell, Justine Murray, Chris Pavey, Jodie Pritchard, Nat Raisbeck-Brown, Ashley Sparrow |

| Role or team | Contributor(s) |
|---|---|
| Geology | CSIRO: Deepak Adhikary, Emanuelle Frery, Mike Gresham, Jane Hodgkinson, Zhejun Pan, Matthias Raiber, Regina Sander, Paul Wilkes<br>Geoscience Australia: Steven Lewis (Discipline Leader) |
| Geographic information systems | CSIRO: Jody Bruce, Debbie Crawford, Dennis Gonzalez, Mike Gresham, Steve Marvanek, Arthur Read<br>Geoscience Australia: Adrian Dehelean |
| Groundwater modelling | CSIRO: Russell Crosbie (Discipline Leader), Tao Cui, Warrick Dawes, Lei Gao, Sreekanth Janardhanan, Luk Peeters, Praveen Kumar Rachakonda, Wolfgang Schmid, Saeed Torkzaban, Chris Turnadge, Andy Wilkins, Binzhong Zhou |
| Hydrogeology | Geoscience Australia: Tim Ransley (Discipline Leader), Chris Harris-Pascal, Jessica Northey, Emily Slatter |
| Information management | Bureau of Meteorology: Brendan Moran (Team Leader), Christine Panton<br>CSIRO: Qifeng Bai, Simon Cox, Phil Davies, Geoff Hodgson, Brad Lane, Ben Leighton, David Lemon, Trevor Pickett, Shane Seaton, Ramneek Singh, Matt Stenson<br>Geoscience Australia: Matti Peljo |
| Information model and impact analysis | Bureau of Meteorology: Carl Sudholz (Project Manager), Mark Dyall, Michael Lacey, Brett Madsen, Eliane Prideaux<br>Geoscience Australia: Trevor Tracey-Patte |
| Products | CSIRO: Becky Schmidt (Products Manager), Maryam Ahmad, Helen Beringen, Clare Brandon, Heinz Buettikofer, Sonja Chandler, Siobhan Duffy, Karin Hosking, Allison Johnston, Maryanne McKay, Linda Merrin, Sally Tetreault-Campbell, Catherine Ticehurst<br>Geoscience Australia: Penny Kilgour |
| Risk and uncertainty | CSIRO: Simon Barry (Discipline Leader), Jeffrey Dambacher, Rob Dunne, Jess Ford, Keith Hayes, Geoff Hosack, Adrien Ickowicz, Warren Jin, Dan Pagendam |
| Surface water hydrology | CSIRO: Neil Viney and Yongqiang Zhang (Discipline Leaders), Santosh Aryal, Mat Gilfedder, Fazlul Karim, Lingtao Li, Dave McJannet, Jorge Luis Peña-Arancibia, Tom Van Niel, Jai Vaze, Bill Wang, Ang Yang |

# Acknowledgements

This technical product was reviewed by several groups:

- Discipline Leaders: Richard Mount (assets), Anthony O’Grady (ecology), Tim Ransley (hydrogeology)
- Senior Science Leaders: David Post (Projects Director), Steven Lewis (Science Director, Geoscience Australia), Becky Schmidt (Products Manager)
- Technical Assurance Reference Group: Chaired by Peter Baker (Principal Science Advisor, Department of the Environment and Energy), this group comprises officials from the NSW, Queensland, South Australian and Victorian governments.

Valuable comments were also provided by Eliane Prideaux.

# Introduction

The Independent Expert Scientific Committee on Coal Seam Gas and Large Coal Mining Development (IESC) was established to provide advice to the federal Minister for the Environment on potential water-related impacts of coal seam gas (CSG) and large coal mining developments (IESC, 2015).

Bioregional assessments (BAs) are one of the key mechanisms to assist the IESC in developing this advice so that it is based on best available science and independent expert knowledge. Importantly, technical products from BAs are also expected to be made available to the public, providing the opportunity for all other interested parties, including government regulators, industry, community and the general public, to draw from a single set of accessible information. A BA is a scientific analysis, providing a baseline level of information on the ecology, hydrology, geology and hydrogeology of a bioregion with explicit assessment of the potential impacts of CSG and coal mining development on water resources.

The IESC has been involved in the development of *Methodology for bioregional assessments of the impacts of coal seam gas and coal mining development on water resources* (the BA methodology; Barrett et al., 2013) and has endorsed it. The BA methodology specifies how BAs should be undertaken. Broadly, a BA comprises five components of activity, as illustrated in Figure 1. Each BA is different, due in part to regional differences, but also in response to the availability of data, information and fit-for-purpose models. Where differences occur, these are recorded, judgments exercised on what can be achieved, and an explicit record is made of the confidence in the scientific advice produced from the BA.

## The Bioregional Assessment Programme

The Bioregional Assessment Programme is a collaboration between the Department of the Environment and Energy, the Bureau of Meteorology, CSIRO and Geoscience Australia. Other technical expertise, such as from state governments or universities, is also drawn on as required. For example, natural resource management groups and catchment management authorities identify assets that the community values by providing the list of water-dependent assets, a key input.

The Technical Programme, part of the Bioregional Assessment Programme, has undertaken BAs for the following bioregions and subregions (see http://www.bioregionalassessments.gov.au/assessments for a map and further information):

- the Galilee, Cooper, Pedirka and Arckaringa subregions, within the Lake Eyre Basin bioregion
- the Maranoa-Balonne-Condamine, Gwydir, Namoi and Central West subregions, within the Northern Inland Catchments bioregion
- the Clarence-Moreton bioregion
- the Hunter and Gloucester subregions, within the Northern Sydney Basin bioregion

- the Sydney Basin bioregion
- the Gippsland Basin bioregion.

Technical products (described in a later section) will progressively be delivered throughout the Programme.

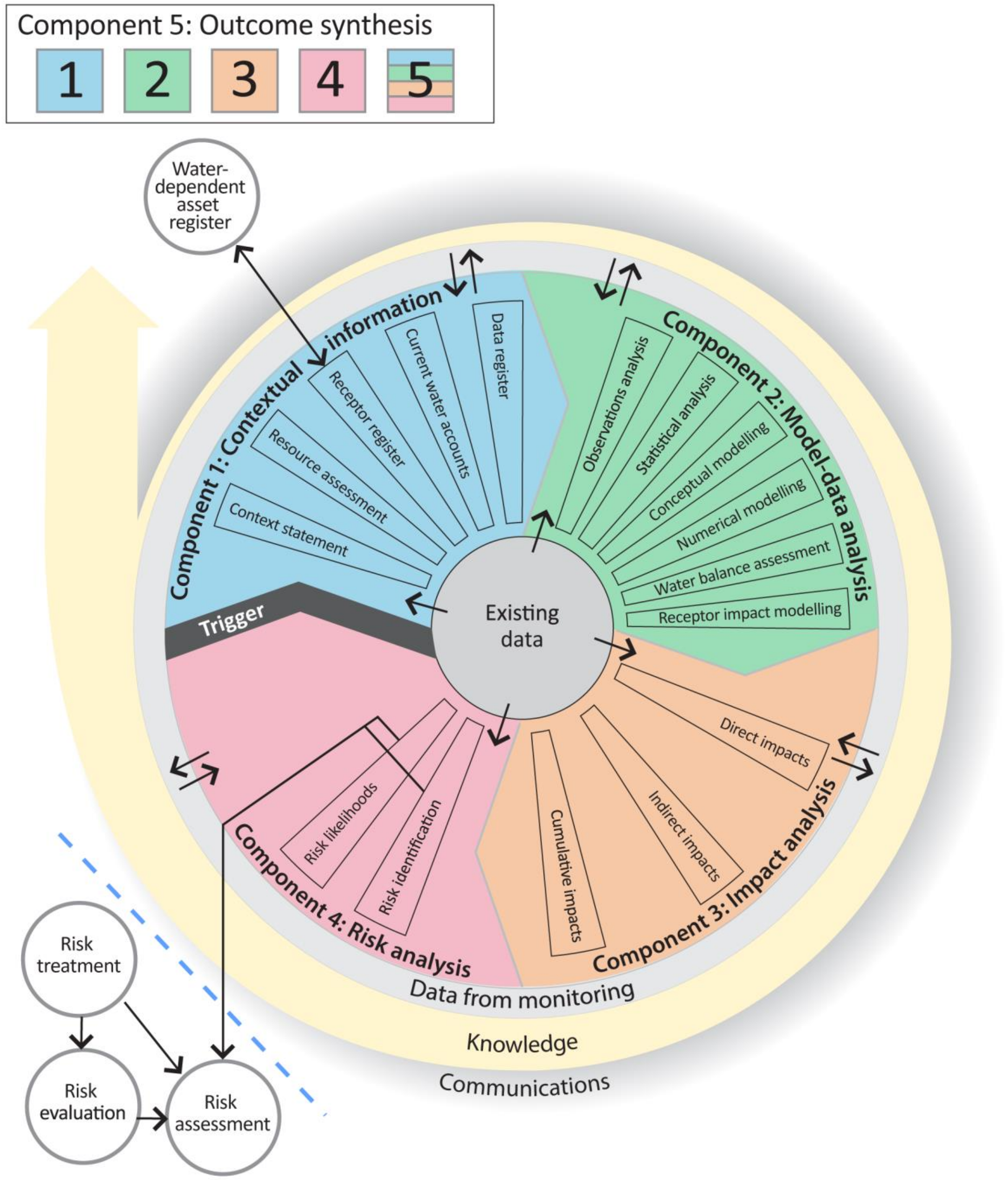

**Figure 1 Schematic diagram of the bioregional assessment methodology**

The methodology comprises five components, each delivering information into the bioregional assessment and building on prior components, thereby contributing to the accumulation of scientific knowledge. The small grey circles indicate activities external to the bioregional assessment. Risk identification and risk likelihoods are conducted within a bioregional assessment (as part of Component 4) and may contribute to activities undertaken externally, such as risk evaluation, risk assessment and risk treatment. Source: Figure 1 in Barrett et al. (2013), © Commonwealth of Australia

# Methodologies

The overall scientific and intellectual basis of the BAs is provided in the BA methodology (Barrett et al., 2013). Additional guidance is required, however, about how to apply the BA methodology to a range of subregions and bioregions. To this end, the teams undertaking the BAs have developed and documented detailed scientific submethodologies (Table 1) to, in the first instance, support the consistency of their work across the BAs and, secondly, to open the approach to scrutiny, criticism and improvement through review and publication. In some instances, methodologies applied in a particular BA may differ from what is documented in the submethodologies.

The relationship of the submethodologies to BA components and technical products is illustrated in Figure 2. While much scientific attention is given to assembling and transforming information, particularly through the development of the numerical, conceptual and receptor impact models, integration of the overall assessment is critical to achieving the aim of the BAs. To this end, each submethodology explains how it is related to other submethodologies and what inputs and outputs are required. They also define the technical products and provide guidance on the content to be included. When this full suite of submethodologies is implemented, a BA will result in a substantial body of collated and integrated information for a subregion or bioregion, including new information about the potential impacts of coal resource development on water and water-dependent assets.

# About this submethodology

The following notes are relevant only for this submethodology.

- All reasonable efforts were made to provide all material under a Creative Commons Attribution 3.0 Australia Licence. It should be assumed that third parties are not entitled to use this material without permission from the copyright owner.
- Visit http://www.bioregionalassessments.gov.au to access metadata (including copyright, attribution and licensing information) for datasets cited or used to make figures in this product.
- In addition, the datasets are published online if they are unencumbered (able to be published according to conditions in the licence or any applicable legislation). The Bureau of Meteorology archives a copy of all datasets used in the BAs. This archive includes datasets that are too large to be stored online and datasets that are encumbered. The community can request a copy of these archived data at http://www.bioregionalassessments.gov.au.

- The citation details of datasets are correct to the best of the knowledge of the Bioregional Assessment Programme at the publication date of this submethodology. Readers should use the hyperlinks provided to access the most up-to-date information about these data; where there are discrepancies, the information provided online should be considered correct. The dates used to identify Bioregional Assessment Source Datasets are the dataset's created date. Where a created date is not available, the publication date or last updated date is used.

**Table 1 Methodologies**

Each submethodology is available online at http://data.bioregionalassessments.gov.au/submethodology/XXX, where 'XXX' is replaced by the code in the first column. For example, the BA methodology is available at http://data.bioregionalassessments.gov.au/submethodology/bioregional-assessment-methodology and submethodology M02 is available at http://data.bioregionalassessments.gov.au/submethodology/M02. Submethodologies might be added in the future.

| Code | Proposed title | Summary of content |
|---|---|---|
| bioregional-assessment-methodology | *Methodology for bioregional assessments of the impacts of coal seam gas and coal mining development on water resources* | A high-level description of the scientific and intellectual basis for a consistent approach to all bioregional assessments |
| M02 | *Compiling water-dependent assets* | Describes the approach for determining water-dependent assets |
| M03 | *Assigning receptors to water-dependent assets* | Describes the approach for determining receptors associated with water-dependent assets |
| M04 | *Developing a coal resource development pathway* | Specifies the information that needs to be collected and reported about known coal and coal seam gas resources as well as current and potential resource developments |
| M05 | *Developing the conceptual model of causal pathways* | Describes the development of the conceptual model of causal pathways, which summarises how the 'system' operates and articulates the potential links between coal resource development and changes to surface water or groundwater |
| M06 | *Surface water modelling* | Describes the approach taken for surface water modelling |
| M07 | *Groundwater modelling* | Describes the approach taken for groundwater modelling |
| M08 | *Receptor impact modelling* | Describes how to develop receptor impact models for assessing potential impact to assets due to hydrological changes that might arise from coal resource development |
| M09 | *Propagating uncertainty through models* | Describes the approach to sensitivity analysis and quantification of uncertainty in the modelled hydrological changes that might occur in response to coal resource development |
| M10 | *Impacts and risks* | Describes the logical basis for analysing impact and risk |
| M11 | *Systematic analysis of water-related hazards associated with coal resource development* | Describes the process to identify potential water-related hazards from coal resource development |

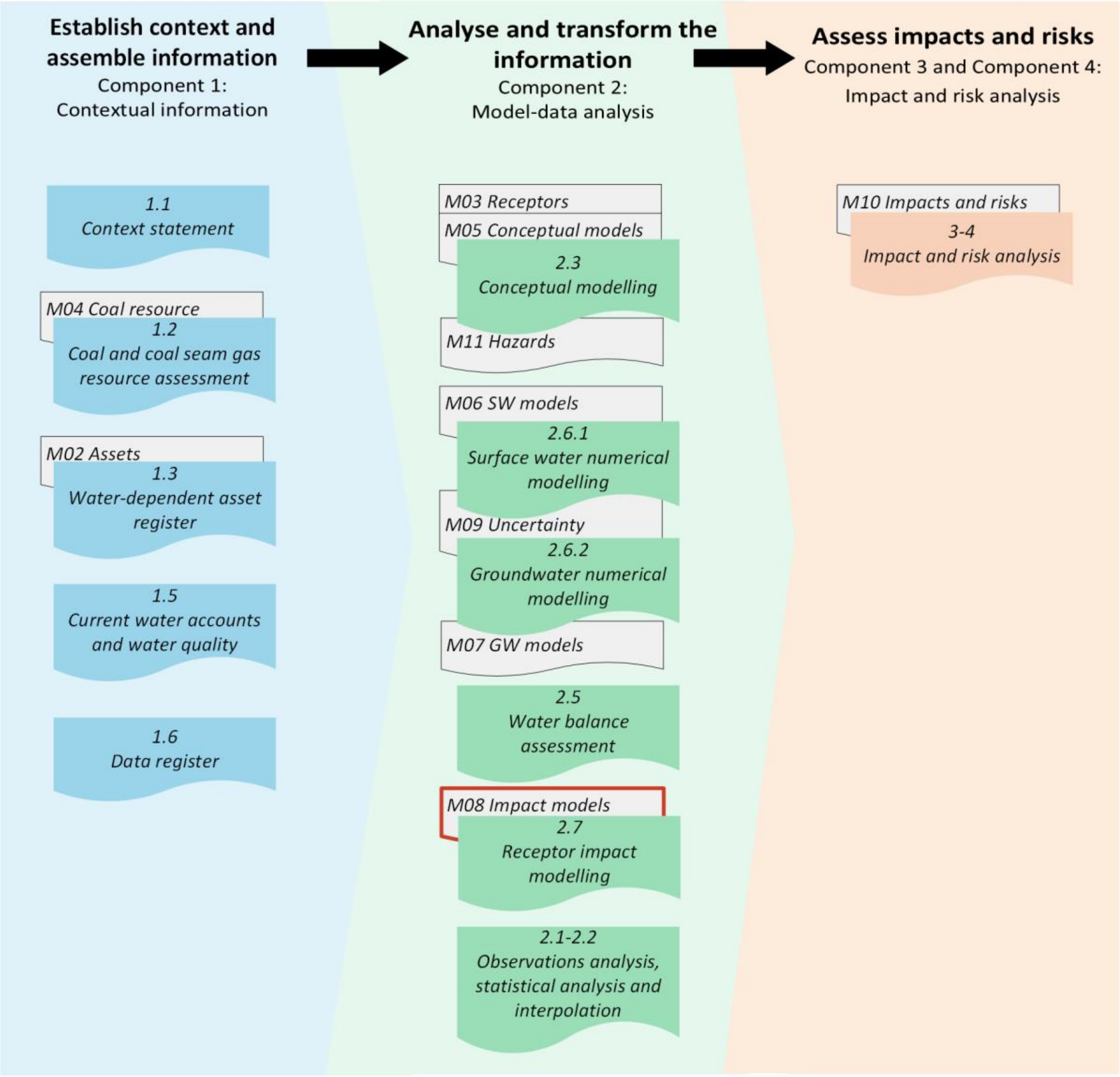

**Figure 2 Technical products and submethodologies associated with each component of a bioregional assessment**

In each component (Figure 1) of a bioregional assessment (BA), a number of technical products (coloured boxes, see also Table 2) are potentially created, depending on the availability of data and models. The light grey boxes indicate submethodologies (Table 1) that specify the approach used for each technical product. The red outline indicates this submethodology. The BA methodology (Barrett et al., 2013) specifies the overall approach.

# Technical products

The outputs of the BAs include a suite of technical products presenting information about the ecology, hydrology, hydrogeology and geology of a subregion or bioregion and the potential impacts of CSG and coal mining developments on water resources, both above and below ground. Importantly, these technical products are available to the public, providing the opportunity for all interested parties, including community, industry and government regulators, to draw from a single set of accessible information when considering CSG and large coal mining developments in a particular area.

The BA methodology specifies the information to be included in technical products. Figure 2 shows the relationship of the technical products to BA components and submethodologies. Table 2 lists the content provided in the technical products, with cross-references to the part of the BA methodology that specifies it.

Technical products are delivered as reports (PDFs). Additional material is also provided, as specified by the BA methodology:

- unencumbered data syntheses and databases
- unencumbered tools, model code, procedures, routines and algorithms
- unencumbered forcing, boundary condition, parameter and initial condition datasets
- lineage of datasets (the origin of datasets and how they are changed as the BA progresses)
- gaps in data and modelling capability.

In this context, unencumbered material is material that can be published according to conditions in the licences or any applicable legislation. All reasonable efforts were made to provide all material under a Creative Commons Attribution 3.0 Australia Licence.

Technical products, and the additional material, are available online at http://www.bioregionalassessments.gov.au.

The Bureau of Meteorology archives a copy of all datasets used in the BAs. This archive includes datasets that are too large to be stored online and datasets that are encumbered. The community can request a copy of these archived data at http://www.bioregionalassessments.gov.au.

**Table 2 Technical products delivered by the Bioregional Assessment Programme**

For each subregion or bioregion in a bioregional assessment (BA), technical products are delivered online at http://www.bioregionalassessments.gov.au. Other products – such as datasets, metadata, data visualisation and factsheets – are also provided online. There is no product 1.4; originally this product was going to describe the receptor register and application of landscape classes as per Section 3.5 of the BA methodology, but this information is now included in product 2.3 (conceptual modelling) and used in products 2.6.1 (surface water modelling) and 2.6.2 (groundwater modelling). There is no product 2.4; originally this product was going to include two- and three-dimensional representations as per Section 4.2 of the BA methodology, but these are instead included in products such as product 2.3 (conceptual modelling), product 2.6.1 (surface water numerical modelling) and product 2.6.2 (groundwater numerical modelling).

| Component | Product code | Title | Section in the BA methodology[a] |
|---|---|---|---|
| **Component 1: Contextual information for the subregion or bioregion** | 1.1 | Context statement | 2.5.1.1, 3.2 |
| | 1.2 | Coal and coal seam gas resource assessment | 2.5.1.2, 3.3 |
| | 1.3 | Description of the water-dependent asset register | 2.5.1.3, 3.4 |
| | 1.5 | Current water accounts and water quality | 2.5.1.5 |
| | 1.6 | Data register | 2.5.1.6 |
| **Component 2: Model-data analysis for the subregion or bioregion** | 2.1-2.2 | Observations analysis, statistical analysis and interpolation | 2.5.2.1, 2.5.2.2 |
| | 2.3 | Conceptual modelling | 2.5.2.3, 4.3 |
| | 2.5 | Water balance assessment | 2.5.2.4 |
| | 2.6.1 | Surface water numerical modelling | 4.4 |
| | 2.6.2 | Groundwater numerical modelling | 4.4 |
| | 2.7 | Receptor impact modelling | 2.5.2.6, 4.5 |
| **Component 3 and Component 4: Impact and risk analysis for the subregion or bioregion** | 3-4 | Impact and risk analysis | 5.2.1, 2.5.4, 5.3 |
| **Component 5: Outcome synthesis for the bioregion** | 5 | Outcome synthesis | 2.5.5 |

[a] *Methodology for bioregional assessments of the impacts of coal seam gas and coal mining development on water resources* (Barrett et al., 2013)

## References

Barrett DJ, Couch CA, Metcalfe DJ, Lytton L, Adhikary DP and Schmidt RK (2013) Methodology for bioregional assessments of the impacts of coal seam gas and coal mining development on water resources. A report prepared for the Independent Expert Scientific Committee on Coal Seam Gas and Large Coal Mining Development through the Department of the Environment. Department of the Environment, Australia. Viewed 27 April 2018, http://data.bioregionalassessments.gov.au/submethodology/bioregional-assessment-methodology.

IESC (2015) Information guidelines for the Independent Expert Scientific Committee advice on coal seam gas and large coal mining development proposals. Independent Expert Scientific Committee on Coal Seam Gas and Large Coal Mining Development, Australia. Viewed 27 April 2018, http://www.iesc.environment.gov.au/publications/information-guidelines-independent-expert-scientific-committee-advice-coal-seam-gas.

# 1 Background and context

Receptor impact modelling is integral to undertaking the impact and risk analysis of a bioregional assessment (BA). The *Methodology for bioregional assessments of the impacts of coal seam gas and coal mining development on water resources* (the BA methodology; Barrett et al., 2013) states:

> Modelling and analysis of the direct, indirect and cumulative impacts of coal seam gas (CSG) and coal mining development on anthropogenic and ecological receptors is the pivotal component of a BA.

The receptor impact model predicts how a receptor impact variable (a variable for which water-related impacts on assets are to be assessed) responds to direct, indirect and cumulative impacts of CSG and coal mining development. Although the BA methodology gives a high-level overview of the components required for a BA, it is not sufficiently detailed to clearly guide Assessment teams performing specific assessment tasks. This submethodology provides the explicit detail needed to develop the receptor impact models that are required to assess the potential impacts from CSG and large coal mining on receptor impact variables. This M08 report is a methodology report that provides the technical detail underlying the reporting in technical products 2.7 and 3-4. A quantitative, numerical and probabilistic modelling approach is described. The Assessment teams can use this method to undertake probabilistic assessment for a receptor impact variable as required by BA. The receptor impact modelling methodology was purposely developed to enable coherent assimilation of empirical data if permitted by available resources within the BA Programme. The methodology describes the generation of a probabilistic prior that can be updated with empirical data where available.

Receptor impact models are functions that translate hydrological changes into the distribution of potential economic, sociocultural and/or ecological outcomes that may arise from those changes (see product 2.7 (receptor impact modelling) for applied examples, Figure 3). Within BAs, hydrological changes are described by hydrological response variables. Hydrological response variables are defined as hydrological characteristics of the system that potentially change due to coal resource development (for example, drawdown or the annual flow volume). They are thought to be instrumental in maintaining and shaping the ecological components, processes and values provided by the ecosystems in each landscape class. Examples of hydrological response variables are found in product 2.7 (receptor impact modelling) (Figure 3). A receptor impact variable is defined as a characteristic of the system that, according to the conceptual modelling, potentially changes due to changes in hydrological response variables (for example, condition of the breeding habitat for a given species, or biomass of river red gums). Economic, sociocultural or ecological outcomes are represented by receptor impact variables, and a receptor impact model portrays the relationship between a particular receptor impact variable (e.g. the percent foliage cover of woody riparian vegetation) and one or more hydrological response variables (e.g. the change in depth to groundwater).

Receptor impact models are only developed for ecological receptor impact variables in BAs. Potential impacts on economic assets were assessed in BAs by estimating the changes to the availability of groundwater and surface water (see product 3-4 (impact and risk analysis) for details, Figure 3). Likewise, some sociocultural assets have a direct relationship with groundwater and surface water and are treated similarly to economic assets. Other sociocultural receptor impact variables may have an ecological component that can be assessed by an ecological receptor impact model.

Ecological receptor impact modelling is the focus of this submethodology: a method is provided that links hydrological response variables to receptor impact variables enabling prediction of ecological responses to coal resource development. The hydrological response variables may interact with and produce cumulative impacts on an ecological receptor. Ecological receptor impact models quantify the potential impacts on water-dependent assets that may have ecological value. In the ecological scientific literature, receptor impact models are often known as 'ecological response functions' (Boulton et al., 2014). This submethodology develops a novel approach to constructing ecological response functions that allows for direct, indirect and cumulative impacts, while coherently incorporating uncertainty, expert assessments and potentially empirical data.

Two potential futures are considered by BAs:

- *baseline coal resource development* (baseline): a future that includes all coal mines and CSG fields that are commercially producing as of December 2012
- *coal resource development pathway* (CRDP): a future that includes all coal mines and CSG fields that are in the baseline as well as those that are expected to begin commercial production after December 2012.

The difference in results between the CRDP and baseline is the change that is primarily reported in a BA. This change is due to the *additional coal resource development* – all coal mines and CSG fields, including expansions of baseline operations, that are expected to begin commercial production after December 2012. BAs focus solely on water-related impacts, and specifically those related to water quantity and availability. Potential water quality hazards and pathways are identified but any (qualitative) analysis is limited to salinity as other water quality impacts beyond the scope of BA.

Ecological receptor impact models in BAs are relevant to specific landscape classes that are potentially impacted by the additional coal resource development over the baseline, encapsulated by the CRDP. A landscape class is defined as an ecosystem with characteristics that are expected to respond similarly to changes in groundwater and/or surface water due to coal resource development (companion submethodology M03 (as listed in Table 1) for assigning receptors to water-dependent assets (O'Grady et al., 2016); companion submethodology M05 (as listed in Table 1) for developing a conceptual model of causal pathways (Henderson et al., 2016)). For each potentially impacted landscape class, the overall conceptual approach is to propagate uncertainty from hydrological models into receptor impact models (see companion submethodology M09 (as listed in Table 1) for propagating uncertainty through models (Peeters et al., 2016)). The probabilistic outputs from the hydrology models, which depend on the choice of futures (baseline or CRDP), become the inputs into the receptor impact models. The receptor impact models, in turn, produce probabilistic predictions of the receptor impact variables conditional on the hydrological inputs. This approach thus provides a coherent probabilistic assessment of the receptor impact variable while accounting for the direct, indirect and cumulative impacts of CSG and coal mining development.

The receptor impact modelling described by this submethodology is a crucial part of the cumulative impact and risk analysis process used in a BA (Component 3 and Component 4; Figure 3). Receptor impact modelling is a key step of the risk analysis as it converts the potentially abstract information about hydrological changes to ecological variables of scientific interest that stakeholders care about and can more readily understand and interpret. In particular, outcomes of the modelling will relate more closely to their ecological values and beliefs and therefore support community discussion and decision making about acceptable levels of coal resource development.

Receptor impact models may be constructed for three types of asset groups: economic, sociocultural and ecological. An overview of the construction of the receptor impact model for the three asset groups is provided in this section. The primary focus of this submethodology is on ecological assets; only ecological receptor impact models were developed. Subsequent sections concentrate entirely on ecological assets. Any subregion-specific deviations from this approach, or gaps and limitations, are addressed in product 2.7 (receptor impact modelling) for that subregion.

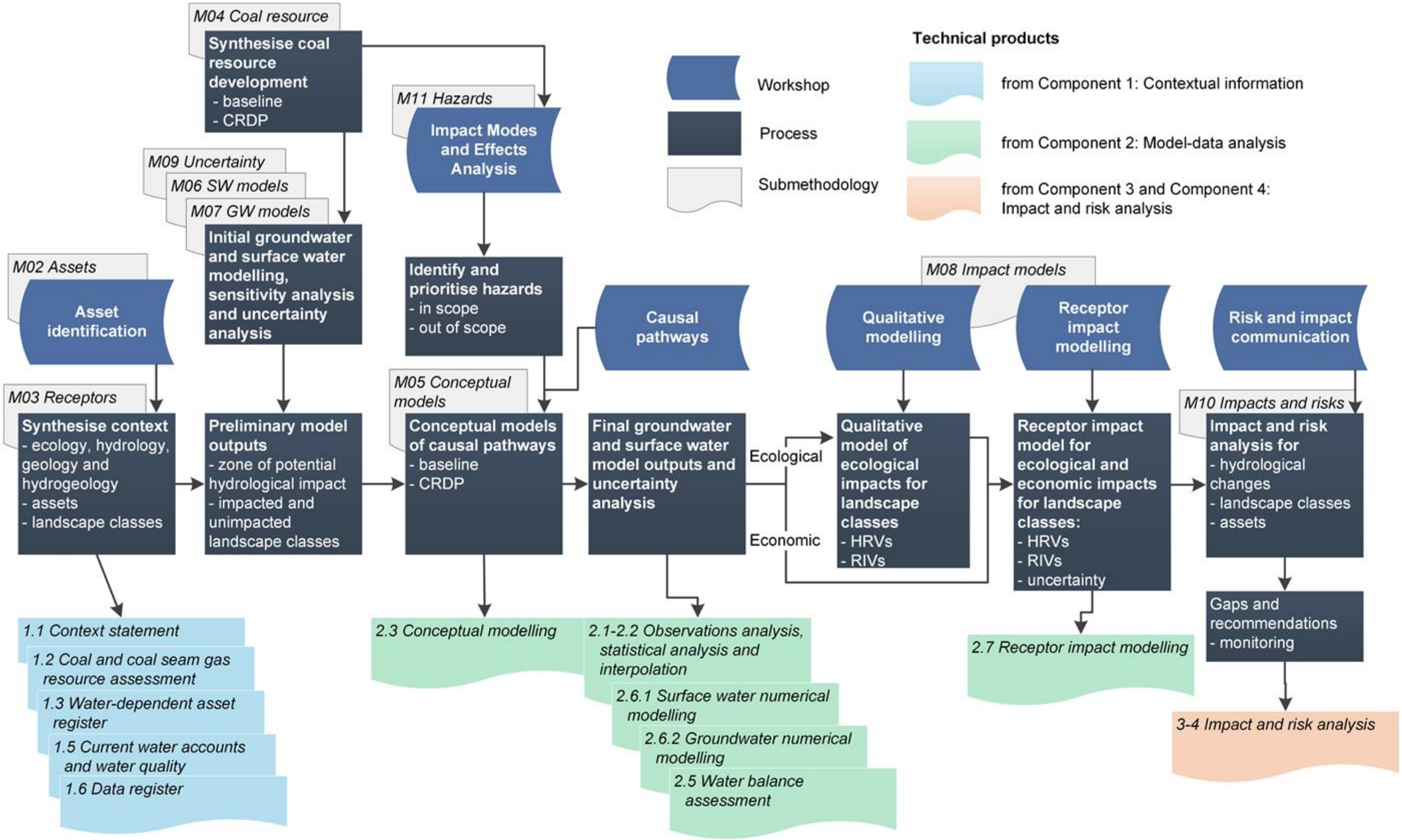

**Figure 3 A bioregional assessment from end to end, showing the relationship between the workflow, technical products, submethodologies and workshops**

CRDP = coal resource development pathway, GW = groundwater, HRVs = hydrological response variables, RIVs = receptor impact variables, SW = surface water

## 1.1 Economic assets

In BAs, economic water-dependent assets are confined to groundwater and surface water management zones or areas and comprise specific water access entitlements or rights and other water supply features or infrastructures. Within these asset subgroups there may be many individual asset elements (e.g. water supply bores within a groundwater management zone).

The scope of the BA methodology (Barrett et al., 2013) does not require detailed analysis of the economic impacts of changes induced by the development of the coal resources. Instead the impact and risk analysis concludes at the assessment of changes to the availability of the relevant groundwater or surface water.

Within the scope of BA, potential impacts to economic assets are tied directly to potential hydrological response variables, which can be assessed against specific management thresholds, whereas indirect impacts and complex feedbacks are anticipated for the ecological assets that are the focus of this submethodology. Within BA, there is no need for economic receptor impact modelling as the range of potential hydrological change can be considered against already defined thresholds such as those specified by the *NSW Aquifer Interference Policy* (NSW Office of Water, 2012) or the requirements of water resource plans under Queensland's *Water Act 2000*.

Economic assets are discussed no further in this submethodology but are discussed in companion submethodology M10 (as listed in Table 1) (Henderson et al., 2018) where there is a broader consideration of impacts and risks.

## 1.2 Sociocultural assets

In BAs, water-dependent sociocultural assets are classified into 'Heritage site', 'Indigenous site' or 'Recreation area' asset classes. Receptor impact modelling is not directly undertaken for sociocultural assets and the impact and risk analysis is limited to characterising the hydrological change that may be experienced by those assets. In some cases, sociocultural assets also have obvious ecological value (e.g. the health of a wetland that is also identified as having Indigenous value for fishing) and the changes to ecological receptor impact variables may be particularly important for identifying potential impacts and risks to them.

Sociocultural assets are discussed no further in this submethodology. However, it is important to note that the receptor impact model for ecological assets will also play an important role in assessing any hydrological changes for those sociocultural assets that have ecological characteristics.

## 1.3 Ecological assets

The construction of receptor impact models for ecological assets and ecosystems presents great challenges. The first overarching challenge is that ecological systems are complex and constantly changing. There are inherent limits to explaining patterns and to predicting outcomes for ecological variables. Logically incorporating this predictability problem into an uncertainty analysis can be practically challenging to implement. It is therefore incumbent for the approach to document and demonstrate how it will quantitatively assimilate the current state of knowledge, and use this information in such a way to ensure the applicability of the general approach to a large number of assets with various spatial footprints and at different points in time.

The second challenge for ecological receptor impact models is the specific nature of the analysis. The analysis requires assessing impacts at different points in time. However, the hydrological impacts of coal resource development in the subregion or bioregion are non-stationary, that is, the impacts vary over time because of the timing of coal resource developments, the lagged response of hydrological response variables to development pressures, and the lagged response of receptor impact variables to changes in the hydrological response variables. Thus, while it is attractive to conceptualise the problem in terms of a comparison of the existing hydrological regime before coal resource development and the new regime after coal resource development, such a simple temporal breakdown does not adequately reflect key aspects of the problem. The hydrological regime will typically change continuously as coal resource developments begin and change their patterns of water use and management during different operational phases. At the end of the coal resource development, many aspects of the hydrological regime could potentially return to their previous state or alternatively undergo a perturbed trajectory. The transitory effects may not completely be exhausted even by the end of the longest-term projection for the receptor impact variable in a BA. Thus, the problem cannot be considered as a simple change of steady state.

The non-stationarity of hydrological impact also restricts the available data that can be used to empirically estimate a receptor impact model. Fundamentally, the sequence of hydrological changes may be a key determinant of change in the receptor impact variable. However, many of the hydrological changes, both in magnitude and sequence, may be novel in these systems. Sparse data will sometimes exist to empirically calibrate these relationships. The use of ecosystem modelling is similarly restricted by the lack of process knowledge and associated data to calibrate the relationships contained in the model. Techniques such as Bayes Nets could be considered (e.g. Marcot et al. (2001)), but they do not naturally accommodate spatial and temporal phenomena, and entail significant resources to parameterise.

The fractured nature of the knowledge base means that the use of expert opinion will typically be needed to construct receptor impact models, which can be assessed and updated with empirical data where possible. The experts can integrate their knowledge base to make predictions about likely outcomes related to hydrological change (O’Hagan, 1998). Carefully constructed questions help experts focus on the key issues that need to be considered and elicit their response in a structured and transparent way.

The elicitation process requires careful construction (O’Hagan et al., 2006). It is well known that poorly designed elicitations can seriously impact on the reliability of the results. Ambiguous questions mean that experts may misunderstand the task and introduce additional uncertainty into the analysis. Poor processes can lead to problems, such as dominance of debate by vocal individuals. Inadequate protocols can lead to issues, such as anchoring, where experts do not explore the full extent of their knowledge and beliefs. Unorganised, inefficient or unclear protocols can confuse experts and lead to burn out or decreased motivation, which leads them to drop out from the elicitation process. There is also a balance between choosing a protocol that is simple to explain versus a complex protocol that requires substantial training and time commitment of experts. A very pragmatic protocol is one that allows experts to contribute their knowledge early without too much education on the elicitation process. On the other hand, overly simplistic protocols may not elicit the experts' knowledge correctly and can also lead to confusion on the nature of what exactly was elicited if the education portion of the elicitation process is reduced too much. These trade-offs are of particular importance for receptor impact models deployed in a BA, where many such models are developed across a subregion or bioregion and the demand for expert involvement and motivated participation is essential for successful completion of each task. Ideally, the same experts will be involved for deriving conceptual models and also the quantitative probabilistic receptor impact models. But the method described here also allows for the pragmatic cases where different experts contribute at various times, for example, because of availability limitations or a shift in the focus of domain expertise for a given landscape class.

The opinions expressed by different experts will sometimes disagree. This fact simply reflects variation in their understandings and beliefs and so does not directly undermine the use of the approach. However, it means that the choice of experts can have a material effect on the analysis, so the elicitation needs to be done in a consistent, flexible and principled way. In some senses, the experts operate as a jury. Provided the experts represent the diverse views across the relevant informed community within their consensus opinion, readers will have confidence that a wide range of informed opinions have been considered and reported in the Assessment. The choice of experts, including their identification, invitation and participation in the process, needs close attention to ensure appropriate expertise is included. In practice, expert availability can be a non-trivial constraint. The expert invitation process was a collaborative effort among the Office of Water Science, the Bureau of Meteorology and the BA ecology discipline teams for each bioregion or subregion, which provided communication channels to key regional institutions and individuals with a wide range of expertise.

Ultimately, however, the best way to assess the experts' judgement is to collect relevant empirical data wherever possible. Experts will be encouraged to include knowledge of existing empirical data from independent sources within their assessments. A very important additional objective of the receptor impact model approach is to allow for the coherent incorporation of relevant empirical data if it were to become available. These data may be obtained through a defined monitoring program that targets a receptor impact variable as part of validating (or invalidating) the risk predictions and the characterisation of the receptor impact model. This goal for coherent data assimilation dictates the choice of model structure, which must allow for potentially very different forms of ecological data, such as continuous (e.g. abundance), non-negative integers (e.g. counts), binomial counts (e.g. percent cover) or binary (e.g. presence–absence) responses. This goal, which seeks to ensure the coherent updating of uncertainty estimates given potential empirical data, also guides the selection of the receptor impact variable, where it forms a clear criterion that any selected receptor impact variable must at least in principle be measurable. Such a requirement additionally ensures that the target receptor impact variable is both well-defined and also accessible to expert assessment.

In summary, the risks in the use of expert-based information can be managed by appropriate protocols and procedures. The workflow for ecological receptor impact modelling is outlined in Figure 4. Input from independent external ecology experts contributes to the workflow at three separate stages, as does hydrological modelling simulation output and the expertise of the BA hydrology modellers. Both the external experts and the hydrologists contribute to the selection of hydrological response variables that are ecologically meaningful and also amenable to hydrology modelling. The careful definition of landscape class forms the spatial context for the expert elicitation and receptor impact model analysis (Chapter 2). Experts should be engaged in a structured way with strong facilitation to ensure clarity of communication and focus on the issue in question. Elucidation of the ecological ecosystem within a landscape class is conducted at a qualitative modelling workshop that maps the key ecological variables, hydrological response variables and the linkages among these variables (Chapter 3). This modelling exercise is used to choose key hydrological response variables and receptor impact variables to progress for a receptor impact model analysis (Chapter 4). Based on the selected hydrological response variables and receptor impact variables, an efficient design of elicitation scenarios is constructed, which directly addresses the resource and time limitations imposed by expert availability (Chapter 5). The expert education and elicitation process as experienced by the experts and its theoretical underpinning is given in Chapter 6. The derivation of the receptor impact model and prediction methods are given in Chapter 7 and Chapter 8. Each stage is described in the following subsections.

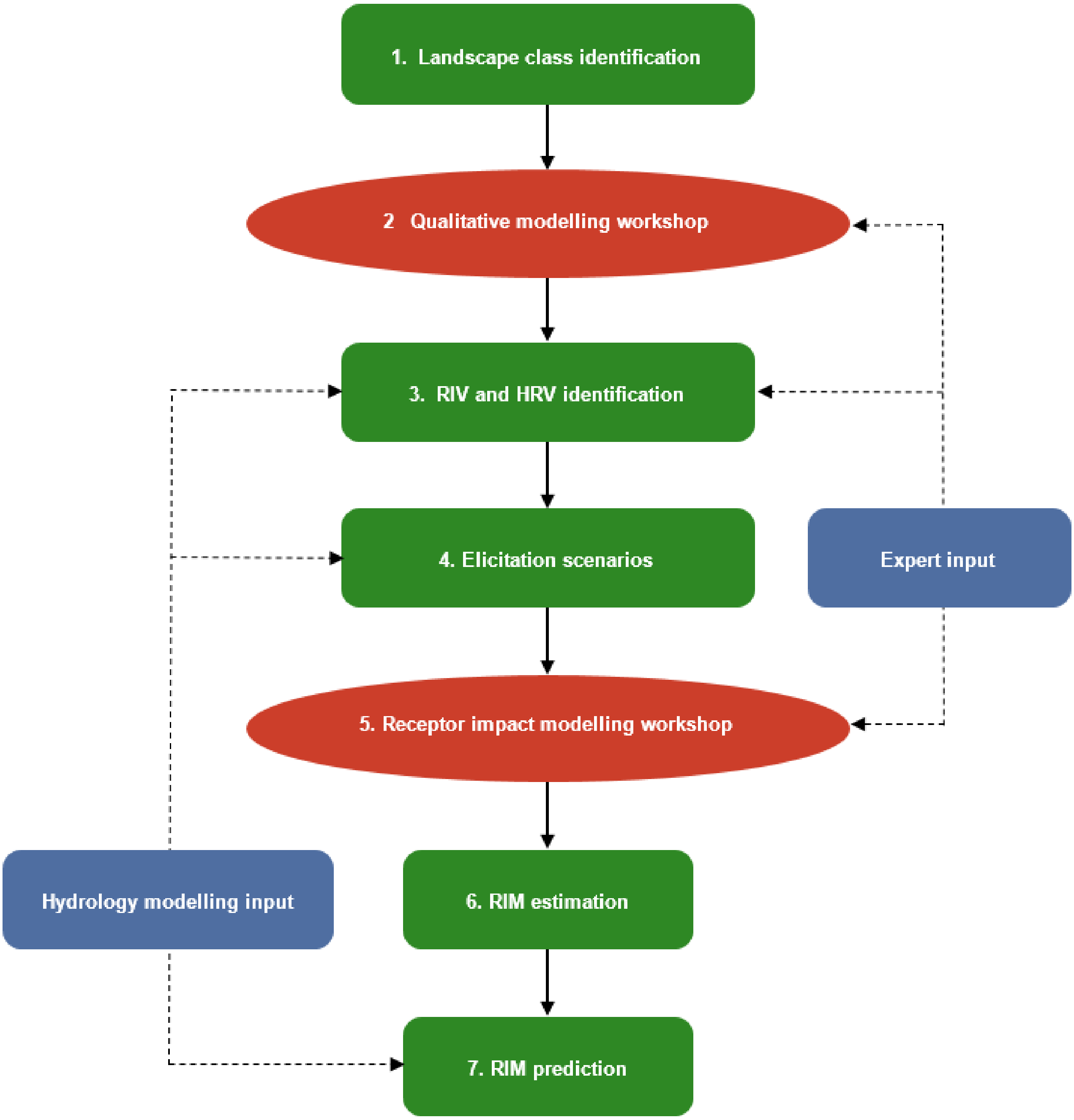

**Figure 4 Outline of the ecological receptor impact workflow with contributions from external independent ecology experts, groundwater hydrology modelling and surface water hydrology modelling identified by stage**

The workflow leads to the construction of a receptor impact model (RIM) that predicts the response of a receptor impact variable (RIV) conditional on hydrological response variables (HRVs). The uncertainty encapsulated by the hydrology modelling is propagated through the RIM when predicting the RIV response to the choice of BA futures (baseline or coal resource development pathway) across a landscape class. Workshop steps are shown in red, ecology and hydrology expert input sources are shown in blue.

# 2 Identification of potentially impacted landscape classes (stage 1)

Within the BAs, all modelling of the potential ecological impacts of coal resource development is organised by, and conditioned upon, landscape classes. This classification provides a structure that enables qualitative and quantitative modelling at a scale and resolution that is appropriate to the objectives and scale of the BA. Each bioregion or subregion has multiple landscape classes grouped into terrestrial and aquatic landscape groups. Examples of landscape classes are found in product 2.7 (receptor impact modelling) (Figure 3).

The purpose of the landscape classification is to partition the landscape into ecologically similar subunits. The ecological similarities mainly manifest as similar vegetation types (terrestrial landscape group) and similar stream classes (aquatic landscape group) resulting from common geological, geomorphological and/or hydrological characteristics. Importantly, the ecosystems within each landscape class are assumed to respond in a similar way to any alterations in groundwater and/or surface water regimes.

In most instances, the BA landscape classification capitalises on existing aquatic and terrestrial classification schemes and hence links to existing conceptual models of these ecosystems and their water dependency. These models provide a conceptual understanding of how key ecosystem components and processes are shaped and maintained by surface water and groundwater regimes. The criteria that delineate the landscape classes and the methods used to identify landscape classes within each bioregion or subregion, are summarised by companion submethodology M03 (as listed in Table 1) for assigning receptors to water-dependent assets (O’Grady et al., 2016).

For a given bioregion or subregion, only a subset of landscape classes may be affected by additional coal resource development. BA identifies landscape classes that could be impacted by coal resource development as those landscape classes that lie wholly or partially within a zone of potential hydrological change. The zone of potential hydrological change is defined as the union of the groundwater zone of potential hydrological change and the surface water zone of potential hydrological change. The groundwater zone is conservatively defined as the area with a greater than 5% chance of exceeding 0.2 m of drawdown in the relevant aquifers.

The surface water zone of potential hydrological change is defined in a similarly conservative manner (companion submethodology M06 (as listed in Table 1) for surface water modelling (Viney, 2016)). For flux-based variables (e.g. the annual flow volume or daily flow rate at the 1st percentile) it is defined by the surface water model nodes that exhibit at least a 1% change in the variable, relative to the baseline value, in at least 5% of the model simulations (replicates). For most of the frequency-based variables (e.g. the number of zero-flow days per year or the number of low-flow days per year), the zone is defined by model nodes that exhibit at least a 5% chance of the same relative magnitude of change (at least 1% compared to baseline) for at least 3 days in any simulated year. For one of the frequency-based metrics (the number of low-flow spells per year in perennial streams), it is defined by a greater than 5% chance of a change in the variable for at least two spells in any year (see companion submethodology M06 (as listed in Table 1) for surface water modelling (Viney, 2016)).

It is important to recognise that the zone of potential hydrological change represents an estimate of those parts of the landscape that are likely to experience at least some level of hydrological change attributable to coal resource development. The zone serves only to distinguish those landscape classes that should be taken into the next step of the receptor impact methodology from those landscape classes that should not, on the grounds that the latter are predicted to experience negligible (or insignificant) exposure to hydrological change due to coal resource development.

# 3 Qualitative mathematical modelling (stage 2)

In this work we have been guided by a strategy of model building that recognises a practical trade-off between realism, generality and precision when building and analysing models of complex systems (Levins, 1966, 1998). To obtain a manageable and useful model, one typically sacrifices one attribute for the other two. Qualitative mathematical models emphasise generality and realism, but lack precision, while numerical simulation models can be both precise and realistic but are not generalisable (i.e. application of the model to changed circumstance requires reparameterisation). A third approach is statistical, and emphasises precision and generality. Here there are precise insights into the general pattern of correlations among variables, but at the cost of causal understanding of the processes involved. In practice we seek a robust strategy that considers combinations of different modelling approaches, such that models are mutually informative and build upon the strengths and insights of other approaches. The impact and risk analysis for BAs is being informed by all three modelling approaches. In this section we describe the basic methods underpinning qualitative mathematical modelling.

Qualitative mathematical modelling serves three roles within BA's receptor impact modelling strategy. Firstly, sign-directed graphs (e.g. see Figure 5), built for each potentially impacted landscape class, document current understanding of how the landscape class ecosystem 'works' including key interactions between the system's physical, chemical and biological components, and their dependence on hydrology.

Secondly, through qualitative predictions of increase, decrease or no change, the models capture the direct and indirect effects that are anticipated to occur following a sustained change to the surface water or groundwater regimes that maintain the ecosystem of the landscape class. Lastly, BAs use the hydrological factors identified in the models, and the direct and indirect effects that they predict, to identify appropriate hydrological response variables and receptor impact variables for the next stages of the receptor impact model strategy.

The BA methodology (Barrett et al., 2013) requires an explicit assessment of the potential direct, indirect and cumulative impacts of CSG and coal mining development on water resources, together with an analysis of the associated uncertainties. A coherent analysis of uncertainty necessitates a probabilistic analysis, so in effect the BA methodology requires a probabilistic landscape-level assessment of the risks that coal resource development poses to water resources. The BA methodology goes on to define direct impacts as those associated with CSG and coal mining developments that impact on natural resources without intervening agents or pathways, whilst indirect impacts are defined as those impacts on receptors (within water-dependent assets) that are produced as a result of a (simple or complex) pathway of cause and effect.

It is important to recognise that the BA methodology sets an ambitious target, seeking to predict how direct, indirect and cumulative impacts on hydrology by development affect receptors at multiple time points, but does not describe how to meet it. The receptor impact modelling strategy is designed to operationalise the aspirations of the BA methodology, and this process begins by distinguishing direct and indirect impacts in a conceptual model of the potential stress imposed on the ecosystem by coal resource development.

Qualitative mathematical modelling depicts the stressor conceptual model as a sign-directed graph wherein the direct effects between model variables are depicted as arcs ending in an arrowhead for positive direct effects, and arcs ending in a filled circle for negative direct effects (Section 3.1.1). The potential direct effect of coal resource development is identified by at least one negatively or positively signed arc between a hydrological response variable and one or more of the model’s other variables.

Within the sign-directed graph, a pathway with one arc is formally defined as a ‘direct effect’, whereas a pathway consisting of two or more arcs, which must therefore involve other intermediate variables, is formally defined as an ‘indirect effect’ (Section 3.1.1). The direct impacts of coal resource development are thereby depicted within the sign-directed graph as arcs of length one between a hydrological response variable and another model variable. Indirect impacts are identified as arcs of length two or more between a hydrological response variable and another model variable. Finally, cumulative impacts are depicted by changes to two or more potentially interacting hydrological response variables, with concomitant direct and indirect impacts arising from this.

When considering potential changes (perturbations) to the hydrological response variables that maintain landscape class ecosystems, it is important to distinguish between ‘press’ and ‘pulse’ perturbations. A press perturbation is defined as a sustained, or long-term, change in the value of a biological or physiochemical parameter that is associated with one or more variables that causes a shift in the equilibrium values of the system’s variables. This is in contrast to a pulse perturbation, which is a sudden increase or decrease in a variable that only moves the system away from its equilibrium temporarily, but does not necessarily result in a permanent shift to a new equilibrium state. Press-type perturbations are typically defined in terms of experimental manipulations of ecosystems (Dambacher et al., 2002; Schmitz, 1998). In a BA qualitative model, press perturbations are caused by CSG or coal mine induced changes to hydrological response variables. It is acknowledged that pulse perturbations and other categories of disturbance, such as ramp disturbance (Lake, 2000), are also of interest to BA. Press perturbations and other perturbation types may be assessed by the quantitative receptor impact model methodology described in the following chapters.

The distinction between press and pulse perturbations is important because the direct and indirect effects of pulse perturbations are typically minimal or short-lived. These perturbations typically occur over time frames that are much smaller than the generation time of the biological variables of interest. Unless the magnitude of the perturbation is so large that the system is moved to a new equilibrium state, there will be no permanent shift in the equilibrium level of the system variables, but only transient variations in their levels of abundance.

The Bioregional Assessment Programme, however, is principally concerned with press perturbations – that is, changes in hydrological response variables that are sustained for a relatively long time period, say many decades – and hence over much larger time frames than the generation times of the potentially impacted biological variables. The sustained nature of this type of perturbation provides time for the knock-on effects to be felt throughout the entire system. Although the quantitative receptor impact modelling strategy may account for both press and pulse perturbations (see following chapters), the implementation of the qualitative mathematical modelling focuses on press perturbations, and deliberately describes potential impacts on hydrological response variables in terms of changes in their long-term (e.g. 30 years) averages. Some pulse perturbations (e.g. the failure of a tailings dam) are assumed to be adequately managed by site-based risk management and mitigation procedures and are not assessed further in BAs.

This approach to defining and identifying direct, indirect and cumulative impacts is consistent with, and operationalises, the definitions in the BA methodology, and importantly provides a transparent platform for a coherent probabilistic analysis of direct effects (cumulative or otherwise) together with a qualitative mathematical analysis of indirect effects where relevant (see below).

# 3.1 Methods for qualitative mathematical modelling

The following section contains a brief overview of qualitative mathematical modelling based on the construction and analysis of sign-directed graphs (or signed digraphs) and also equivalent approaches using matrix algebra. The signed digraph models in this section are very simple models presented only for pedagogical purposes. Applied BA examples can be found in product 2.7 (receptor impact modelling) (Figure 3).

## 3.1.1 Signed digraph methods for qualitative mathematical modelling

### *3.1.1.1 System structure and signed digraphs*

Qualitative mathematical modelling proceeds by first determining system structure, which is defined by the variables of the system and the relationships by which they are linked (Puccia and Levins, 1985). In biological systems, variables are typically interacting populations of different species, and their dynamics can be accounted for by generalised Lotka–Volterra equations, where each contributes towards the birth or death of another. Similarly, the dynamics of human social and economic systems can be described by the interactions of different sectors and entities of society that control flows of resources, goods and services.

The variables and relationships in a model system are portrayed by sign-directed graphs, or signed digraphs, where a link from one variable to another ending in an arrow (→) represents a positive direct effect, such as a rate of birth that is increased by consumption of prey, and a link ending in a filled circle (—●) represents a negative direct effect, such as a rate of death due to predation. All possible ecological relationships can thus be described:

- predator–prey or parasitism (●→)
- mutualism (↔)
- commensalism (→)
- interference competition (●—●)

- amensalism (—●).

Self-effects are shown by links originating and ending in the same variable, and are typically negative, as in self-regulated variables, but can also be positive (↻), where variables are self-enhancing. Usually, relationships between species can be restricted to being linear (e.g. Lotka–Volterra predator–prey interactions), although non-linear relationships (e.g. Holling functional responses) can also be incorporated. Here, the sign of a relationship can change when the system passes a threshold. This then leads to construction and analysis of alternative model structures (Dambacher and Ramos-Jiliberto, 2007).

Once the structure of a system is defined, it is possible to analyse the system's feedback which determines the qualitative conditions for system stability and perturbation response. These methods can proceed via analysis of the signed digraph through graphical algorithms or through equivalent algebraic analyses of the system's community matrix. As an illustrative example, the Assessment team consider a signed digraph of a predator–prey community of size $\mathrm{n} = 3$, where the top predator is an omnivore that feeds on the basal species (Figure 5):

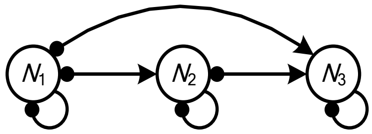

**Figure 5 A signed digraph of a predator–prey community of size $\mathrm{n} = 3$, where the top predator $\mathrm{N}_3$ is an omnivore that feeds on the two other prey species**

### 3.1.1.2 System stability

The stability of a system can be judged and understood according to criteria that depend on the relative sign and balance of the system's feedback cycles (Puccia and Levins, 1985; Levins, 1974, 1975, 1998; Dambacher et al., 2003b). In general, stability requires that the net feedback in a system is negative and that feedback at lower levels is stronger than feedback at higher levels in the system. Negative feedback ensures that a system's dynamics are self damped, and stronger feedback at lower levels ensures that a system will not over correct and exhibit unrestrained oscillations. The conditions can be interpreted through specific algebraic arguments. In general, though, stability in this system depends on the relative weakness of feedback cycles involving omnivory. Here, the feedback cycle $+a_{3,1}a_{1,2}a_{2,3}$ has the potential to destabilise the system through positive feedback, and the feedback cycle $-a_{2,1}a_{1,3}a_{3,2}$, even though it is negative in sign, has the potential to introduce excessive higher level feedback if it is too strong.

### 3.1.1.3 Perturbation response

The signed digraph (or its matrix equivalent) provides a mechanism to predict qualitatively how species abundance or biomass in the system as a whole change as a result of a sustained change to the rate of birth, death or migration in any one of the species (Puccia and Levins, 1985; Levins, 1974, 1975, 1998; Dambacher et al., 2002). As an example perturbation scenario, consider a positive input to $\mathrm{N}_2$, such as food supplementation that increases its rate of birth. The qualitative effect of this input to the other variables is determined by accounting for all of the feedback cycles of length $n - 1$ that emanate from $\mathrm{N}_2$. This is accomplished by tracing all paths from the input variable to a responding variable and multiplying each path by its complementary subsystem, the resulting product is defined as a feedback cycle. The complementary subsystem is defined by the feedback of the variables not on the path from the input to the response variable. If the sign of this subsystem's feedback is positive then it will switch the sign of the path to the response variable, otherwise the sign of the path will be unchanged. The signed digraphs below illustrate the formation of feedback cycles that are used to predict perturbation response. All links that enter the input variable and all links leaving the response variable have been removed; products of the remaining links then become the feedback cycles, which determine the sign of the response. For response of $\mathrm{N}_1$, feedback cycles will be composed of the following links (Figure 6):

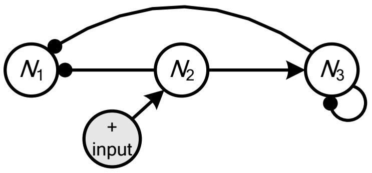

Figure 6 A signed digraph of a predator–prey community of size $\mathrm{n} = 3$, showing a perturbation scenario where there is a positive input to $\mathrm{N}_2$

Here two feedback cycles determine the sign of the response of $\mathrm{N}_1$ due to an input to $\mathrm{N}_2$. One feedback cycle, $-a_{1,2}a_{3,3}$, is formed by a path which goes directly from $\mathrm{N}_2$ to $\mathrm{N}_1$, and it has a complementary subsystem in the negative self-effect of $\mathrm{N}_3$. The other cycle, $-a_{3,2}a_{1,3}$, is composed of a path with negative sign of length two. This path lacks a complementary subsystem, in which case the sign of the path remains negative. Since both feedback cycles are negative, the equilibrium abundance of $\mathrm{N}_1$ is predicted to decrease as a result of supplementation of $\mathrm{N}_2$.

Next, consider the response of $\mathrm{N}_3$ when there has been a negative input to $\mathrm{N}_2$, say through an increased rate of death, and note that for negative inputs the signs of the feedback cycles are switched. The sign of the response of $\mathrm{N}_3$ is determined by the following links (Figure 7),

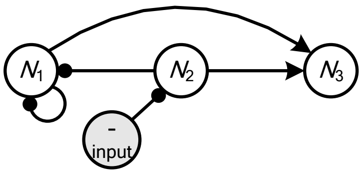

Figure 7 A signed digraph of a predator–prey community of size $\mathrm{n} = 3$, showing a perturbation scenario where there is a negative input to $\mathrm{N}_2$

which form feedback cycles $+a_{1,2}a_{3,1}$and $-a_{3,2}a_{1,1}$. Here the response is ambiguous, as it is determined by feedback cycles of opposing sign.

### 3.1.1.4 Ambiguity and weighted predictions

The ambiguity in the response of $\mathrm{N}_3$ can be resolved through consideration of symbolic inequalities. For instance, if it is believed that $a_{3,2}a_{1,1} > a_{1,2}a_{3,1}$, then the predicted response of $\mathrm{N}_3$ will be negative. Dealing with ambiguity in this manner requires a knowledge of relative strength of interaction and an ability to make sense of contingencies presented by symbolic arguments. However, in larger systems, complex inequalities can arise, which are too difficult to interpret or comprehend. In these instances the Assessment team can employ a heuristic technique of weighting the net number of feedback cycles to the absolute number in a response (i.e. the *weighted prediction* for a response prediction is equal to the net number of feedback cycles divided by the total number of cycles (Dambacher et al., 2002)).

For instance, the predicted response of $N_3$ for an input to $N_2$ is completely ambiguous, as there is the same number of positive and negative feedback cycles. But if there were, say, a total of four feedback cycles in a perturbation response, three of which were positive and one negative, then the net number of cycles would be two and the weighted prediction of the response would be 2/4 = 0.5. The sign determinacy of responses with weighted predictions ≥0.5 has been shown to generally be greater than 90% through simulations using random parameter space (Dambacher et al., 2003a; Hosack et al., 2008); below this threshold the sign determinacy of responses declines to zero for weighted predictions equal to zero.

### 3.1.2 Matrix algebra methods for qualitative mathematical modelling

The preceding section presents an overview of qualitative modelling based on a general interpretation of signed digraphs, this section presents an equivalent analysis based on a system's community matrix, as derived from a system of equations. Computer programs for these analyses can be found in the most recent revision of Supplement 1 of Dambacher et al. (2002).

Rates of growth for *n* interacting populations can be described by a system of linear equations:

$$\frac{\mathrm{d}N_i}{N_i\mathrm{d}t} = \sum_{j=1}^{n} \alpha_{ij} N_j + \beta_i - \delta_i + \iota_i - \varepsilon_i \quad (i = 1, \ldots, n) \tag{1}$$

where the per capita rate of change in population abundance of $\mathrm{N_i}$ is controlled by density-independent rates of birth $(\beta_i)$, death $(\delta_i)$, immigration $(\iota_i)$ and emigration$(\varepsilon_i)$, and $\alpha_{ij}$ density-dependent interactions. Generally stated as: $e_1$

$$\frac{\mathrm{d}N_i}{N_i\mathrm{d}t} = g_i\left(N_1, N_2, \ldots, N_n; p_1, p_2, \ldots, p_m\right) \tag{2}$$

the growth function of each population $(\mathrm{g_i})$ is determined by the system's variables $(\mathrm{N_i})$ and growth rate parameters $\left(\alpha_{ij}, \beta_i, \delta_i, \iota_i, \varepsilon_i\right)$; the latter can be collectively referred to as a vector p of length m. At equilibrium, population abundances (*Ni**) are unchanging and defined by:

$$\mathrm{d}N_i / N_i\mathrm{d}t\big|_{N^*} = 0 \tag{3}$$

In the neighborhood of an equilibrium point, density-dependent interactions, formally organised in the community matrix **A** (Levins, 1968), determine the balance between the column vectors of

$N^*$ (the equilibrium population abundance), and **k** (the density-independent rates of growth), via $AN^* = -k$, where $\mathrm{k_i} = \beta_\mathrm{i} - \delta_\mathrm{i} + \iota_\mathrm{i} - \varepsilon_\mathrm{i}$. Elements of the community matrix are calculated as:

$$a_{ij} = \left.\frac{\partial\left(\frac{\mathrm{d}N_i}{N_i\mathrm{d}t}\right)}{\partial N_j}\right|_{N^*} = \left.\frac{\partial g_i}{\partial N_j}\right|_{N^*} \tag{4}$$

and essentially define the relationships or direct effects between system variables. Alternately, consider the system's Jacobian matrix (**A′**), which is derived from the $\mathrm{dN_i/d_t}$ form of Equation (2) (i.e. $\mathrm{a'_{ij}} = \partial(\mathrm{dN_i/dt})/\partial \mathrm{N_j}|_{\mathrm{N^*}} = \partial(\mathrm{N_i g_i})/\partial \mathrm{N_j}|_{\mathrm{N^*}}$). The sign structure of the community and Jacobian matrices are identical, and thus for the purpose of qualitative modelling, either one can be used. Formal quantitative stability analyses, however, require use of the Jacobian matrix, although calculations of perturbation response can proceed with either the community or Jacobian matrix.

As previously discussed, the relationships between variables can be portrayed by signed digraphs, and for our example system with omnivory (Figure 5), the corresponding community matrix:

$$\mathbf{A} = \begin{bmatrix} -a_{1,1} & -a_{1,2} & -a_{1,3} \\ +a_{2,1} & -a_{2,2} & -a_{2,3} \\ +a_{3,1} & +a_{3,2} & -a_{3,3} \end{bmatrix} \tag{5}$$

is equivalent in information content to the signed digraph.

A qualitative analysis of local or Lyapunov stability is based on the sign of the eigenvalues ($\lambda$) of **A**, which are the roots of the characteristic equation formed by the equality:

$$\det(\mathbf{A} - \lambda\mathbf{I}) = 0 \tag{6}$$

where **I** is the identity matrix and 'det' is the matrix determinant. The resulting characteristic polynomial is of the form:

$$\mathrm{a_0}\lambda^\mathrm{n} + \mathrm{a_1}\lambda^{n-1} + \mathrm{a_2}\lambda^{\mathrm{n-2}} + \ldots + \mathrm{a_n} = 0 \tag{7}$$

where $\mathrm{a_0}, \mathrm{a_1}, \mathrm{a_2}, \ldots, \mathrm{a_n}$, are the system's polynomial coefficients. A system can be stable if and only if all eigenvalues of **A** have negative real parts:

$$\mathrm{Re}\,\lambda_i\,(\mathbf{A}) < 0\; i = 1, 2, 3, \ldots, n. \tag{8}$$

For systems with five or more variables, an explicit solution of the roots is not possible, however, conditions for Equation (8) can be ensured through a series of *n* Hurwitz determinants constructed from the system's polynomial coefficients:

$$\Delta_1 = a_1, \quad \Delta_2 = \begin{vmatrix} a_1 & a_3 \\ a_0 & a_2 \end{vmatrix}, \ldots, \quad \Delta_n = \begin{vmatrix} a_1 & a_3 & a_5 & a_7 & \cdots & a_{2n-1} \\ a_0 & a_2 & a_4 & a_6 & \cdots & a_{2n-2} \\ 0 & a_1 & a_3 & a_5 & \cdots & a_{2n-3} \\ 0 & a_0 & a_2 & a_4 & \cdots & a_{2n-4} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \cdot & \cdot & \cdot & \cdot & \cdot & a_n \end{vmatrix} \tag{9}$$

Additionally, a necessary, but not sufficient condition for Equation (8) is that the polynomial coefficients have the same sign. The value of $a_0$ is arbitrarily +1 or –1, and for the convention of equating negative feedback with self damping it can be set equal to –1. This permits an alternative restatement of the conditions for stability via the criteria that: (i) polynomial coefficients $a_0, a_1, a_2, \ldots, a_n$ are all negative, where $a_0 = -1$; and, (ii) Hurwitz determinants $\Delta_2, \Delta_3, \Delta_4, \ldots, \Delta_{n-1}$ all are positive, where $a_0 = +1$.

By these two criteria it is possible to understand stability in terms of the sign and balance of a system feedback (Puccia and Levins, 1985; Dambacher et al., 2003b). The polynomial coefficients describe the feedbacks *F* at each of *n* levels in a dynamical system, such that $a_0 = F_0 = -1$, $a_1 = F_1$, $a_2 = F_2, \ldots, a_n = F_n$. The first criterion requires that feedback at each level be negative such that the system's dynamics are self damped. Feedback cycles at the first level of a system are composed of the variable's self-effects. For the example system, there are three feedback cycles at the first level: $F_1 = -a_{1,1} - a_{2,2} - a_{3,3}$. In higher levels of the system, feedback cycles are products of either conjunct or disjunct links. At the second level there are six feedback cycles of length two that are products of either pairwise predator–prey links or the disjunct self-effects. All of these cycles have a negative sign: $F_2 = -a_{2,2}a_{3,3} - a_{2,3}a_{3,2} - a_{1,1}a_{3,3} - a_{1,1}a_{2,2} - a_{2,1}a_{1,2} - a_{3,1}a_{1,3}$. Feedback at the highest level of the system is: $F_3 = -a_{1,1}a_{2,2}a_{3,3} - a_{1,1}a_{2,3}a_{3,2} - a_{2,1}a_{1,2}a_{3,3} - a_{2,1}a_{1,3}a_{3,2} - a_{3,1}a_{1,3}a_{2,2} + a_{3,1}a_{1,2}a_{2,3}$. Here, there is one positive feedback cycle that involves the benefit that the top predator receives from omnivory. Stability in this system thus depends on the condition that the strength of this positive feedback cycle be less than the sum of the five negatively signed cycles.

The second stability criterion can be generally understood as a balance between higher and lower levels of feedback. A system that is dominated by higher level feedbacks can exhibit over correction and thus be prone to undamped oscillations. For a three-variable system the second criterion is:

$$\Delta_2 = F_1F_2 + F_3 > 0 \tag{10}$$

which for our example system requires $a_{2,1}a_{1,3}a_{3,2} < a_{2,2}a_{3,3}^2 + a_{3,3}a_{2,3}a_{3,2} + a_{1,1}a_{3,3}^2 + 2a_{1,1}a_{2,2}a_{3,3} + a_{3,3}a_{3,1}a_{1,3} + a_{2,2}^2a_{3,3} + a_{2,2}a_{2,3}a_{3,2} + a_{1,1}a_{2,2}^2 + a_{2,2}a_{2,1}a_{1,2} + a_{1,1}^2a_{3,3} +$

$a_{1,1}^2 a_{2,2} + a_{1,1} a_{2,1} a_{1,2} + a_{1,1} a_{3,1} a_{1,3} + a_{3,1} a_{1,2} a_{2,3}$. Here stability depends on the feedback cycle with the negative effect that the basal species receives from omnivory be less than the sum of all other feedback cycles of length three and the products of lower level feedbacks. Stability in this system thus depends on the relative weakness of the omnivory interaction. Moreover, it can be inferred that if omnivory did not occur, then there would be no conditions for instability, in which case the system would be sign stable.

Any long-term impact on an ecosystem can be interpreted and evaluated as a sustained change in one of the system's $\mathrm{p_h}$ growth parameters. This is the case whether the change comes from within the system via a density-dependent parameter, as in Mendelian selection, or externally by way of a density-independent parameter, as in change coming from the environment or change from management, development or experimental purpose. The Assessment team is interested, then, in predicting change in the equilibrium level of each variable, and, through the implicit function theorem, the solution is by differentiation of Equation (2) with respect to $\mathrm{p_0}$:

$$\frac{\partial \mathbf{N}^*}{\partial p_h} = -\mathbf{A}^{-1} \frac{\partial \mathbf{g}}{\partial p_h} \tag{11}$$

Given the matrix equality:

$$\mathbf{A}^{-1} = \frac{1}{\det(\mathbf{A})} \operatorname{adj}(\mathbf{A}) \tag{12}$$

where 'det' is the matrix determinant and 'adj' is the adjoint matrix, also known as a classical adjoint matrix, Equation (11) can be expressed more conveniently as:

$$\partial \mathbf{N}^* = \underbrace{\frac{1}{\det(-\mathbf{A})}}_{\substack{\text{overall}\\ \text{feedback}}} \underbrace{[\operatorname{adj}(-\mathbf{A})]}_{\substack{\text{complementary}\\ \text{feedback}}} \underbrace{\left(\frac{\partial \mathbf{g}}{\partial p_h}\right)\partial p_h}_{\substack{\text{strength of}\\ \text{input or}\\ \text{perturbation}}} \tag{13}$$

Here is, via Cramer's Rule (Lay, 2003), the solution for $\partial \mathrm{N}^*$ (the difference between the old and new equilibrium abundance for each population), and $(\partial \mathbf{g}/\partial p_h)\partial p_h$ (the strength or magnitude of a given input or perturbation). For the example system, the solution for the effects of parameter change will be:

$$\partial\mathbf{N}^{*}=\frac{\begin{matrix} a_{2,2}a_{3,3}+a_{2,3}a_{3,2} & -a_{1,2}a_{3,3}-a_{1,3}a_{3,2} & a_{1,2}a_{2,3}-a_{1,3}a_{2,2} \\ a_{2,1}a_{3,3}-a_{2,3}a_{3,1} & a_{1,1}a_{3,3}+a_{3,1}a_{1,3} & -a_{1,1}a_{2,3}-a_{1,3}a_{2,1} \\ a_{2,1}a_{3,2}+a_{2,2}a_{3,1} & a_{1,1}a_{3,2}-a_{1,2}a_{3,1} & a_{1,1}a_{2,2}+a_{2,1}a_{1,2} \end{matrix}}{a_{1,1}a_{2,2}a_{3,3}+a_{1,1}a_{2,3}a_{3,2}+a_{2,1}a_{1,2}a_{3,3}+a_{2,1}a_{1,3}a_{3,2}+a_{3,1}a_{1,3}a_{2,2}-a_{3,1}a_{1,2}a_{2,3}}\left(\frac{\partial\mathbf{g}}{\partial p_h}\right)\partial p \qquad (14)$$

From Equation (14) it is evident the matrix **A** has two separate functions in determining a population's response to a perturbation. Through the adjoint of **A**, all direct and indirect effects in the system are combined in complementary feedback cycles (Dambacher et al., 2002), which mediate the relative variation in the response of each population. The determinant of **A** constitutes the overall feedback of the system and scales the magnitude of each variable's response to a perturbation. For example, for a given input, if the overall feedback is relatively weak, then the effect of the complementary feedback cycles on population abundance will be relatively large. For this system to be stable requires that det(–**A**) be positive, and from the above stability analysis, this requires that $a_{3,1}a_{1,2}a_{2,3}$ be relatively weak. Note that use of the determinant of –**A** maintains a sign convention in the adjoint matrix for even- and odd-sized systems (Dambacher et al., 2002). However, as previously mentioned, system feedback is normally considered to be negative in stable systems; this convention is reversed in Equation (13) and Equation (14) to analyse perturbation response.

In interpreting the adjoint matrix, a positive input to a variable – through an increase in birth rate or decrease in death rate – is read down a column, and response predictions for each variable are read along the rows. Here the correspondence between the matrix adjoint matrix in Equation (14) with the signed digraph analysis in the main text is seen. For a positive input to $\mathrm{N}_2$ the prediction of decreased abundance of $\mathrm{N}_1$ is determined by two feedback cycles of like sign (i.e. adjoint $(-\boldsymbol{A})_{1,2}=-a_{1,2}a_{3,3}-a_{1,3}a_{3,2}$), while the ambiguous response of $\mathrm{N}_3$ is determined by two feedback cycles of opposing sign (i.e. adjoint$(-\boldsymbol{A})_{3,2}=-a_{1,1}a_{3,2}-a_{1,2}a_{3,1}$). Where inputs to a variable are negative – through a decrease in birth rate or increase in death rate – then the signs of the adjoint matrix elements are simply reversed. In analysis of linear systems, multiple inputs will have an additive effect on the equilibrium of a variable through the superposition principle. Thus if there were simultaneous inputs that increased the birth rate of $\mathrm{N}_1$ (e.g. by supplementation) and increased the death rate of $\mathrm{N}_2$ (e.g. by culling), then the predicted response of $\mathrm{N}_3$ would be calculated as:

$$\partial N_3^{*}=\frac{1}{\det(-\mathbf{A})}\left(\left(a_{2,1}a_{3,2}+a_{2,2}a_{3,1}\right)\frac{\partial g_1}{\partial\beta_1}\partial\beta_1-\left(a_{1,1}a_{3,2}-a_{1,2}a_{3,1}\right)\frac{\partial g_2}{\partial\delta_2}\partial\delta_2\right) \qquad (15)$$

# 4 Identification of hydrological response variables and receptor impact variables (stage 3)

Receptor impact models translate predicted changes in the hydrological regimes that help maintain landscape classes into predicted changes in one or more ecosystem variables. The regional landscape-level scope of the BA, however, provides for many choices during this process, whilst the complexity of the assessment combined with the time and resource limitations imposed on expert participation place many restrictions on what can be practically achieved within the time available for the BA. This section presents the set of criteria and processes that can be used to progress cumulative impact assessments for receptor impact variables.

The two most important choices during the receptor impact modelling are:

- *hydrological response variables*. The BA groundwater and surface water models could be used to predict many different characteristics of the hydrological regimes that operate in a landscape class, but these will not be equally important in maintaining and shaping the ecosystem. The choice of hydrological response variables to include in the receptor impact models should be guided by the results of the IMEA hazard analysis (see companion submethodology M11 (as listed in Table 1) for hazard analysis (Ford et al., 2016)), the capabilities of the surface water and groundwater models, and most importantly the hydrological variables identified by ecological experts when the signed digraphs of the landscape class are created.
- *receptor impact variables*. Typically, there are hundreds to thousands of assets identified within a BA bioregion or subregion and tens of variables identified within the sign-directed graphs of the landscape classes, with ecological assets either explicitly recognised as a model variable, or encompassed within a broader functional group. For example, a particular species of bird (the asset) may be encompassed within a function group of nocturnal predators. Given that expert participation is required to construct each receptor impact model, by necessity the receptor impact model process can only construct a few receptor impact models per landscape class, which accommodates no more than a few receptor impact variables per landscape class.

The choice of hydrological response variables ultimately will be a compromise between verbal descriptions of the hydrological regimes given by ecologists during the construction of the landscape class signed digraphs, and the numerical indices that are produced by the surface water and groundwater models. In this context it is important to recognise that the outputs of the surface water and groundwater models are typically finalised before the qualitative modelling workshops are conducted, and hence the model outputs are not specifically designed to address the hydrological regimes that are subsequently described. In most instances, however, the surface water and groundwater outputs are either directly relevant, or provide a sufficient basis to construct, the hydrological variables identified and verified by the expert ecologists as critical to the landscape class ecosystems.

The choice of receptor impact variables is framed by the complexities of the potential direct and indirect effects associated with press perturbations and the BA's expected audience. There will be a wide range of interests represented across the Independent Expert Scientific Committee on Coal Seam Gas and Large Coal Mining Development (IESC, after 27 November 2012), regulators, industry and community. There will be experts in particular taxa and experts about particular locations. There will be experts about particular agricultural or ecological systems. Furthermore, many readers of the assessments will be members of the general public who will each have their own beliefs and understandings of these systems, and an associated set of values.

The Bioregional Assessment Programme (the Programme) and the Assessment team need to be conscious of the expectations of this community, and the natural tension that arises between ensuring that the risk analysis is achievable (within operational constraints) but at the same time relevant to as much of the community as possible. This can be achieved, for example, by choosing impact variables that simplify the elicitation task, by speaking to broad sections of the community, and by using narrative in the analysis to broaden the relevance of the assessment to other more specific sections of the community. For example, basal vegetation variables are relevant to diverse segments of the expected audience when interpreted in terms of forest cover and such.

To navigate the choice across the large set of possibilities, the BA approach is to adopt a selection process that is efficient and consistent across the Programme, and leads to effective and efficient communication to stakeholders. For example, choosing a cryptic species with unknown relationships to overall environmental health would be inefficient and potentially misleading. Impact variables that are identified as representative of core stakeholder and community values within that bioregion or subregion will be central to the assessment. Impact variables that are more directly and unambiguously connected to the hydrological response, and that are measurable in some sense will lead to more certain and useful receptor impact models. Impact variables that strongly speak to the condition or abundance of other components of the system are also clearly advantageous. For example, if the status of river red gums also informs about the health of other floodplain trees and organisms, then it is a good choice.

To assist in these considerations the BA developed the following selection criteria to help guide the choice of impact variable:

- *Is the response variable directly affected by changes in hydrology?* These variables typically have a lower trophic level, and focusing on direct (signed digraph arcs of length one) impacts helps alleviate the elicitation burden imposed on experts during the construction of the receptor impact models.
- *Is its status important in maintaining other parts of the landscape class?* Variables (or nodes) within the qualitative model that other components of that ecosystem or landscape class depend on will speak more broadly to potential impacts. Again, these types of variables will typically have a lower or mid-trophic level.
- *Is it something that the available expertise can provide an opinion on?* There is a need to be pragmatic and make a choice of receptor impact variable that plays to the capabilities and

knowledge base of the experts that are available at the time the receptor impact models are created.

- *Is it something that is potentially measurable?* This is essential for (in)validation of the predicted impacts and in the subsequent design of monitoring strategies that close the risk analysis loop by testing, comparing its predictions with observations.
- *Will the community understand and accept the relevance and credibility of the receptor impact variables for a given landscape class?* This reflects the communication value of the receptor impact variable.

The challenge in applying these selection criteria is that a priori it is unlikely that any single impact variable will simultaneously satisfy all criteria. Examples of ecological variables that are likely to perform well when judged against these criteria include: the percent cover of a dominant canopy-forming species or plant association, the abundance or biomass of an iconic species, or the abundance/biomass of a particular key functional group (e.g. invertebrate food source for waterbirds and other predators).

The importance of a dominant canopy-forming species or plant association is reflected in its key role in forming habitat for other species. It is also likely to have a low trophic level and therefore be directly impacted by changes in hydrological variables and a major change in the component of the ecosystem will likely be noticed by the non-expert and casual observer alike. Iconic species are recognisable to large numbers of people and would reflect community concerns about particular changes but some life stages may be relatively insensitive to hydrological change. Changes in the abundance or biomass of key prey groups (such as aquatic invertebrates) are again likely to be relatively low in the food chain and potentially more proximal to hydrological changes, and also be important in maintaining many other groups within the ecosystem, for example iconic predators such as the platypus (*Ornithorhynchus anatinus*), and hence are more easily aligned with ecological features of the landscape class that are valued by the community.

To illustrate the potential choices, complexities and the implications the choice of impact variable holds for the estimation of direct and indirect effects, the Assessment team can identify three impact scenarios of increasing complexity:

1. Press perturbation involving only a single hydrological response variable, which leads only to direct impacts (i.e. single link in causal pathway). In this most simple scenario, the outputs from the coal resource development pathway (CRDP), the hazard analysis, and signed digraph (stressor conceptual model) predict only direct impacts between an individual hydrological response variable and receptor impact variable. The quantification of the impact proceeds by eliciting the expected change in the receptor impact variable (y-axis in the figure on right-hand side of Figure 8) conditional on the change in hydrological response variable (x-axis in the figure on the right-hand side of Figure 8), via the scenario creation and receptor impact modelling process. Note this scenario is shown here for illustrative purposes only and may never occur in reality.
2. Press perturbation to a single hydrological response variable, leading to direct and indirect impacts. In this more complex situation, the signed digraph may indicate the existence of both direct and indirect causal pathways leading from the hydrological response variable to the receptor impact variable of interest. In Figure 9, the indirect pathways from FR3 to

HF MI involve additional intervening variables. In this example not only is there a choice of receptor impact variable, but, if an interest in HF MI is maintained, the quantification of the direct and indirect effects on HF MI must either be explicitly accounted for in the statistical model (an option not practical given the time and resource limitations imposed on expert participation) or the expert must be asked for the predicted change in HF MI averaging over the possible changes in the intervening variables.

3. Press perturbations to multiple hydrological response variables, leading to direct and indirect impacts on multiple potential receptor impact variables. In this more complex but realistic scenario, the signed digraph identifies direct and indirect impacts on multiple receptor impact variables arising from simultaneous press perturbations to multiple hydrological response variables (Figure 10). The complexity of this process can be described in the scenario creation step prior to the receptor impact modelling, but again there is a choice of receptor impact variable, and, if an interest in variable WRV is maintained, the expert must account for the effects of FR1 and FR2 on Seedl, and their indirect impact on WRV as well as the direct effects of FR1, FR2 and GW on WRV. The statistical model leads to a response surface that allows for interactions and non-linear relationships between the expert response and the hydrological response variables.

By representing stressor conceptual models as sign-directed graphs (see Chapter 3), the BA methodology is able to make qualitative predictions of long-term, press perturbation impacts to variables (i.e. direction but not magnitude of change) in the landscape class stressor conceptual model. This facet of the analysis provides the possibility of comparing the predicted direction of change of the qualitative mathematical analysis with the quantitative analysis of the receptor impact variable, and also provides whole-of-system qualitative predictions for the landscape class that may be informative during the post-assessment monitoring, in addition to the quantitative predictions available from the receptor impact model analyses. The sign-directed graphs were constructed from expert participation within the qualitative modelling workshop (Figure 4).

In conjunction with input from the hydrological modelling, the sign-directed graph in turn advised the design of the quantitative receptor impact modelling workshop (Figure 4). The quantitative receptor impact model predicts the magnitude of the receptor impact variable response given changes in the identified hydrological response variables (Figure 8, Figure 9, Figure 10). The quantitative receptor impact model additionally captures not only long-term responses to press perturbations but also potentially transient responses to pulse perturbations. The quantitative receptor impact model quantifies the uncertainty in the response of the receptor impact variable due to uncertainty in the direct impacts, indirect impacts and other factors or pathways that are not explicitly represented. It is important to recognise that the risk methodology proposed here is capable of quantifying the direct impacts under scenario 1, and direct and indirect impacts under scenarios 2 and 3 including the case where these are mediated by indirect effects, but the elicitation task is made increasingly more difficult through scenarios 2 and 3. As noted in the above discussion, it is a matter of choice on what factors to explicitly include within a quantitative receptor impact model. Whereas including more factors or hydrological response variables may seem more comprehensive, the cost of the increased complexity is that a more elaborate and complicated model that could increase the difficulty of an elicitation is required. For example, fundamental relationships among identified water availability metrics and the receptor impact variable could be obscured by additionally conditioning on intermediary variables.

In practice, the BA seeks to minimise the complexity of the analysis by judicious choice of hydrological response variables and receptor impact variables so that, wherever possible (and acceptable to managers and stakeholders), the analysis is constrained to receptor impact modelling within scenario 1. Where this is not possible, the elucidation of the system framework in the graphical depiction of the sign-directed graph can help the expert allow for different alternative pathways while contributing their assessment.

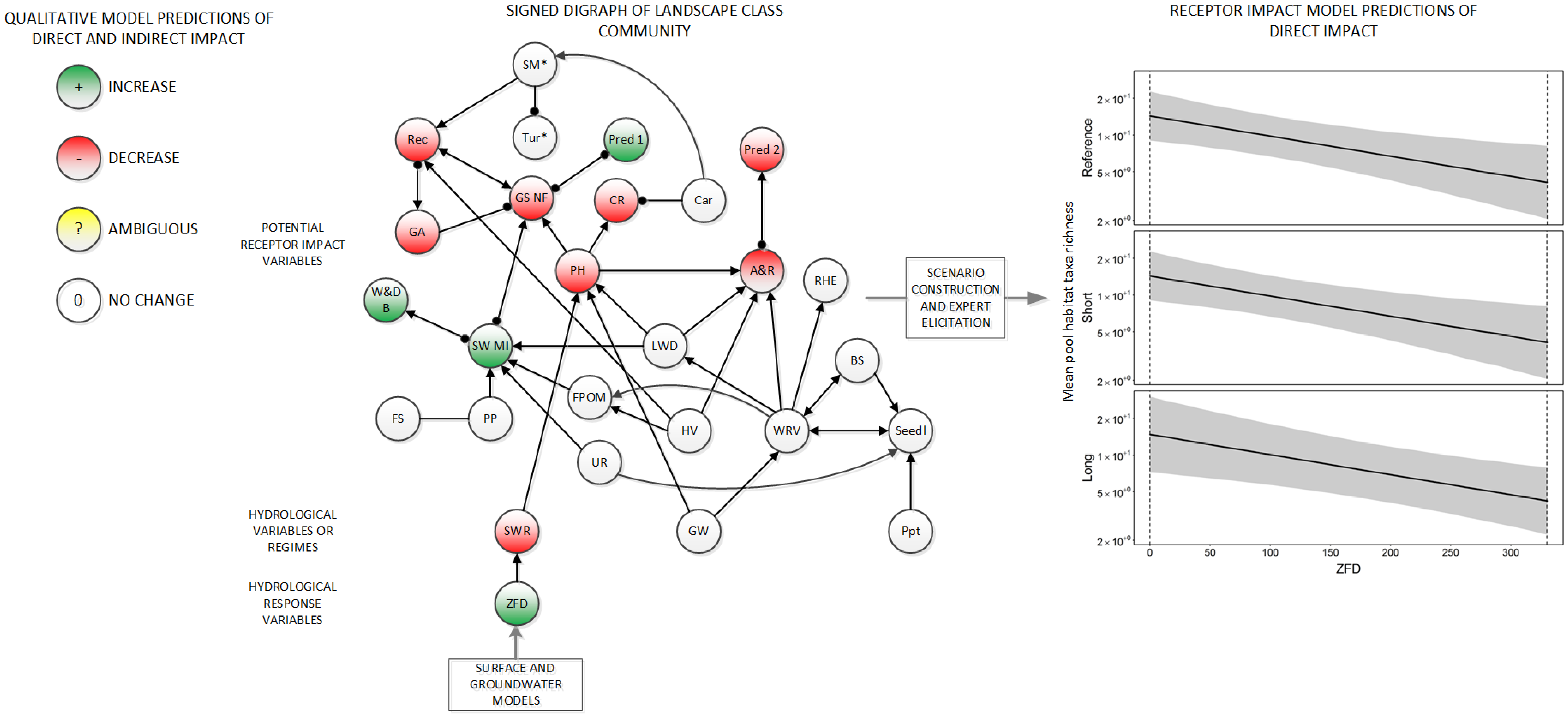

**Figure 8 Schematic showing how direct and indirect impacts are qualitatively and quantitatively assessed within the bioregional assessments for a simple impact scenario**

The schematic shows the simplest impact scenario focused on a particular receptor impact variable within a landscape class community. In this scenario, a decrease in hydrological variable surface water replenishment (SWR) is modelled as an increase in hydrological response variable zero-flow days (averaged over 30 years) (ZQD). Under the stressor conceptual model (represented here by the sign-directed graph) this is deemed to have a direct impact on the receptor impact variable Invertebrate Taxa within Pool Habitat (PH). The signed digraph also provides qualitative predictions (node colours) of the (long-term) knock-on effects through the ecosystem on to other variables. The impact of this direct effect is quantified through the receptor impact model (right-hand panel). The receptor impact model (essentially a generalised linear model with expert opinion as data) predicts how the median value of the receptor impact variable changes in response to changes in the hydrological response variable (HRV; black line in right-hand plots) together with the uncertainty associated with these predictions (grey polygons in the right-hand plot), for the reference (2012), short (2042) and long (2102) assessment years. Note that all self-effects are omitted from the sign-directed graph for clarity. Dashed vertical lines on the right-hand panels show the minimum and maximum HRV values provided to the elicitation team prior to the receptor impact modelling workshops. This scenario is shown for illustrative purposes and may never occur in reality. Readers should refer to the 'Intermittent – gravel/cobble streams' landscape class in companion product 2.7 for the Gloucester subregion (Hosack et al., 2018) for further information on the nodes shown in the signed digraph.

This figure has been optimised for printing on A3 paper (420 mm x 297 mm).

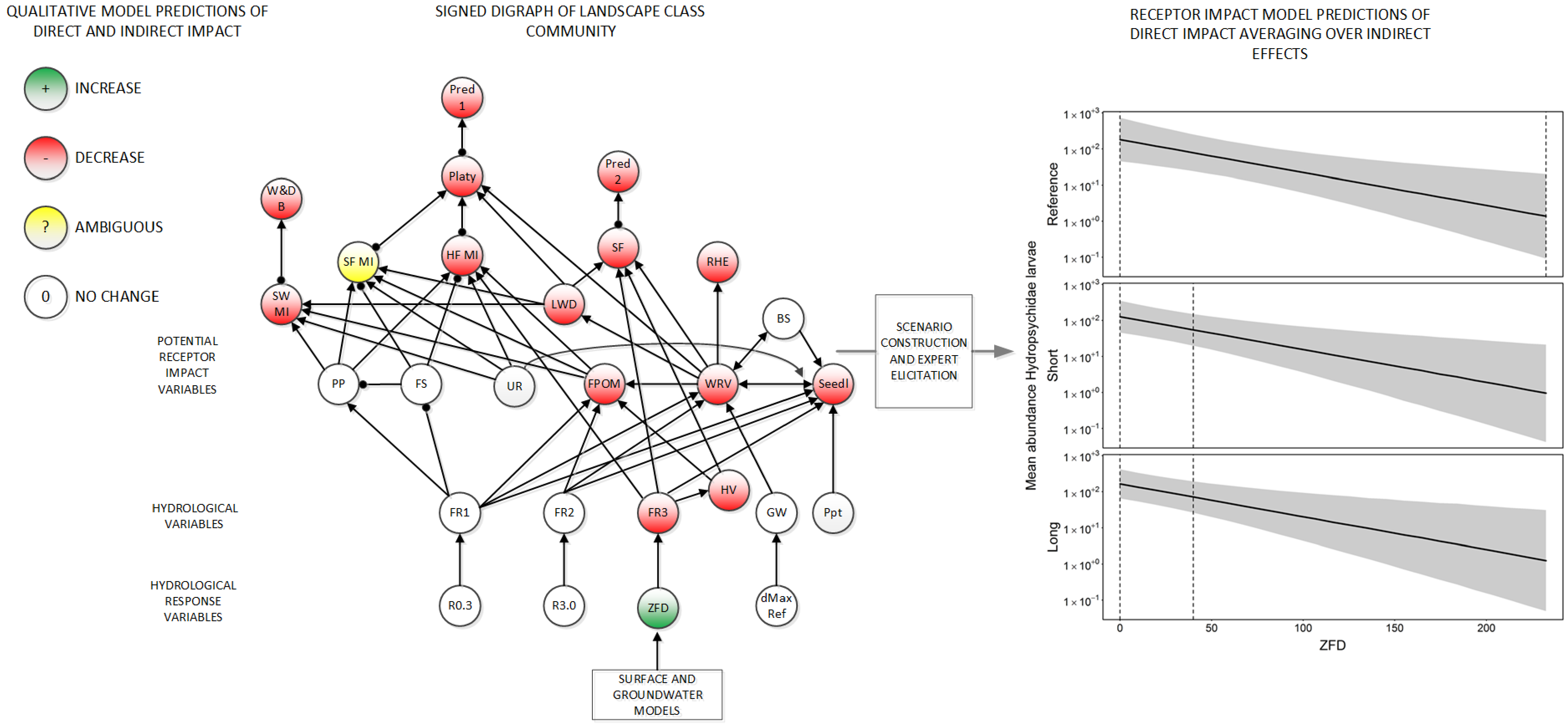

**Figure 9 Schematic showing how direct and indirect impacts are qualitatively and quantitatively assessed within the bioregional assessments for an impact scenario of intermediate complexity**

In this scenario, the development's impact on the hydrological variable flow regime 3 (FR3) is modelled as a decrease in the hydrological response variable zero-flow days (averaged over 30 years) (ZQD) for another landscape class. Under the stressor conceptual model this has a direct impact on receptor impact variable high-flow macroinvertebrates (HF MI), where increasing zero-flow days corresponds to decreasing abundance of HF MI, but it also impacts herbaceous vegetation (HV) which influences fine organic particulate matter (FPOM), and woody riparian vegetation (WRV through seedlings) which provide habitat to platypus (Platy) that consume HF MI. The direct and indirect impact on receptor impact variable (the abundance of Hydropsychidae larvae) caused by changes in ZQD is quantified in the receptor impact model. In this situation the model must either explicitly account for the indirect interactions of ZQD on HF MI as mediated by FPOM and Platy, or the expert must account for the direct and indirect interactions internally during the elicitation by marginalising over the indirect pathways from ZQD as mediated by FPOM and Platy. The receptor impact model (essentially a generalised linear model with expert opinion as data) predicts how the median value of the receptor impact variable changes in response to changes in the hydrological response variable (HRV; black line in right-hand plots) together with the uncertainty associated with these predictions (grey polygons in the right-hand plot), for the reference (2012), short (2042) and long (2102) assessment years. Note all self-effects are omitted from the sign-directed graph for clarity. Dashed vertical lines on the right-hand panels show the minimum and maximum HRV values provided to the elicitation team prior to the receptor impact modelling workshops. Readers should refer to the 'Perennial – gravel/cobble streams' landscape class in companion product 2.7 for the Gloucester subregion (Hosack et al., 2018) for further information on the nodes shown in the signed digraph.

This figure has been optimised for printing on A3 paper (420 mm x 297 mm).

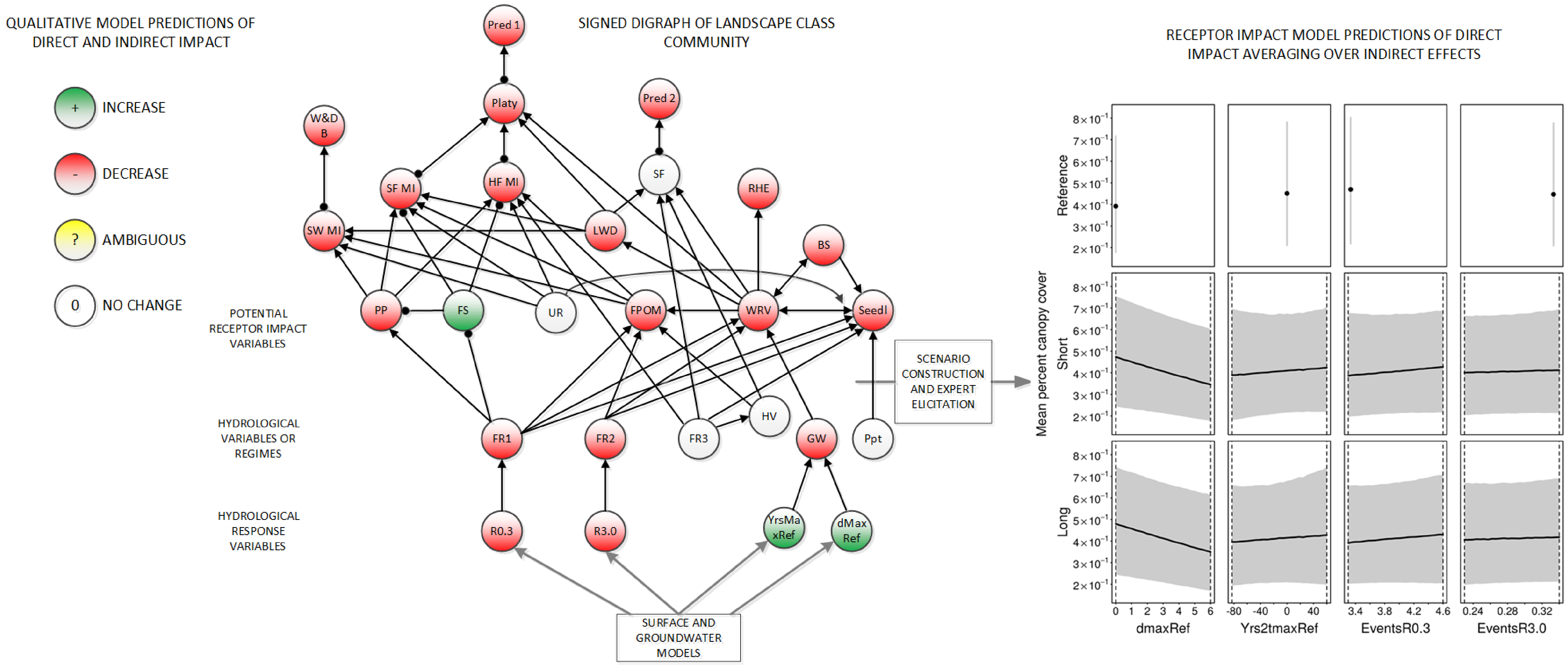

**Figure 10 Schematic showing how direct and indirect impacts are qualitatively and quantitatively assessed within the bioregional assessments for an impact scenario of higher complexity**

In this scenario, the development simultaneously changes hydrological variables flow regime 1 (FR1), flow regime 2 (FR2) and groundwater depth (GW). These changes are modelled as decrease in hydrological response variables Events R0.3 (R0.3) and Events R3.0 (R3.0), and an increase in hydrological response variable maximum groundwater depth relative to the reference period (dMaxRef) and years to maximum groundwater depth (YrsMaxRef). Under the stressor conceptual model this leads to direct impacts on receptor impact variables wood riparian vegetation (WRV), fine sediment (FS), primary production (PP) and fine particulate organic matter (FPOM). To quantify the direct and indirect impacts of the hydrological response variables on WRV in this instance, the model must either explicitly account for the direct effect of FR1 and FR2 on the additional ecological variable seedlings (Seedl) and its indirect interaction with WRV, or the expert is asked to marginalise over the direct and indirect effects during the elicitation. In this case the plot on the right-hand side shows how the median value of the receptor impact variable changes in response to changes in these hydrological response variables (HRV; black line in right-hand plots) together with the uncertainty associated with these predictions (grey polygons in the right-hand plot), holding all other hydrological response variables at their median values, for the reference (2012), short (2042) and long (2102) assessment years. Note all self-effects are omitted from the sign-directed graph for clarity. Dashed vertical lines on the right-hand panels show the minimum and maximum HRV values provided to the elicitation team prior to the receptor impact modelling workshops. Readers should refer to the 'Perennial – gravel/cobble streams' landscape class in companion product 2.7 for the Gloucester subregion (Hosack et al., 2018) for further information on the nodes shown in the signed digraph.

This figure has been optimised for printing on A3 paper (420 mm x 297 mm).

# 5 Development of scenarios for receptor impact model expert elicitation (stage 4)

As described in Section 1.3, it is essential to have an efficient design to collect expert information given the large number of receptor impact models, landscape classes and bioregions or subregions to address with limited resources. The design must also reflect the predicted hydrological regimes as summarised by hydrological modelling outputs. Without this information, design points may present hydrological scenarios that are unrealistically beyond the bounds suggested by the landscape class definition. Alternatively, insufficiently wide bounds on hydrological regimes lead to an overextrapolation problem when receptor impact model predictions are made conditional on hydrological simulations at the risk estimation stage. The design must further reflect the feasibility of the design space, which may be constrained by mathematical relationships between related hydrological response variables. The design must accommodate the requirement to predict to past and future assessment years. The design must also allow for the estimation of potentially important interactions and non-linear impacts of hydrological response variables on the receptor impact variable. Moreover, the predictions of the receptor impact variable in future years may depend on the state of the receptor impact variable in past years. For example, a forest stand may persist over time periods longer than the time frame considered by BA. The past states of the forest stand may have an important influence on future states; this temporal relationship may also interact with hydrological response variables. Therefore, this ecological temporal relationship, which if ignored may confound the relationship between the receptor impact variable and the hydrological response variables, must also be accommodated for by the design. These above requirements must all be met by the design to enable prediction of receptor impact variable response to hydrological changes at the time points and spatial scales required by BA.

For each landscape class, a set of receptor impact variables (e.g. woody riparian vegetation) and a set of hydrological response variables are defined. For each landscape class, different receptor impact variables may not be connected to the same hydrological response variables. The following section describes the general framework of the elicitation approach.

As noted in Chapter 0, in BAs, hydrological response variables are defined as the hydrological characteristics of the system that: (i) are thought to be instrumental in maintaining and shaping the ecosystem components, processes and functions provided by the ecosystems in each landscape class, and (ii) have the potential to change due to coal resource development. Receptor impact variables are defined as the components of the ecosystem that, according to the qualitative mathematical model developed at the preceding qualitative modelling workshop, potentially change due to changes in the hydrological response variables. Receptor impact models describe how changes in hydrological response variables may impact particular aspects of ecological systems at a project-defined spatial and temporal scale. A receptor impact model describes the distribution of outcomes that would be expected to be observed in a receptor impact variable given a particular change to one or more hydrological response variables.

Where possible, receptor impact models would be based on data. In practice, the empirical information is usually incomplete so structured elicitation with experts is used to integrate their knowledge into the models. The quantitative relationship between receptor impact variables and hydrological response variables is the focus of the receptor impact modelling workshop that follows the qualitative modelling workshop in the overall workflow of the BAs (Figure 3).

The proposed approach allows the elicitation to proceed either on magnitude (e.g. abundance or percent cover) or presence–absence responses. If experts prefer to think about presence–absence for some combinations of covariates, but magnitude for other combinations, then both can be accommodated within a single model. The elicitation target is continuous (i.e. mean abundance or probability). Technical details are given in the following sections.

The structure of the elicitation is as follows. Time is indexed using $t_k$ for receptor impact variable responses. Here, $t_k$ refers to a specific year of assessment for the receptor impact variable that corresponds to the pre-specified time points of interest. The complete set of assessment years is given by $\left\{t_{ref} = 2012, t_{short} = 2042, t_{long} = 2102\right\}$, which corresponds to past, short-term future and long-term future assessment time points, respectively. Associated with each assessment year $t_k$ (that is, for $k \in \{ref, short, long\}$) are relevant time periods of hydrological history $\tau_k^w$, which vary depending on whether the histories are with respect to groundwater ($w = G$) or surface water ($w = S$) hydrological response variables. For surface water hydrological response variables, the $\tau_k^S$ are 30-year time periods with end-date $t_k$ (inclusive). For example, the reference period for all surface water hydrological response variables is given by the closed interval, $\tau_{ref}^S = [1983, 2012]$. The groundwater hydrological response variables in the future period are defined relative to the same reference period used for the surface water, $\tau_{ref}^G = \tau_{ref}^S = [1983, 2012]$. The value of the groundwater hydrological response variable depends only on whether or not the assessment year occurs in the reference period or not, and does not depend on whether or not the assessment year occurs in either the short-term or long-term assessment year. For groundwater hydrological response variables, the $\tau_k^G$ for the *short* and *long* periods are therefore the same ($\tau_{short}^G = \tau_{long}^G$); in this case, the hydrological response variable is, for instance, a summary statistic considered over the whole of the future period. An example is the maximum drawdown in the future period relative to the groundwater level in the reference period. The $t_k$ and $\tau_k^w$, are visualised in Figure 11.

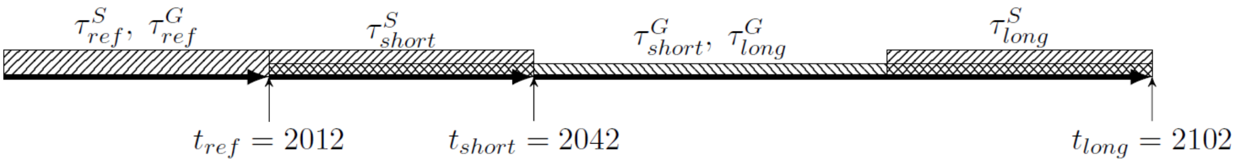

**Figure 11 Visualisation of $t_k$, a specific year of assessment for receptor impact variable response, and $\tau^w_k$, the relevant time periods of hydrological history, that correspond to the pre-specified time points of interest**

The assessment years, given by $t_k$ with $k \in \{ref, short, long\}$, were chosen to assess the potential response of the receptor impact variable early in the development period ($t_{ref} = 2012$), over an intermediate period of development ($t_{short} = 2042$) and also the enduring impact of the developments ($t_{long} = 2102$). The hydrological response variables (for example, the annual mean number of low-flow days over a 30-year period) were defined by a summary statistic derived for the intervals $\tau^w_k$ that depended on whether the hydrological response variable was surface water or groundwater ($w \in \{S, G\}$). The reference period was shared across the hydrology disciplines, $\tau^S_{ref} = \tau^G_{ref} = \tau_{ref} = [1983,\ 2012]$. The surface water hydrological response variable intervals $\tau^S_k$ spanned the 30 years preceding $t_k$. The groundwater hydrological response variable intervals only depended on whether the corresponding assessment year was in the reference or future period; the intervals for the short-term and long-term future assessment years were therefore equivalent for groundwater, $\tau^G_{short} = \tau^G_{long} = \tau^G_{future}$. The groundwater future interval $\tau^G_{future}$ was the only interval that varied depending on bioregion. In this example, $\tau^G_{future} = [2013,\ 2102]$.

Let $y(t_k)$ be a receptor impact variable at time $t_k$. It is assumed that $y(t_k)$ can depend on:

- hydrological response variables derived for the corresponding intervals $\tau^w_k$
- previous values of the receptor impact variable $y(u), \quad u < t_k$.

The goal of the receptor impact modelling workshop is to predict how the receptor impact variable changes at future time points given the reference value of the receptor impact variable and the hydrological response variable values. The Assessment team therefore elicits subjective probability distributions for the receptor impact variable for particular hydrological scenarios from the experts. The probability distributions express uncertainty in the impact of the hydrological change across the region. For example, the canopy cover of woody riparian vegetation may respond to an increase or decrease of a particular flood expected frequency of occurrence in the 30 years preceding the time point of interest. The predicted receptor impact variable may also depend on the level of the receptor impact variable (e.g. canopy cover) in the past. For example, a change in the hydrological response variable at a future time period relative to the reference period may be relevant if starting from a high level of canopy cover, but perhaps less important if starting from a very low canopy cover. The Assessment team therefore conditions the predictions of receptor impact variables on a stated level of the receptor impact variable at the reference assessment year, which is assumed known. However, the receptor impact variable (e.g. percent cover) at the reference assessment year is actually unknown. Therefore, the experts first assess how the receptor impact variable is distributed across the landscape class at the reference assessment year.

Given the above rationale, the receptor impact modelling workshop elicitation targets are defined by the conditional probability distributions of:

- the receptor impact variable value at the reference assessment year 2012 given any hydrological response variables in the reference period
- the receptor impact variable value at a future year given a value at the reference year and the hydrological response variables during the future time period.

The statistical model and elicitation process is summarised in Figure 12. Note that the final estimated receptor impact variable at time $t_{k'}$ depends on the hydrological history prior to $t_{ref}$, but the elicitation has simplified this by conditioning instead on the value of the receptor impact variable at time $t_{ref}$ in the dependence structure of Figure 12. For example, a scenario considered in elicitation 2 (see Figure 12) conditions on the past receptor impact variable ($y(t_{ref})$) instead of the possibly multiple hydrological response variable covariates that correspond to time $t_{ref}$ and period $\tau_{ref}$.

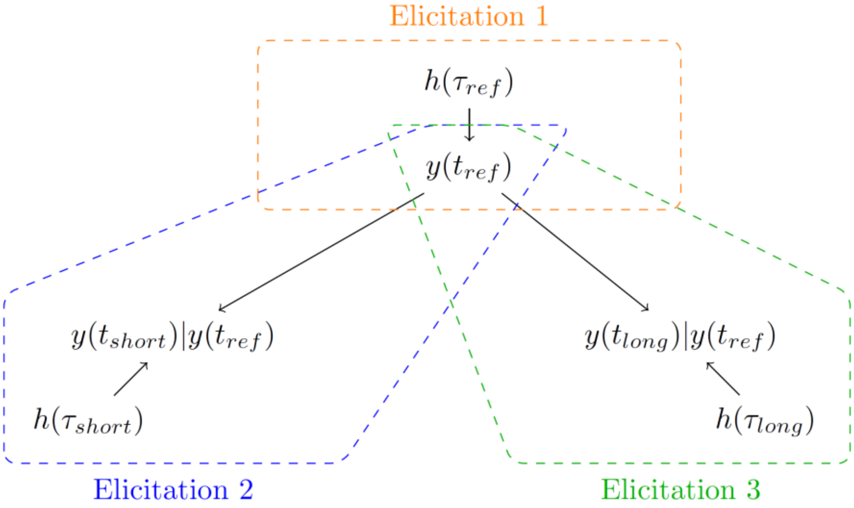

**Figure 12 Elicitation relationship diagram, with the notations used in the related sections**

For any $k' \in \{short, long\}$, the receptor impact variable at time $t_{k'}$ depends on the current hydrological response variables (i.e. in the period $\tau_{k'}$) and the value of the receptor impact variable at time $t_{ref}$.

## 5.1 The general statistical model

The receptor impact model framework specifies a statistical model for the response function parameterised by expert opinion (Hosack et al., 2016, 2017). A modelling approach is adopted that allows for different types of empirical data, $z$, such as abundance, density, quadrat counts or presence–absence using generalised linear models (McCullagh and Nelder, 1989). The observation model for data $z_i$, $p(z_i|y_i, \xi)$, is assumed to be from the exponential family and is conditioned on the expected response, $\mathbb{E}[z_i] = y_i$, with possibly additional parameters $\xi$ that pertain only to the observation model. This expected response is mapped to the linear predictor, $\eta_i$, by an invertible link function, $g(y_i) = \eta_i$. The linear predictor depends on the design point defined by the $p \times 1$ vector of known covariate values $x_i$ and the $p \times 1$ vector of unknown coefficients $\beta$ through the linear function:

$$\eta_i = x_i^{\top} \beta \tag{16}$$

The above assumptions lead to the construction of a prior for the unknown parameters with a generalised linear model with defined observation model, link function and design point $x_i$.

### 5.1.1 Choice of link function and observation models

The generalised linear model allows for a wide variety of potentially observable responses that can be identified within a receptor impact model. Many options of link function and observation model will be available for a given receptor impact model. Guidance is provided here to assist this important selection process. The elicitation approach assumes that counts over the non-negative integers follow a Poisson distribution (with unknown varying intensity) and counts from a finite sample size follow a binomial distribution (with unknown varying probability); these are standard choices from the exponential family for count data. For the Poisson case, two link functions will be required (log and complementary log-log, 'cloglog'), complicating the model fitting. For the binomial case, only the cloglog is required (with and without an offset). For both cases the above approach can be accommodated by the currently available methods for eliciting subjective probability distributions. These choices establish a correspondence among non-negative counts, bounded non-negative counts and presence–absence receptor impact variable models. Although the below choices develop log link models for discrete observations $z$, note that the analogue for non-negative continuous data is given by the Gaussian observation model coupled with log link.

#### *5.1.1.1 Poisson*

Let $p(z_i|y_i) = \frac{y_i^{z_i}\exp{-y_i}}{z_i!}$ with $y_i = g^{-1}(\eta_i) = \exp(\eta_i)$ with mean given by the inverse of the canonical log link function of the linear predictor $\eta_i = \log(y_i) = x_i\beta$ with known covariates $x_i$ and unknown parameters $\beta$. The intensity of the inhomogeneous Poisson process is given by $y_i$.

##### 5.1.1.1.1 Support

The support of $p(z)$ is over the non-negative integers, $z = 0,1,\ldots$. For example, it is applicable to a setting with a countable number of individuals in a given area.

##### 5.1.1.1.2 Elicitation target

The expected value of the response is the average number of individuals in a given area over a specified time period.

##### 5.1.1.1.3 Probability of presence

The probability of observing a zero count is $P(z_i = 0) = e^{-y_i}$. Let $p_i$ equal the probability of presence:

$$\begin{aligned} 1 - p_i &= P(z_i = 0) = e^{-y_i} \\ -\log(1 - p_i) &= z_i \\ \log(-\log(1 - p_i)) &= \log(z_i) \\ &= \eta_i, \end{aligned} \qquad (17)$$

where the function $h(p_i) = \log(-\log(1 - p_i))$ is called the complementary log log function (cloglog). Therefore, eliciting the probability of presence given the complementary log log link function is a replacement for eliciting the expected abundance, when the expected abundance is very low (near zero). If the expected abundance is far from zero, then the probability of presence is effectively 1 and so will not provide much information on $y_i$. It is then better to stick to the log link and target the magnitude (e.g. abundance).

#### *5.1.1.2 Binomial*

Let $p(z_i|n_i, y_i) = \binom{n_i}{z_i} y_i^{z_i}(1 - y_i)^{n-z_i}$ with mean given by the inverse complementary log log link function $y_i = h^{-1}(\eta_i) = 1 - \exp(-\exp \eta_i)$ and linear predictor $\eta_i = h(y_i) = \log(-\log(1 - y_i))$. The expected proportion of presences (successes) is given by $y_i$.

##### 5.1.1.2.1 Support

The support of $p(z)$ is over the bounded non-negative integers, $z = 0, \dots, n$. For example, it is applicable to a setting with the number of observed presences given a total number of observations in a given area or transect.

##### 5.1.1.2.2 Elicitation target

The expected value of the response is the expected proportion of presences given the $n_i$ observations

##### 5.1.1.2.3 Probability of presence

The probability of observing a zero count is $P(z_i = 0) = (1 - y_i)^{n_i}$. Let $p_i$ equal the probability of presence:

$$\begin{aligned} 1 - p_i &= P(z_i = 0) = (1 - y_i)^{n_i} \\ \log(-\log(1 - p_i)) &= \log(-\log(1 - y_i)^{n_i}) \\ &= \log(-\log(1 - y_i)) + \log n_i \\ &= \eta_i + \log n_i, \end{aligned} \tag{18}$$

where $\log n_i$ is an offset. The left-hand side on the second line in Equation (18) is again the complementary log log link function, which is also the assumed link function for this binomial generalised linear model. Therefore, the probability of presence given the complementary log log link function is equivalent to eliciting the expected proportion of presences with an offset.

### 5.1.2 Structure of design matrix

The elicitation of $y(t_k)$ for the short- and long-term future periods depend on the hydrological response variables and a realised value of $y(t_{ref})$. The model formulation uses a quadratic surface to describe the relationship between hydrological response variables and receptor impact variables. The curvature allowed for the fact that most ecological variables will have optimal values at intermediate levels of an environmental gradient. For example, not enough water can lead to tree mortality due to drought, whereas too much water can lead to tree mortality due to flooded conditions.

A fundamental issue in developing the receptor impact models is that for a significant number of variables their current state across the landscape class is unknown, which obviously impacts on the ability to make predictions. A key example is groundwater depth. Change in depth can be modelled but there will not be detailed maps of groundwater depth across all subregions or bioregions. Another example is information on the presence, absence or condition of an ecological community. This data will typically be incomplete or missing entirely. Given these constraints, models will sometimes need to accept covariates defined in terms of deviations in hydrological response variables relative to ‘reference’ conditions.

The model structure is determined by the design matrix $X$ that is composed of the design points $x_i, i = 1, \dots, n$. For hydrological response variables that vary in both the reference and future periods, the functional form is a second order polynomial on the linear predictor that allows interactions among hydrological response variables, the reference period and the future period (Equation 19). For hydrological response variables that are defined relative to the reference period, the values of the hydrological response variables are fixed by definition in the reference period. For a given receptor impact model, enumerate the hydrological response variables with varying values in the reference period by $h_j, j = 1, \cdots, V$, and the hydrological response variables without varying values in the reference period by $h_j, j = V + 1, \cdots, N$:

$$\begin{aligned}\eta = \quad & \beta_0 x_0 + \beta_f x_f + \beta_l x_l + \beta_r x_r \\ & + \sum_{j=1}^{N} \beta_{h_j} x_{h_j} + \sum_{j=1}^{N-1} \sum_{k=j+1}^{N} \beta_{h_j \times h_k} x_{h_j} x_{h_k} + \sum_{j=1}^{N} \beta_{h_j^2} x_{h_j}^2 \\ & + \sum_{j=1}^{V} \beta_{h_j \times f} x_{h_j} x_f + \sum_{j=1}^{V-1} \sum_{k=j+1}^{V} \beta_{h_j \times h_k \times f} x_{h_j} x_{h_k} x_f + \sum_{j=1}^{V} \beta_{h_j^2 \times f} x_{h_j}^2 x_f\end{aligned} \qquad (19)$$

The coefficients are defined in Table 3.

**Table 3 The coefficient notation, attributed names and corresponding covariates as defined by the structure of the design matrix for the full model**

| Coefficient | Name | Covariate description |
|---|---|---|
| $\beta_0$ | Intercept | All-ones vector |
| $\beta_f$ | Future | Binary: scored 1 if case is in a short- or long-term assessment year |
| $\beta_l$ | Long | Binary: scored 1 if case is in the long-term future assessment year |
| $\beta_r$ | $y_{ref}$ | Continuous: value of RIV in the reference assessment year on the link transformed scale, $g\left(y(t_{ref})\right) = \eta_{ref}$; set to zero if case is in the reference assessment year |
| $\beta_{h_j}$ | Linear | Continuous: linear trend with HRV $h_j$ |
| $\beta_{h_j \times h_k}$ | Interaction | Continuous: interaction between HRVs $h_j$ and $h_k$ |
| $\beta_{h_j^2}$ | Quadratic | Continuous: square of HRV $h_j$ |
| $\beta_{h_j \times f}$ | Linear:future | Continuous: interaction of linear trend with HRV $h_j$ and future period |
| $\beta_{h_j \times h_k \times f}$ | Interaction:future | Continuous: interaction between HRVs $h_j$ and $h_k$ and future period |
| $\beta_{h_j^2 \times f}$ | Quadratic:future | Continuous: interaction between square of HRV $h_j$ and future period |

HRV = hydrological response variable, RIV = receptor impact variable

### 5.1.2.1 *Influence of the receptor impact variable from the reference assessment year*

Note that by construction the covariate $x_r$ named with the symbolic shorthand as $y_{ref}$ in Table 3, which has zero values in the reference period, can be interpreted as an interaction between $g\left(y(t_{ref})\right) = \eta_{ref}$ and the future period binary factor, $x_f$. To see the ecological interpretation of $\beta_r$ in the above equation for $\eta$ that is associated with the covariate $x_r$ (Table 3), which is equivalent to the entrywise product between $x_f$ and $\eta_{ref}$, first consider a model developed for the unknown quantity $\Delta_t = \eta_t - o_t$, with $o_t$ a known offset,

$$\Delta_t = \beta_0 x_0 + \beta_f x_f + \beta_l x_l + H_t + \beta_{fr} x_f \eta_{ref}, \tag{20}$$

where the term:

$$H_t = \sum_{j=1}^{N} \beta_{h_j} x_{h_j} + \sum_{j=1}^{N-1} \sum_{k=j+1}^{N} \beta_{h_j \times h_k} x_{h_j} x_{h_k} + \sum_{j=1}^{N} \beta_{h_j^2} x_{h_j}^2 + \sum_{j=1}^{V} \beta_{h_j \times f} x_{h_j} x_f + \sum_{j=1}^{V-1} \sum_{k=j+1}^{V} \beta_{h_j \times h_k \times f} x_{h_j} x_{h_k} x_f + \sum_{j=1}^{V} \beta_{h_j^2 \times f} x_{h_j}^2 x_f \tag{21}$$

captures the hydrological response variable effects on the receptor impact variable. The offset is specified as the function:

$$o_t = \begin{cases} 0, & \text{if } t = ref \\ \eta_{ref}, & \text{if } t = \{short, long\} \end{cases} \tag{22}$$

with the value of $\eta_{ref}$ in the last line assumed known. Given this specification of the offset, the unknown quantity $\Delta_t$ is defined by:

$$\Delta_t = \begin{cases} \eta_{ref}, & \text{if } t = ref \\ \eta_t - \eta_{ref}, & \text{if } t = \{short, long\} \end{cases} \tag{23}$$

again, with the value of $\eta_{ref}$ in the last line assumed known.

In the above model for $\Delta_t$, the term $\beta_{fr} x_f \eta_{ref}$ measures the association between the receptor impact variable $y(t_{ref})$ and the future change in $y(t)$ on the linear predictor scale, and so it includes an interaction with the binary covariate $x_f$ such that this term has no influence on the receptor impact variable in the reference period. The coefficient $\beta_{fr}$ therefore defines the relationship between $\eta_{ref}$ and the future change $\Delta_t$ when $t \in \{short, long\}$.

The above model for $\Delta_t$ can be rewritten with the known offset moved to the right-hand side:

$$\eta_t = \beta_0 x_0 + \beta_f x_f + \beta_l x_l + H_t + \beta_{fr} x_f \eta_{ref} + o_t. \tag{24}$$

Consider this model applied to a scenario in the reference period:

$$\eta_{ref} = \beta_0 x_0 + \beta_f x_f + \beta_l x_l + H_t + \beta_{fr} x_f \eta_{ref} + o_t = \beta_0 + H_t \tag{25}$$

where the known offset is set to zero when $t = ref$, so that $\Delta_{ref} = \eta_{ref} - o_t = \eta_{ref}$. The unknown $\eta_{ref}$ only appears on the left-hand side, as desired (it is the response variable). For a scenario in the future period with $t \in \{short, long\}$, the model instead takes the form:

$$\begin{aligned} \eta_t &= \beta_0 x_0 + \beta_f x_f + \beta_l x_l + H_t + \beta_{fr} x_f \eta_{ref} + o_{t\in\{short,long\}} \\ &= \beta_0 + \beta_f + \beta_l x_l + H_t + \beta_{fr} \eta_{ref} + \eta_{ref} \\ &= \beta_0 + \beta_f + \beta_l x_l + H_t + (\beta_{fr} + 1)\eta_{ref}, \end{aligned} \tag{26}$$

where $\Delta_t = \eta_t - o_t = \eta_t - \eta_{ref}$ with known value of $\eta_{ref}$ for $t \in \{short, long\}$. It can be seen that $(\beta_{fr} + 1) = \beta_r$ for a future scenario. The coefficient $\beta_r$ associated with the covariate $x_r$ in the original model for $\eta$ above provides the relationship between $\eta_{ref}$ and the future $\eta_t$ when $t \in \{short, long\}$.

All parameters thus have an ecological interpretation. Importantly, dependence of the receptor impact variable on the hydrological response variables is modelled jointly across both reference and future scenarios. Moreover, all unknown parameters appear linearly in the above equations. This linear property will be used by the method of estimation for the unknown parameters $\beta$ described in Chapter 7. For estimation, the choices for the values of $\eta_{ref}$ in the known offset function are derived from the elicitation scenarios as described in Chapter 7. Future predictions for $t \in \{short, long\}$ will depend on known $\eta_{ref}$ generated in the reference assessment year as described in Chapter 8. Example receptor impact models with specified design matrices are given in Section 5.3.

## 5.2 Design construction

The selection of design points (or scenarios) for consideration by the experts occurs prior to the receptor impact modelling workshop in two stages. First, potential candidate design points are identified. This stage uses hydrological model output and hydrology expertise to identify plausible bounds on the relevant hydrological response variables. Second, design points are selected from the set of candidate points in such a way to optimise the design subject to the structure of the design matrix as described above.

### 5.2.1 Candidate point selection

The number of samples $n$ is set equal to the total number of parameters in the full model that describes the quadratic surface, which fully interacts with the reference period and future period, along with the parameters that correspond to the intercept, the long-term assessment year and the influence of the receptor impact variable from the reference assessment year.

For unconstrained designs, the candidate design points are defined by a $3^V$ factorial design with corner points determined by the ranges of the $V$ hydrological response variables in the reference period. The centre values are typically set to the mid-point of the hydrological response variable ranges for each marginal (Figure 13, top) but occasionally modified, for example, to the logarithmic scale. In the future period, the candidate points are similarly drawn from a $3^N$ factorial design for the $N$ hydrological response variables in the future period augmented by low and high values for the terms: future period, long-term assessment year in the future period and $y_{ref}$; this results in a $2^3 \times 3^N$ factorial design used to generate the candidate set in the future period.

The case of constrained or restricted design regions is considered in Section 5.2.3.

### 5.2.2 Design point selection

The $D$-criterion (Chaloner and Verdinelli, 1995) seeks to maximise the objective function:

$$\zeta^* = \arg\max_{\zeta} \log\det M(\zeta) \qquad (27)$$

where $M(\zeta) = \sum \zeta_i x_i^{\top} x_i$, with the notation that $x_i^{\top}$ is a row vector of $X$, and where $\zeta$ is a probability measure on the design region $\mathcal{X}$ with $\zeta_i = n_i/n$, where $n_i$ is the number of samples at the $i$th design point and $\sum_i n_i /n = 1$ with $n$ the total number of samples. The numerical solution uses the optimisation algorithm of Federov (1972) as implemented by Wheeler (2004) and Wheeler (2014) separately applied to the candidate design points generated for the reference and future assessment years. The resulting optimised solutions are randomly ordered within each assessment year. In all cases, the elicitation procedure begins with elicitations in the reference year before progressing to the short-term assessment year and finishing with elicitation in the long-term assessment year (Figure 12).

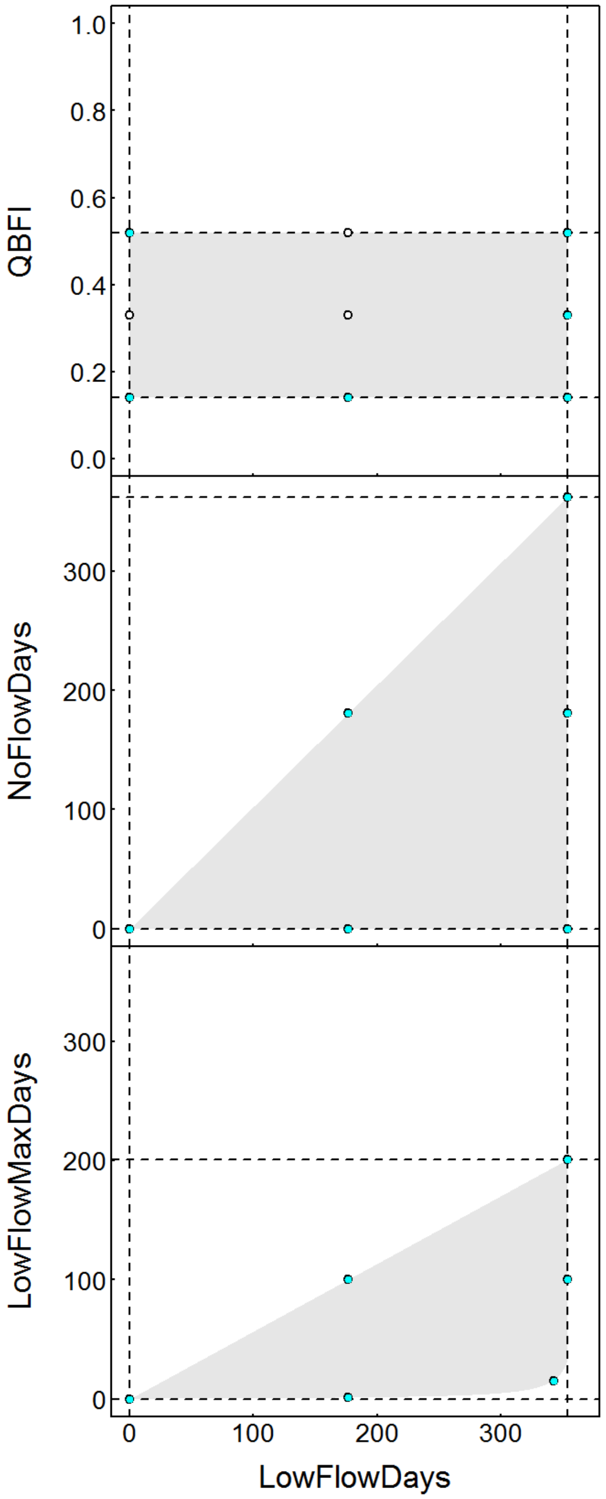

**Figure 13 Constraints and design point selection for simplified surface water configurations considered only within the reference period**

The average number of low-flow days over a 30-year period is given on the x-axis. *Top*: QBFI, an index of the ratio of surface flow to baseflow, ranges between 0 and 1. The ranges estimated from the stochastic hydrology modelling output are given by black dashed lines. The resulting feasible design region is the shaded area. Candidate design points generated from the 3 by 3 factorial design given by the design region ranges and mid-points are shown by open circles. The cyan points are those selected by the D-criterion optimisation algorithm. *Middle*: NoFlowDays is similarly defined as LowFlowDays but uses a more extreme flow threshold. NoFlowDays cannot be greater than LowFlowDays, which produces the constrained design region (shaded area). *Bottom*: LowFlowMaxDays is the maximum duration of low-flow events. The defined relationships between these hydrological response variables result in a complicated set of constraints. The candidate design points affected by the lower bound were adjusted and the revised design points are shown for this example.

### 5.2.3 Restricted design regions

There are two general strategies to deal with constrained design regions. The first approach would be to propose a set of candidate design points for the unconstrained design region, then optimise the selection of candidate design points from the subset that meet the constraints. The second approach would be to modify the candidate points so as to meet the constraints before optimising the set of design points.

For some elicitations, the first approach is sufficient. For example, in the Gloucester subregion ‘Perennial – gravel/cobble streams’ landscape class percent canopy cover receptor impact model, EventsR0.3 and EventsR3.0, respectively, serve as proxies for overbench and overbank flood events. The number of overbank flood events (EventsR3.0) cannot be greater than the number of overbench flood events (EventsR0.3). Another example is the relationship between the number of no-flow days (equivalently, zero-flow days) and the number of low-flow days. A no-flow day is a day when the flow does not exceed a negligible value, whereas a low-flow day occurs if the daily flow does not exceed a higher value that corresponds to a defined level of low flow. The number of no-flow days therefore cannot exceed the number of low-flow days. This latter example is shown in Figure 13 (middle) for the reference period.

For more restrictive constraint relationships, the simple lattice structure approach to generating candidate points may lose too many candidate points once the constraints are applied. However, in such cases for the receptor impact models considered here, a small adjustment may be applied that brings some of these excluded lattice points back into the feasible design region. The D-optimality algorithm can then be applied to these adjusted candidate points. An example is given below.

This second approach is necessary for combinations of hydrological response variables composed of average number of days of low flow and the average maximum duration of the event, defined as the number of contiguous days separated by a full day over the low-flow threshold. The definitions for these two hydrological response variables impose a complicated set of constraints. The maximum duration of low-flow events within a year, given by $m$, must be less than the number of total low-flow days ($m < l$), but not below the curve given by $f(l) = \frac{l}{365-l}$, thus, $f(l) \leq m \leq l$. Let the coordinates of a non-compliant candidate point drawn from the $3^N$ factorial lattice be given by $(l_f, m_f)$, where $0 \leq m_f < f(l_f)$. The distance from this point to an arbitrary point $(l, m)$ is given by:

$$\begin{aligned} \delta(l,m) &= \sqrt{(l-l_f)^2 + (m-m_f)^2} \\ \delta(l) &= \sqrt{(l-l_f)^2 + l^2/(365-l)^2 - 2lm_f/(365-l) + m_f^2} \end{aligned} \tag{28}$$

where the second line gives the distance to an arbitrary point on the curve that is obtained by substituting $m = f(l_f)$ into the first line. Differentiating this result gives:

$$\frac{d\delta}{dl} = \frac{l - l_f + l/(365 - x)^2 + l^2/(365 - l)^3 - m_f/(365 - l) - lm_f/(365 - l)^2)}{\sqrt{(l - l_f)^2 + l^2/(365 - l)^2 - 2lm_f/(365 - l) + m_f^2}} \tag{29}$$

Setting $d\delta/dl = 0$ gives the quartic polynomial equation:

$$\begin{aligned} \lambda^4 \quad & +(-1095 - l_f)\lambda^3 + (399675 + 1095l_f)\lambda^2 \\ & +(-365m_f - 399675l_f - 48627490)\lambda + 133225m_f + 48627125l_f = 0 \end{aligned} \tag{30}$$

The derivative $f'(l)$ is positive for $0 \leq l \leq 365$. The curve $f(l)$ is therefore monotonically increasing within the feasible region and a single real root $\lambda_1$ satisfies the constraints, $f(\lambda_1) \leq \lambda_1 \leq l_f$. Thus, the solution for the sought-after feasible candidate point is given by the coordinates, $(\lambda_1, f(\lambda_1))$. A graphical example is given in Figure 13 (bottom). For example, the candidate point in the lower right corner has been visibly adjusted inward to the closest point within the feasible design region.

## 5.3 Example designs

In this section, two examples are provided from the Gloucester subregion (see companion product 2.7 for the Gloucester subregion (Hosack et al., 2018)). The elicitation scenarios as presented to the experts are provided, and the corresponding design matrices for the full model are presented. The qualitative modelling workshop identified several hydrological response variables to include in the receptor impact modelling for the Gloucester subregion. These are summarised in Table 4. The first example has varying hydrological response variables only in the future period that are defined relative to the reference period. This receptor impact model conditioned on the hydrological response variables: dmaxRef and tmaxRef (groundwater) and EventsR0.3 and EventsR3.0 (surface water). The second example has a single hydrological response variable that varies in both the reference and future periods: ZQD (surface water).

**Table 4 Summary of the hydrological response variables used in the receptor impact models for the 'Perennial – gravel/cobble streams' landscape class, together with the signed digraph variables that they correspond to for the Gloucester subregion**

| Hydrological response variable | Definition of hydrological response variable | Signed digraph variable |
|---|---|---|
| dmaxRef | Maximum difference in drawdown under the baseline future or under the coal resource development pathway future relative to the reference period (1983 to 2012) | GW |
| tmaxRef | The year that the maximum difference in drawdown relative to the reference period (1983 to 2012) (dmaxRef) occurs. | GW |
| EventsR0.3 | The mean annual number of events with a peak daily flow exceeding the threshold (the peak daily flow in flood events with a return period of 0.3 years as defined from modelled baseline flow in the reference period (1983 to 2012)). This metric is designed to be approximately representative of the number of overbench flow events in future 30-year periods. | FR1 |
| EventsR3.0 | The mean annual number of events with a peak daily flow exceeding the threshold (the peak daily flow in flood events with a return period of 3.0 years as defined from modelled baseline flow in the reference period (1983 to 2012)). This metric is designed to be approximately representative of the number of overbank flow events in future 30-year periods. | FR2 |
| ZQD | The number of zero-flow days per year, averaged over 30 years | FR3 |
| QBFI | Ratio of total baseflow generation to total streamflow generation, averaged over a 30-year period | FR3 |

FR1 = flow regime 1, FR2 = flow regime 2, FR3 = flow regime 3, GW = groundwater

### 5.3.1 Example with non-varying hydrological response variables in the reference period

One example receptor impact variable in the Gloucester subregion is the mean percent canopy cover of riparian vegetation in the 'Perennial – gravel/cobble streams' landscape class. The mean percent canopy cover of riparian vegetation was considered as a temporal average over the assessment year. The area of the hypothetical transect was 2000 m$^2$ covering from the bottom of the stream bench to the high bank. This transect would typically have dimensions 20 m wide by 100 m length but is envisioned to vary with the local stream topography.

This receptor impact model is conditioned on the hydrological response variables (see Table 4): dmaxRef and tmaxRef (groundwater) and EventsR0.3 and EventsR3.0 (surface water). The elicitation scenarios presented to the experts are shown in Table 5. The values of the hydrological response variables are fixed in the reference period by definition. The value of $y(t_{ref})$, which forms the covariate $y_{ref}$ in Table 3, is not conditioned on elicitations within the reference assessment year of 2012. In the future period, a low and high value for $y(t_{ref})$ were determined from the 1/4th and 3/4th elicited fractiles (see Section 6.2.2), respectively, from the reference assessment year.

**Table 5 Elicitation scenarios considered by the experts for percent canopy cover of riparian vegetation along perennial gravel/cobble streams in the Gloucester subregion**

| $y(t_{ref})$ | dmaxRef | tmaxRef | EventsR0.3 | EventsR3.0 | Year |
|---:|---:|---|---:|---:|---|
| na | na | na | 3.33 | 0.33 | 2012 |
| 0.3 | 6 | 2102 | 4 | 0.34 | 2042 |
| 0.6 | 6 | 2019 | 4 | 0.23 | 2042 |
| 0.6 | 1.7 | 2102 | 4.6 | 0.29 | 2042 |
| 0.6 | 0 | 2019 | 4.6 | 0.23 | 2042 |
| 0.6 | 6 | 2060 | 4.6 | 0.34 | 2042 |
| 0.3 | 0 | 2060 | 3.3 | 0.34 | 2042 |
| 0.3 | 6 | 2019 | 3.3 | 0.34 | 2042 |
| 0.3 | 6 | 2102 | 3.3 | 0.23 | 2042 |
| 0.3 | 6 | 2019 | 4.6 | 0.29 | 2102 |
| 0.6 | 6 | 2060 | 3.3 | 0.29 | 2102 |
| 0.6 | 1.7 | 2102 | 3.3 | 0.34 | 2102 |
| 0.6 | 0 | 2102 | 3.3 | 0.23 | 2102 |
| 0.3 | 0 | 2019 | 3.3 | 0.23 | 2102 |
| 0.6 | 0 | 2019 | 4 | 0.34 | 2102 |
| 0.3 | 0 | 2102 | 4.6 | 0.34 | 2102 |
| 0.6 | 6 | 2102 | 4.6 | 0.23 | 2102 |
| 0.3 | 1.7 | 2060 | 4 | 0.23 | 2102 |

Hydrological response variables are defined in Table 4. Yref is value of receptor impact variable in the reference assessment year on the linear predictor scale.
na = not applicable

The design matrix of the full model with quadratic surface for the hydrological response variable is split into three tables: Table 6, Table 7 and Table 8. The subset of covariates that are retained for all possible models is given in Table 6; compare with the definitions found in Table 3. The new covariate Yrs2tmaxRef, which is assigned zero in the reference period, is simply the difference between tmaxRef and the assessment year, $Yrs2tmaxRef = tmaxRef - t_k, k \in \{2012, 2042, 2102\}$. The interaction terms for the hydrological response variables are given in Table 7. Note that there are no interaction terms between the hydrological response variables and the future period binary factor because the hydrological response variables do not vary in the reference period. The quadratic terms for the hydrological response variables are given in Table 8. As described above, the covariates that appear in Table 6 are common to all models. The covariates that appear in Table 7 and Table 8 may potentially be dropped (see Section 7.2), however, forming a candidate set of simpler alternative models relative to the full model.

**Table 6 Partial design matrix of the full model for mean percent canopy cover of riparian vegetation along perennial gravel/cobble streams in the Gloucester subregion: essential covariates**

| Intercept | future | long | Yref | dmaxRef | Yrs2tmaxRef | EventsR0.3 | EventsR3.0 |
|---:|---:|---:|---:|---:|---:|---:|---:|
| 1 | 0 | 0 | 0 | 0 | 0 | 3.33 | 0.33 |
| 1 | 1 | 0 | –1.03 | 6 | 60 | 4 | 0.34 |
| 1 | 1 | 0 | –0.09 | 6 | –23 | 4 | 0.23 |
| 1 | 1 | 0 | –0.09 | 1.7 | 60 | 4.6 | 0.29 |
| 1 | 1 | 0 | –0.09 | 0 | –23 | 4.6 | 0.23 |
| 1 | 1 | 0 | –0.09 | 6 | 18 | 4.6 | 0.34 |
| 1 | 1 | 0 | –1.03 | 0 | 18 | 3.3 | 0.34 |
| 1 | 1 | 0 | –1.03 | 6 | –23 | 3.3 | 0.34 |
| 1 | 1 | 0 | –1.03 | 6 | 60 | 3.3 | 0.23 |
| 1 | 1 | 1 | –1.03 | 6 | –83 | 4.6 | 0.29 |
| 1 | 1 | 1 | –0.09 | 6 | –42 | 3.3 | 0.29 |
| 1 | 1 | 1 | –0.09 | 1.7 | 0 | 3.3 | 0.34 |
| 1 | 1 | 1 | –0.09 | 0 | 0 | 3.3 | 0.23 |
| 1 | 1 | 1 | –1.03 | 0 | –83 | 3.3 | 0.23 |
| 1 | 1 | 1 | –0.09 | 0 | –83 | 4 | 0.34 |
| 1 | 1 | 1 | –1.03 | 0 | 0 | 4.6 | 0.34 |
| 1 | 1 | 1 | –0.09 | 6 | 0 | 4.6 | 0.23 |
| 1 | 1 | 1 | –1.03 | 1.7 | –42 | 4 | 0.23 |

Hydrological response variables are defined in Table 4. Yref is value of receptor impact variable in the reference assessment year on the linear predictor scale. Yrs2tmaxRef, which is assigned zero in the reference period, is simply the difference between tmaxRef and the assessment year, $Yrs2tmaxRef = tmaxRef - t_k, k \in \{2012, 2042, 2102\}$.

**Table 7 Partial design matrix of the full model for mean percent canopy cover of riparian vegetation along perennial gravel/cobble streams in the Gloucester subregion: interaction terms**

| dmaxRef: Yrs2tmaxRef | dmaxRef: EventsR0.3 | dmaxRef: EventsR3.0 | Yrs2tmaxRef: EventsR0.3 | Yrs2tmaxRef: EventsR3.0 | EventsR0.3: EventsR3.0 |
|---:|---:|---:|---:|---:|---:|
| 0 | 0 | 0 | 0 | 0 | 1.11 |
| 360 | 24 | 2.04 | 240 | 20.4 | 1.36 |
| –138 | 24 | 1.38 | –92 | –5.29 | 0.92 |
| 102 | 7.82 | 0.49 | 276 | 17.4 | 1.33 |
| 0 | 0 | 0 | –105.8 | –5.29 | 1.06 |
| 108 | 27.6 | 2.04 | 82.8 | 6.12 | 1.56 |
| 0 | 0 | 0 | 59.4 | 6.12 | 1.12 |
| –138 | 19.8 | 2.04 | –75.9 | –7.82 | 1.12 |
| 360 | 19.8 | 1.38 | 198 | 13.8 | 0.76 |
| –498 | 27.6 | 1.74 | –381.8 | –24.07 | 1.33 |
| –252 | 19.8 | 1.74 | –138.6 | –12.18 | 0.96 |
| 0 | 5.61 | 0.58 | 0 | 0 | 1.12 |
| 0 | 0 | 0 | 0 | 0 | 0.76 |
| 0 | 0 | 0 | –273.9 | –19.09 | 0.76 |
| 0 | 0 | 0 | –332 | –28.22 | 1.36 |
| 0 | 0 | 0 | 0 | 0 | 1.56 |
| 0 | 27.6 | 1.38 | 0 | 0 | 1.06 |
| –71.4 | 6.8 | 0.39 | –168 | –9.66 | 0.92 |

Hydrological response variables are defined in Table 4. Yrs2tmaxRef, which is assigned zero in the reference period, is simply the difference between tmaxRef and the assessment year.

**Table 8 Partial design matrix of the full model for mean percent canopy cover of riparian vegetation along perennial gravel/cobble streams in the Gloucester subregion: quadratic terms**

| $\text{dmaxRef}^2$ | $\text{Yrs2tmaxRef}^2$ | $\text{EventsR0.3}^2$ | $\text{EventsR3.0}^2$ |
|---:|---:|---:|---:|
| 0 | 0 | 11.11 | 0.11 |
| 36 | 3600 | 16 | 0.12 |
| 36 | 529 | 16 | 0.05 |
| 2.89 | 3600 | 21.16 | 0.08 |
| 0 | 529 | 21.16 | 0.05 |
| 36 | 324 | 21.16 | 0.12 |
| 0 | 324 | 10.89 | 0.12 |
| 36 | 529 | 10.89 | 0.12 |
| 36 | 3600 | 10.89 | 0.05 |
| 36 | 6889 | 21.16 | 0.08 |
| 36 | 1764 | 10.89 | 0.08 |
| 2.89 | 0 | 10.89 | 0.12 |
| 0 | 0 | 10.89 | 0.05 |
| 0 | 6889 | 10.89 | 0.05 |
| 0 | 6889 | 16 | 0.12 |
| 0 | 0 | 21.16 | 0.12 |
| 36 | 0 | 21.16 | 0.05 |
| 2.89 | 1764 | 16 | 0.05 |

Hydrological response variables are defined in Table 4. Yrs2tmaxRef, which is assigned zero in the reference period, is simply the difference between tmaxRef and the assessment year.

## 5.3.2 Example with varying hydrological response variables in the reference period

Another example receptor impact variable from the Gloucester subregion is hyporheic (the area beneath the streambed where surface water mixes with groundwater) taxa richness in the ‘Intermittent – gravel/cobble streams’ landscape class. The units are taxa richness per 6 L of water. This receptor impact model is conditioned on the surface water hydrological response variable zero-flow days (averaged over 30 years) (ZQD), which varies in both the reference and future periods. The elicitation scenarios presented to the experts for this example are shown in Table 9.

**Table 9 Elicitation scenarios considered by the experts for hyporheic taxa richness within the 'Intermittent – gravel/cobble streams' landscape class in the Gloucester subregion**

| Yref | ZQD | Year |
|---:|---:|:---|
| na | 330.3 | 2012 |
| na | 165 | 2012 |
| na | 0 | 2012 |
| 6 | 0 | 2042 |
| 12 | 165 | 2042 |
| 6 | 330 | 2042 |
| 12 | 330 | 2102 |
| 6 | 165 | 2102 |

Hydrological response variables are defined in Table 4.
Yref is value of receptor impact variable in the reference assessment year on the linear predictor scale.
na = not applicable, ZQD = zero-flow days (averaged over 30 years)

The design matrix of the full model with quadratic surface for the hydrological response variable is given in Table 10. Note that, in this case, there are no interaction terms between hydrological response variables given the single hydrological response variable. On the other hand, there are interaction terms between the hydrological response variables and the future period binary factor because the values for ZQD vary in both the reference and future periods. The covariates that include interactions or quadratic terms may be dropped, which forms a candidate set of simpler alternative models relative to the full model.

**Table 10 Design matrix for the full receptor impact model of hyporheic taxa richness**

| Intercept | future | long | Yref | ZQD | ZQD$^2$ | future: ZQD | future: ZQD$^2$ |
|---:|---:|---:|---:|---:|---:|---:|---:|
| 1 | 0 | 0 | 0 | 330.3 | 109098.1 | 0 | 0 |
| 1 | 0 | 0 | 0 | 165 | 27225 | 0 | 0 |
| 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 1 | 1 | 0 | 1.79 | 0 | 0 | 0 | 0 |
| 1 | 1 | 0 | 2.48 | 165 | 27225 | 165 | 27225 |
| 1 | 1 | 0 | 1.79 | 330 | 108900 | 330 | 108900 |
| 1 | 1 | 1 | 2.48 | 330 | 108900 | 330 | 108900 |
| 1 | 1 | 1 | 1.79 | 165 | 27225 | 165 | 27225 |

Hydrological response variables are defined in Table 4. Yref is value of receptor impact variable in the reference assessment year on the linear predictor scale.
ZQD = zero-flow days (averaged over 30 years)

# 6 Receptor impact modelling workshop (stage 5)

Given the hydrological scenarios defined in the previous stage, the predicted response for the receptor impact variables is elicited from the external experts at the receptor impact modelling workshop. The expert education and preparation process leading up to and including the receptor impact modelling workshop is detailed in Section 6.1. The elicitation procedure used at the workshop to elicit expert judgement conditional on the hydrological scenarios, which were developed in the previous stage, is detailed in Section 6.2. Incorporating additional or alternative experts would likely lead to alternative receptor impact models. The quality of a constructed receptor impact model can only be assessed by evaluating the expert assessment against empirical data. The statistical nature of the receptor impact models (Section 5.1) and scenario construction (Section 5.2) explicitly enables this empirical evaluation of expert contributions by allowing incorporation of independent empirical data through the likelihood function within a generalised linear model. This section discusses the expert preparation, elicitation structure and probability assessment that ensure the potential coherent assimilation of the contributed expert assessments with empirical data, while accommodating for expert uncertainty in the relationship between hydrological response variables and the receptor impact variable.

## 6.1 Expert preparation

The protocol for eliciting receptor impact models is structured around each landscape class defined within each bioregion or subregion. The fundamental question ‘How might selected receptor impact variables change under various scenarios of change for the hydrological response variables?’ is addressed by the experts. A very important step in the process therefore is the formalisation and specification of the relevant impact and response variables and the relationships among these variables (Chapter 4).

The protocol methodology asks experts to provide estimates of the chosen receptor impact variable at a range of values of the relevant hydrological response variables. It then uses statistical modelling to build a relationship between the hydrological response variables and the receptor impact variables, where the former are treated as covariates and the latter as the response of a linear model with potentially non-linear basis functions (sensu Hosack et al., 2016). This approach provides a repeatable protocol to develop these models via experts.

The elicitation session is designed to tackle a challenging problem, and it depends on the collaboration and cooperation of the experts. For this session, experts are asked to contribute their knowledge and expertise in a small group setting, although sometimes individually when only a single expert is available. The group format permits experts to confer and seek a consensus opinion when responding to presented scenarios. A group approach not only allows for the entire group to contribute but also permits the opportunity for feedback and group learning while responding to hypothetical scenarios.

The elicitations take place at receptor impact modelling workshops, which are separately held for each bioregion and subregion. The attending experts represent a wide range of expertise and experience. At the workshop, parallel sessions are run with experts grouped by expertise. The groups may range in size from a single individual up to a small number of individuals. The expert preparation for these group elicitation sessions is as follows.

The following steps were involved prior to the receptor impact modelling workshop:

1. Experts are selected for the elicitation process. The choice of experts and their invitation into the process ensures that appropriate expertise is identified and included in the process. Relevant experts are contacted by staff collaborating among the Office of Water, the Bureau of Meteorology and the BA ecology discipline teams for each bioregion or subregion. Ideally, experts will have previously attended the relevant qualitative modelling workshop for each bioregion or subregion. This allows experts to take advantage of their previous exposure to and discussion about the bioregion or subregion, the relevant landscape classes and the receptors and hydrological response variables identified in the qualitative models. This continuity helps experts smoothly transition into a quantitative assessment for one or more receptor impact variables, while understanding the context given to a particular receptor impact variable by its relationship with respect to the landscape class and ecosystem. As noted above, the qualitative models help experts identify both potential direct and indirect hydrological impacts. Also, experts that have attended the qualitative modelling workshop will have had previous exposure to the scope and objectives of the Bioregional Assessment Programme, which leads to further gains in efficiency.
2. Prior to the first meeting the experts receive material outlining the Programme, the nature and purpose of the receptor impact models and a description of the landscape classes, hydrological response variables and potential receptor impact variables. It also describes how the experts' responses are to be used to develop the receptor impact model.
3. The quality of the group elicitation depends on a collegial, open-minded and focused atmosphere. The experts are provided information on the group elicitation approach. In particular, the following guidelines are provided:
    a. Respect alternative opinions. Everyone has different points of view and experience that will be called upon over the course of the session.
    b. Practice patience. The problem is new and challenging. It will take some thought and work to assess and talk about, which can take time. Occasionally, points will need to be considered again (and even again).
    c. Ask questions of the session facilitators. Assessing expert opinion is a communication exercise, and please ask if you are unsure about what is being asked of you. We much prefer to address any confusion as it arises.
    d. Ask questions of your fellow experts. Do not assume that 'you know what they mean', even if you have worked together closely in the past. Everyone has different ways for expressing their views, and clarifying questions can help resolve where points of view are both different and similar.

e. Recognise the minority opinion. Please speak up! The group elicitation approach depends on everyone participating. If you are uncertain, that is ok, this is accommodated by the elicitation approach that will be used in this session.

At the receptor impact modelling workshop, the following steps are completed before beginning the elicitation sessions:

1. The experts are convened and given a brief introduction about BAs, landscape class definitions and hydrological response variable definitions. They are then encouraged to ask questions about the process.
2. Before undergoing an elicitation session, experts are first educated about subjective probability, trained on heuristics and biases commonly encountered in expert elicitation exercises (O'Hagan et al., 2006).
3. The elicitation procedure is presented (Section 6.2) to the experts. Practice examples are worked through until experts are comfortable with the method before beginning the elicitation.
4. The experts are reminded of the group elicitation approach and its accompanying challenges and guidelines for expected group behaviour, as documented above. The experts are then split into their parallel sessions by expertise.
5. Within each parallel session, the facilitator uses their experience of the process to highlight challenges experts may have and strategies to overcome them. The choice of receptor impact variables is discussed to identify any fundamental conceptual modelling modifications from the preceding qualitative modelling workshop.

Throughout each parallel elicitation session, a BA contact is made available to respond to all queries from the experts about questions of hydrology, ecology or the BA process.

## 6.2 Conditional elicitation step

### 6.2.1 Structure of elicitation

The target of the elicitation is the unknown value of the receptor impact variable, $y$. The receptor impact variable $y$ is defined such that it has a direct interpretation relative to potential observables and expert knowledge (Chapter 4). For example, in a Bernoulli response model considered at the $i$th hydrological scenario, a (hypothetical) observation $z_i$, corresponds to either a presence or absence, and $y_i = \mathbb{E}[z_i]$ is interpreted as the probability of presence. In a Poisson response model where $z_i$ corresponds to a count, $y_i$ is interpreted as an intensity that may relate to the annual average abundance over a defined spatial scope. The experts provide subjective probability distributions describing the receptor impact variable estimates conditional on a hydrological scenario summarised within the design point $x_i^\top$ (see Section 5.2.2). The elicited subjective probability distributions are assumed independent conditional on the hydrological scenarios.

### 6.2.2 Elicitation of subjective conditional probabilities

Conditional on the design point $x_i$, the goal is to elicit a normal distribution for $\eta_i|x_i$ with the mean and variance parameters summarised into the parameter vector $\phi_i = [m_i, v_i]^\top$. An elicitation of fractiles (equivalently, percentiles or quantiles) for the target $y_i = g^{-1}(\eta_i)$, given the monotonic link function, directly translates into fractiles for the conditional normal distribution of the linear predictor, $g(y_i|\eta_i) = \eta_i|x_i \sim N(m_i, v_i)$. The experts are asked for each design point to perform judgements of equal odds in three steps:

1. What value do you believe gives a 50% chance that the true receptor impact variable is lower? (This obtains a prediction of the median, $f_{1/2}$)
2. Assume that the true value is really below $f_{1/2}$. Given this information, what value do you believe gives a 50% chance of being above or below the true value of the receptor impact variable? (This is the first quartile, $f_{1/4}$)
3. Assume that the true value is really above $f_{1/2}$. Given this information, what value do you believe gives a 50% chance of being above or below the true value of the receptor impact variable? (This is the third quartile, $f_{3/4}$).

The parameters $\phi_i$ are then chosen to minimise the information lost by approximating the elicited fractiles by a normal distribution. The elicited fractiles are assembled into the vector, $f = [f_{0/4}, \dots, f_{j/4}, \dots, f_{4/4}]^\top$ where $j = 0, \dots, 4$. The extreme fractiles $f_{0/4}$ and $f_{4/4}$ (possibly infinite) are determined by the support of $y_i$. Let the probability intervals determined by these elicited fractiles be denoted by $l_k = f_{k/4} - f_{(k-1)/4} = 0.25$ for $k = 1, \dots, 4$. Let $P_{e|i}$ denote the histogram constructed by these elicited fractiles.

Fitted probability intervals for the approximating normal distribution are obtained by:

$$z_k = \int_{f_{(k-1)/4}}^{f_{k/4}} d\,P_s(\eta_i|\phi_i) = P_s(f_{k/4}|\phi_i) - P_s(f_{(k-1)/4}|\phi_i), k = 1, \dots, 4 \tag{31}$$

where $P_s(\eta_i|\phi_i)$ is the subjective probability distribution function parametrised by $\phi_i$.

The Kullback-Leibler divergence of $P_{s|i}$ from the elicited $P_{e|i}$, where subscripts denote dependence on $\phi_i$ and $x_i$, is approximated by:

$$\mathrm{KL}(P_{e|i}||P_{s|i}) = \int \log \frac{dP_{e|i}}{dP_{s|i}} dP_{e|i} \approx \sum_k \log\left(\frac{l_k}{z_k}\right) l_k, \tag{32}$$

where $P_{e|i}$ is absolutely continuous with $P_{s|i}$. The parameters $\phi_i$ are chosen so as to minimise the approximate Kullback-Leibler divergence of $P_{s|i}$ from $P_{e|i}$:

$$\hat{\phi}_i = \arg\min_{\phi_i} \sum_k \log\left(\frac{l_k}{z_k}\right) l_k \tag{33}$$

For each scenario, the values of the design point are portrayed and the group discusses the potential ecological response. The goal is to develop a probability distribution that is an acceptable representation of the experts' beliefs. Of course, with unlimited time resources a suite of distribution families could be presented for consideration by experts. Resources in BA are not unlimited, however, and for consistency and efficiency the parametric distribution considered by the experts is defined by the models specified and described in Section 5.1. The following steps are then performed by the experts:

1. Initial fractile assessments are elicited using the quartile method. These are plotted graphically as vertical dashed lines. Note that more fractiles than free parameters are elicited in an approach that is referred to as 'overfitting' (O'Hagan et al., 2006), which will permit feedback between the model representation and the group's final probabilistic statement as detailed in the following steps.
2. The corresponding fitted fractiles and density curve from the fitted probability density function $p_{s|i}(y_i|\hat{\phi}_i)$ are plotted as three vertical blue dashed lines and blue curve, respectively. The overfitting approach uses the parametric model to average across multiple probability statements. The parametric model is unlikely to exactly match the elicited fractiles from Step 1, and so this process encourages the group to evaluate the parametric model with respect to their beliefs. If the values for the matching fractiles are not acceptable, then the group returns to Step 1 and adjusts the elicited fractiles. These two steps are repeated as often as necessary until the group accepts the parametric model quantiles as acceptable.
3. The extreme deciles from $p_{s|i}(y_i|\hat{\phi}_i)$ are plotted as dashed blue lines. The group considers these new predictions and returns to Step 1 if the predictions are unacceptable.
4. The group is allowed to consider other fractiles or cumulative probabilities as predictions from the parametric model. The group returns to Step 1 if these predictions are unacceptable.
5. After completing the above feedback steps, the group is allowed to accept the elicited subjective probability distribution $p_{s|i}(y_i|\hat{\phi}_i)$ as a reasonable assessment of the expert opinion.

This process thus elicits from the group a subjective probability distribution of ecological response given the covariate values that make up the defined scenario. Note that the elicitation focuses on the distribution, not the raw fractile assessments (e.g. Step 1 above). The raw fractiles are used by the experts as 'parameters' to iteratively build a probability distribution that is an acceptable representation of their beliefs. The probability predictions made by the elicited probability distribution provide the final products that are assessed by experts for either acceptance or rejection and further iteration until the experts accept the probability distribution as a reasonable model of their beliefs. This process is repeated for each design point.

# 7 Receptor impact model estimation (stage 6)

This stage constructs the receptor impact models given the specified model structure and the elicited responses contributed by experts, conditional on the hydrological scenarios presented at the receptor impact modelling workshop. The receptor impact models constructed for each landscape class within a bioregion or subregion are reported along with the relevant contextualisation in terms of landscape class definitions and qualitative modelling results in the corresponding product 2.7 (receptor impact modelling) (Figure 3). The method used for the receptor impact model construction is detailed below.

## 7.1 Prior distribution of beta

Conditional on a design point $x_i$, the elicited subjective probability distribution is Gaussian,

$$\eta_i | x_i \sim N(m_i, v_i), \quad i = 1, \dots, n. \tag{34}$$

The $m_i$ and $v_i$ are assumed conditionally independent from other elicitations given the design matrix $X$. Let $m = [m_1, m_2, \dots, m_n]^\top$ and $V = \mathrm{diag}(v_1, v_2, \dots, v_n)$ denote a diagonal covariance matrix. The conditional distribution of $\eta$ is proportional to:

$$\begin{aligned}\prod_{i=1}^{n} p\left(\eta_i | m_i, v_i\right) &= \prod_{i=1}^{n} p\left(x_i^\top \beta | m_i, v_i\right) \\ &\propto \exp\left\{-\frac{1}{2}(X\beta - m)^\top V^{-1}(X\beta - m)\right\}.\end{aligned} \tag{35}$$

The distribution for the unknown $\beta$ conditional on $m$, $V$ and $X$ is then proportional to a multivariate normal with mean and covariance given by:

$$\begin{aligned}\mu &= (X^\top V^{-1} X)^{-1} X^\top V^{-1} m, \\ \Sigma &= (X^\top V^{-1} X)^{-1}.\end{aligned} \tag{36}$$

## 7.2 Model structure uncertainty

The full design matrix allows for a quadratic surface response function between the hydrological response function and the receptor impact variables. This includes linear terms, pairwise interactions between the linear terms and quadratic terms for all hydrological response variables. If the hydrological response variables vary in the reference period, then an interaction between the quadratic surface response function and the future period is also included. The full design model structures the elicitation scenario so that this complex model can be elucidated.

However, the full richness of the model may be excessive for simple relationships between hydrological response variables and receptor impact variables. Simpler models are therefore considered. Optional model terms include: interactions between hydrological response variables and the future period, the pairwise interactions between different hydrological response variables and the quadratic terms. The alternative models are ranked using a Bayesian information criterion (BIC) (proportional to the Schwarz criterion; Schwarz, 1978) metric:

$$BIC_j \propto \left(X\hat{\beta} - m\right)^{\mathsf{T}} V^{-1}\left(X\hat{\beta} - m\right) + p_j \log n, \tag{37}$$

where $p_j$ is the dimension of the vector $\beta|M_j$ and $\hat{\beta}\,|M_j = \mu|M_j$. The model with the lowest BIC value is selected as the best model.

Note that several terms are required to be retained within all the candidate models. These include: the intercept, the future-period factor, the short-term factor, the influence of $y\left(t_{ref}\right)$ and at least the linear terms of all hydrological response variables to be considered in the receptor impact modelling elicitation workshop. The inclusion of this minimal subset ensures that the covariates which provide the structure for the elicitation scenarios are also represented in the estimation and prediction steps of the receptor impact modelling.

# 8 Receptor impact model prediction (stage 7)

This stage applies the receptor impact model methodology to predict the response of the receptor impact variables to simultaneous changes in one or more of the hydrological response variables. The general framework allows for the receptor impact model to be applied either at a single or at multiple receptor locations, for example multiple locations that are considered to be representative of a landscape class within a bioregion or subregion. The receptor impact modelling elicitations, however, are designed for all locations within a landscape class given the same hydrology. The uncertainty in the elicited responses therefore represents the natural variability that one would expect in receptor impact variables that experience the same hydrological conditions at different locations in the landscape class, together with the experts' uncertainty about this response.

The hydrology modelling produces simulated values from stochastic groundwater and surface water models that describe the uncertain impact of coal resource development pathways (CRDPs) on the hydrological response variables at a particular location (see companion product M09 (as listed in Table 1) for propagating uncertainty through models (Peeters et al., 2016)). The uncertainty from the hydrology modelling is propagated through the receptor impact model at each receptor location to deliver the predicted distribution of receptor impact variables at different time points for the two futures considered by BA (baseline and CRDP). Integrating over all of these receptors produces the overall predicted response of the receptor impact variable for the landscape class given the choice of the BA future. These landscape class results are summarised in product 3-4 (impact and risk analysis) for each bioregion or subregion (Figure 3). The method used for receptor impact model predictions is detailed below.

BAs assume that the only difference between the CRDP and baseline development pathway are those differences captured by the joint distribution of hydrological response variables provided by the stochastic hydrology modelling. All other variables that might also influence the receptor impact variables are assumed to behave in an identical fashion in a future world under baseline conditions, and a future world under the CRDP. To predict how receptor impact variables will respond to future changes in the hydrological response variables, the receptor impact modelling draws surface water and groundwater simulations from the joint distribution function, $p(h(b), h(c))$, where $\mathrm{h(b)}$ represents the hydrological predictions under the baseline pathway ($d = b$) and $\mathrm{h(c)}$ represents hydrological conditions under the CRDP ($d = c$). For some hydrological response variables, such as surface water hydrological response variables that are aggregated to 30-year periods (Figure 11), the stochastic hydrological model output varies depending on the period of interest, $h(d, \tau_{ref})$, $h(d, \tau_{short})$ and $h(d, \tau_{long})$. Other hydrological response variables, however, are defined over the entire future period, such as the maximum depth of groundwater drawdown in the future period, in which case the values are the same for the short and long period, $h(d, \tau_{short}) = h(d, \tau_{long})$.

Figure 14 shows the temporal dependence in the hydrological response variables across years given a choice of development pathway. This is depicted, for example, by the connecting arrows in the first row of Figure 14, which shows that the future hydrological response variables for the baseline development pathway depend on what has occurred in the past. The second row shows that same temporal dependence for hydrological response variables in the coal resource development pathway. Within a particular period, the hydrological response variables may also depend on common factors that are shared across the two development pathways. Thus, Figure 14 shows between-year dependence and within-year dependence, for example, between hydrological response variables in the reference period for the two development pathways.

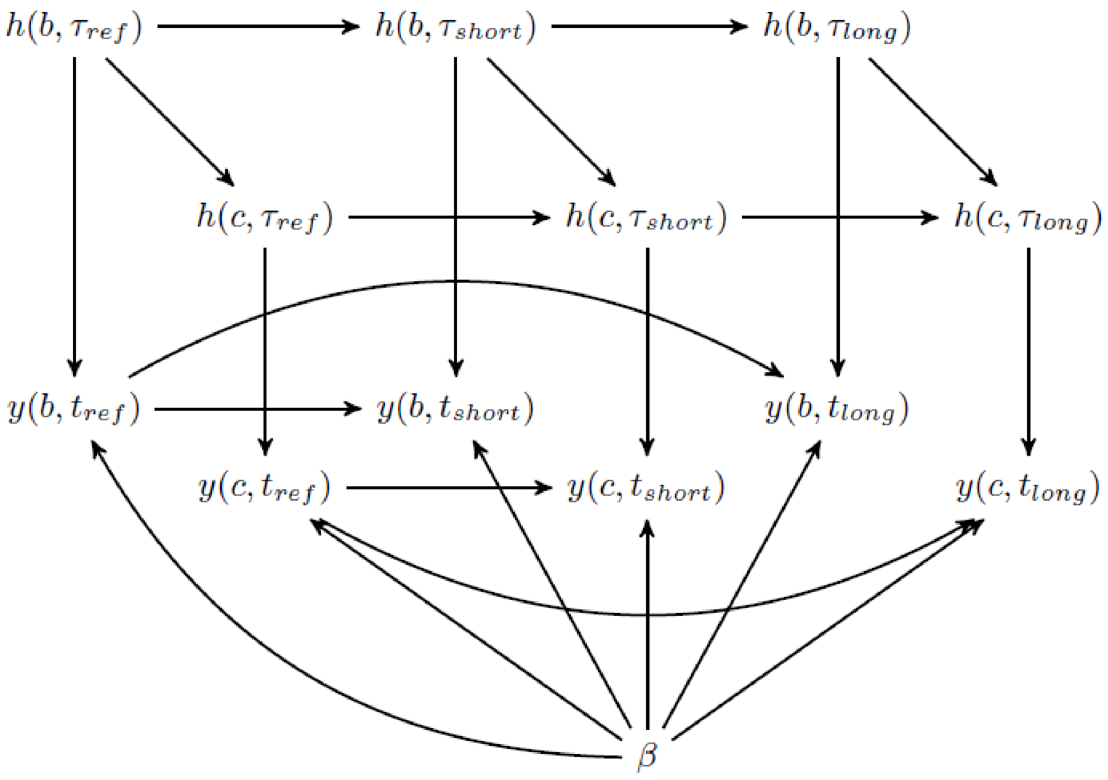

**Figure 14 A directed acyclic graph that shows dependencies among hydrological response variables (*h*) and receptor impact variables (*y*)**

The choices of development pathways are b = baseline and c = crdp for assessment years 2012 (*ref*), 2042 (*short*) and 2102 (*long*). The parameter $\beta$ describes how the receptor impact variable in the current assessment year relates to hydrological response variables in the current assessment year and, for the future assessment years, the receptor impact variable in the 2012 assessment year.

Define the design point $(h', y', t')^{\top}$, which depends on the model structure, the known hydrological response variable values $h'$, the known value for the receptor impact variable in the reference year $y'$, and the assessment year $t'$. In the reference year, the value of $y'$ is fixed at zero and absorbed into the intercept (Section 5.1.2). Conditional on a set of known hydrological response variable values, $h(b,\tau_{ref})$ and $h(c,\tau_{ref})$, the joint distribution of the receptor impact variables in the reference assessment year for the two development pathways (Figure 14) is given by:

$$
\begin{aligned}
&p(y(c,t_{ref}),y(b,t_{ref})|h(c,\tau_{ref}),h(b,\tau_{ref})) = \\
&k^{-1}g^{-1}\left(x(h(c,\tau_{ref}),0,t_{ref})^{\top}\beta\right)g^{-1}\left(x(h(b,\tau_{ref}),0,t_{ref})^{\top}\beta\right)p(\beta)
\end{aligned}
\tag{38}
$$

where $g^{-1}(\cdot)$ is the inverse link function, $p(\beta) = N(\beta|\mu,\Sigma)$ is the elicited prior and the normalising constant is $k = \int g^{-1}\left(x(h(c,\tau_{ref}),0,t_{ref})^{\top}\beta\right)g^{-1}\left(x(h(b,\tau_{ref}),0,t_{ref})^{\top}\beta\right)p(\beta)d\beta$.

The joint distribution of the receptor impact variables for both development pathways in the short-term assessment year is conditioned on the hydrological response variable values in the short-term assessment period, and also the receptor impact variables in the reference assessment year (Figure 14). This joint distribution conditional on the hydrological response variables is given by:

$$
\begin{aligned}
&p\left(y(c,t_{short}),y(b,t_{short})\middle|h(c,\tau_{short}),h(b,\tau_{short}),h(c,\tau_{ref}),h(b,\tau_{ref})\right) \\
&\propto \int g^{-1}\left(x(h(c,\tau_{short}),y(c,t_{ref}),t_{short})^{\top}\beta\right)g^{-1}\left(x(h(b,\tau_{short}),y(b,t_{ref}),t_{short})^{\top}\beta\right) \\
&\times p(y(c,t_{ref}),y(b,t_{ref})|h(c,\tau_{ref}),h(b,\tau_{ref}))p(\beta)dy(c,t_{ref})dy(b,t_{ref})
\end{aligned}
\tag{39}
$$

Similarly, the distribution of the receptor impact variable in the long-term assessment year conditional on hydrological response variables is given by:

$$
\begin{aligned}
&p\left(y(c,t_{long}),y(b,t_{long})\middle|h(c,\tau_{long}),h(b,\tau_{long}),h(c,\tau_{ref}),h(b,\tau_{ref})\right) \\
&\propto \int g^{-1}\left(x(h(c,\tau_{long}),y(c,t_{ref}),t_{short})^{\top}\beta\right)g^{-1}\left(x(h(b,\tau_{long}),y(b,t_{ref}),t_{long})^{\top}\beta\right) \\
&\times p(y(c,t_{ref}),y(b,t_{ref})|h(c,\tau_{ref}),h(b,\tau_{ref}))p(\beta)dy(c,t_{ref})dy(b,t_{ref})
\end{aligned}
\tag{40}
$$

Realisations from the joint distributions in Equations 38, 39 and 40 are obtained using Monte Carlo simulation (Section 8.2). During these simulations, the BAs impose perfect positive dependence in the samples drawn from $p(\beta)$ between development pathways within an assessment year in accordance with the assumption of BAs that, after accounting for the effect of hydrological response variables, receptor impact variables behave in an identical fashion under the baseline and CRDP.

Consider predictions for a particular location, or 'assessment unit' (defined as a geographic area that is used to partition the entire assessment extent into square polygons that do not overlap), denoted $a_i$. The predicted distributions for the assessment and future years are given by:

$$\begin{aligned} p\big(y(c,t_{ref}),y(b,t_{ref})\big|a_i\big) = {} & \int p\left(y(c,t_{ref}),y(b,t_{ref})\middle|h(c,\tau_{ref}),h(b,\tau_{ref})\right) \\ & \times\ p\big(h(c,\tau_{ref}),h(b,\tau_{ref})|a_i\big)dh(c,\tau_{ref})dh(b,\tau_{ref}) \\ p(y(c,t_{short}),y(b,t_{short})|a_i) = {} & \int p(y(c,t_{short}),y(b,t_{short})|h) \\ & \times p(h(c,\tau_{short}),h(b,\tau_{short}),h(c,\tau_{ref}),h(b,\tau_{ref})|a_i) \\ & \times dh(c,\tau_{short})dh(b,\tau_{short})dh(c,\tau_{ref})dh(b,\tau_{ref}) \\ p\big(y(c,t_{long}),y(b,t_{long})\big|a_i\big) = {} & \int p\big(y(c,t_{long}),y(b,t_{long})\big|h\big) \\ & \times p\big(h(c,\tau_{long}),h(b,\tau_{long}),h(c,\tau_{ref}),h(b,\tau_{ref})|a_i\big) \\ & \times dh(c,\tau_{long})dh(b,\tau_{long})dh(c,\tau_{ref})dh(b,\tau_{ref}) \end{aligned} \qquad (41)$$

The distributions of the hydrological response variables in Equation 41 depend on both the choice of development pathway and the assessment unit. Simulated values from the joint distribution of the hydrological response variables for each assessment unit, $p(h(c,\tau),h(b,\tau)|a_i)$ are provided by the surface water and groundwater models.

In addition to the above predictions, it is also of interest to consider functions of these unknowns. Two possibilities are considered for the future period: the actual change, $\delta(t) = y(c,t) - y(t,b)$, and the relative change, $\rho(t) = \delta(t)/\ y(b,t)$ for $t \in \{t_{short}, t_{long}\}$. These predictions depend on the hydrological response variables and are thus also spatially explicit, that is, dependent on the choice of the assessment unit:

$$\begin{aligned} & p(\delta(t)|a_i) = p(y(c,t) - y(b,t)|a_i) \\ & p(\rho(t)|a_i) = p\left(\frac{\delta(t)}{y(b,t)}|a_i\right) \end{aligned} \qquad (42)$$

for the future assessment with $t \in \{t_{short}, t_{long}\}$.

## 8.1 Prediction to landscape class

The companion submethodology M10 (as listed in Table 1 and Figure 3) for analysing impacts and risks (Henderson et al., 2018) lays out the definition of a receptor impact model in terms of what it needs to predict and the appropriate conditioning and uncertainty that it should incorporate. As discussed in companion submethodology M10 (Henderson et al., 2018), a receptor impact model is defined with respect to a particular landscape classification. It is associated with a particular receptor impact variable and hydrological response variables chosen from the conceptual modelling output. It is explicit about the temporal definition of the receptor impact variable. It describes how the receptor impact variable changes as the hydrological response variables change across the entire landscape class.

The receptor impact model predicts the value of the receptor impact variable at all locations across the landscape class that experience specific hydrological conditions. The uncertainty in these predictions represents the natural variability in the value of the receptor impact variable that would occur at different locations within the landscape class that experience the same hydrological conditions, and the experts' uncertainty about how receptor impact variables respond to different hydrological conditions. The predicted value of a receptor impact variable is thus not the predicted response at a particular location within a landscape, which would depend on many additional localised factors such as land use.

The receptor impact models are designed to facilitate prediction to the landscape class level while integrating across the spatially explicit hydrological response variables. This integration requires a quantifiable estimate of the amount of each landscape class within each assessment unit. Thus, a non-negative weight $w_i$ is associated with each of the $A$ assessment units, $a_i$, $i = 1, \ldots, A$. The weight is defined by the amount of the landscape class within each assessment unit. For linear landscape classes defined along stream reaches, it is the length of reach assigned to the landscape class within each assessment unit. For areal landscape classes, it is the area assigned to the landscape class within each assessment unit.

Since it is assumed that the receptor impact variable predictions for each assessment unit are conditionally independent given the hydrological response variables (Chapter 8), the landscape class averaged predictions are given by the weighted averages:

$$y(d,t)|LC = \frac{1}{\sum_{i=1}^{A} w_i} \sum_{i=1}^{A} w_i y(d,t)|a_i \tag{43}$$

where $p(y(c,t)|a_i) = \int p(y(c,t), y(b,t)|a_i) dy(b,t)$, $p(y(b,t)|a_i) = \int p(y(c,t), y(b,t)|a_i) dy(c,t)$ and $LC$ denotes dependence for the landscape class average on all assessment units, $a_i$, $i = 1, \ldots, A$, that form the landscape class. Aggregating to the landscape class requires conditioning on all assessment units in the landscape class.

The landscape class averaged actual change and relative change are obtained by:

$$\begin{aligned} \delta(t)|LC &= \frac{1}{\sum_{i=1}^{A} w_i} \sum_{i=1}^{A} w_i\, \delta(t)|a_i \\ \rho(t)|LC &= \frac{1}{\sum_{i=1}^{A} w_i} \sum_{i=1}^{A} w_i \rho(t)|a_i \end{aligned} \tag{44}$$

which is a weighted average of the individual assessment units. Relative change can be expressed as percentage change by multiplying the relative change by 100.

## 8.2 Monte Carlo approximation

The realisations of the unknown coefficients $\beta'$ are independent of the hydrological response variables (Figure 14). In particular, the realisations of $\beta'$ are independent across the assessment units. Spatial dependence may nevertheless be introduced into the receptor impact variable predictions across assessment units to the extent that spatial dependence is captured by the hydrological response variables.

Monte Carlo approximations to the above joint distributions are available through the method of composition (Tanner, 1996). To sample from $p(y(c,t_{ref}), y(b,t_{ref})\ |a_i)$ (Equation 41):

1. Draw $(h(c,\tau_{ref})', h(b,\tau_{ref})', h(c,\tau_{short})', h(b,\tau_{short})', h(c,\tau_{long})', h(b,\tau_{long})')$ from $p(h|a_i)$.
2. Draw $\beta'$ from $p(\beta)$.
3. Calculate $(y(c,t_{ref})', y(b,t_{ref})')$ from $p\left(y(c,t_{ref}), y(b,t_{ref})|h(c,\tau_{ref})', h(b,\tau_{ref})', \beta'\right)$.

Repeat the above steps $J$ times to obtain $J$ simulations and store all values. Samples from the marginal $p(y(d,t_{ref})|a_i)$ are obtained by considering only the simulations $y(d,t_{ref})'$.

Next, to sample from $p(y(c,t_{short}), y(b,t_{short})|a_i)$:

4. Calculate $(y(c,t_{short})', y(b,t_{short})')$ from $p(y(c,t_{short}), y(b,t_{short})|h(c,\tau_{ref})', h(b,\tau_{ref})', y(c,t_{short})', y(b,t_{short})', \beta')$.

Repeat $J$ times.

Similarly, to sample from $p(y(c,t_{long}), y(b,t_{long})|a_i)$:

5. Calculate $(y(c,t_{long})', y(b,t_{long})')$ from $p(y(c,t_{long}), y(b,t_{long})|h(c,\tau_{ref})', h(b,\tau_{ref})', y(c,t_{long})', y(b,t_{long})', \beta')$.

Repeat $J$ times for each assessment unit $a_i$.

Given the $J$ simulated values of the jointly dependent receptor impact variable $y$, Monte Carlo approximations of any function of $y$, $\zeta = f(y)$ are also available. For example, $J$ simulated values of the actual change and relative change for the short-term assessment year are given by:

$$\begin{aligned}\delta(t_{short}|\ a_i)' &= (y(c,t_{short})' - y(b,t_{short})'|a_i)\\ \rho(t_{short}|\ a_i)' &= \delta(t_{short}|\ a_i)'/(y(b,t_{short})'|a_i)\end{aligned} \tag{45}$$

and likewise for the long-term assessment year, $t_{long}$. Predictions are thus available for all assessment units.

Aggregated predictions to the landscape class are also available with simulated values of the landscape class level weighted averages. Given $J$ simulated values of a quantity $\zeta|a_i$ for each assessment unit, such as a prediction of the receptor impact variable or some other function such as the actual or relative change, the landscape class prediction is given by:

$$\zeta'|LC = \frac{1}{\sum_{i=1}^{A} w_i}\sum_{i=1}^{A}(\zeta'|a_i)w_i. \tag{46}$$

For each landscape class and for each assessment unit, a set of quantiles, which includes the 5th, 10th, 15th, 20th, 25th, 30th, 35th, 40th, 45th, 50th, 55th, 60th, 65th, 70th, 75th, 80th, 85th, 90th, and 95th quantiles, was estimated through the above Monte Carlo approach for each unknown quantity of interest that depends on the receptor impact variable. Let $P^{-1}(u)$ be the inverse cumulative distribution function for $0 \leq u \leq 1$. The $q$th quantile is given by, $\gamma = P(q)$ for $0 \leq q \leq 1$. Let the empirical cumulative distribution function of the $J$ samples of $\zeta$ be given by $P_J(s)$. A Monte Carlo estimate of the $q$th quantile is then given by:

$$\hat{\gamma} = \operatorname*{argmin}_{s} P_J(s) \geq q \qquad (47)$$

Extreme quantiles are more sensitive to approximation error by the Monte Carlo method, and larger sample sizes reduce Monte Carlo error. The number of receptor impact variable simulations can be made arbitrarily large. However, the composition sampling that preserves the joint dependence between the receptor impact variables and the hydrological response variables is limited by the number of realisations available from the hydrological modelling simulations.

Rather than quantiles, it may also be of interest to report the average of the above unknown quantities either at the level of assessment units or landscape class or both. A mean estimate for an unknown quantity is trivially obtained by taking the sample mean of the above Monte Carlo simulations.

## 8.3 Conclusion

The Monte Carlo simulations permit the probabilistic prediction of receptor impact variables. The receptor impact models include the ability to capture direct, indirect and cumulative impacts through the hydrological response variables. In addition, ecological lags imposed through the receptor impact variable are also captured by the above method. For example, the long-term response for a long-lived terrestrial species such as large woody riparian vegetation may be more sensitive to the previous history of the receptor within an assessment unit compared to a short-lived aquatic species. The approach can flexibly estimate a function of the receptor impact variable, such as actual or relative change, quantiles or averages. The range or distribution of receptor impact variable outcomes will be summarised for baseline (all years) and for the change due to additional coal resource development for future assessment years at landscape class level. The change due to additional coal resource development may be expressed in relative terms or actual terms and may depend on the specific receptor impact variable. These responses can be aggregated from the assessment units to the landscape class level to assess overall impacts of development within the bioregion or subregion.

# 9 Content for product 2.7 (receptor impact modelling)

Although some examples are interspersed throughout this submethodology, it is recommended that M08 be read in conjunction with its application in product 2.7 (receptor impact modelling) for each bioregion or subregion. The structure in product 2.7 closely follows the methodology outlined in M08. Table 11 shows the recommended content for product 2.7. It identifies those landscape classes (or landscape groups) that may experience some hydrological change and those that are very unlikely to do so. For those that do, product 2.7 provides a summary of the landscape class (or landscape group) and how it works, details the qualitative mathematical modelling expert workshop and sign-directed graph output for that landscape class, describes the choice of hydrological response variables and receptor impact variables selected, summarises the elicitation scenarios presented to experts, and concludes with a description of the receptor impact model that is constructed and any interpretations that may be made around it.

Table 11 Recommended content for product 2.7 (receptor impact modelling)

| Section number | Title of section | Main content to include in section |
|---|---|---|
| 2.7.1 | Methods | **Summary**<br>**Outline:**<br>• construction of qualitative models for landscape classes<br>• choice of hydrological impact variables and receptor impact variables<br>• construction of receptor impact models<br>• expert elicitation process |
| 2.7.2 | Prioritising landscape classes for receptor impact modelling | **Summary**<br>**Outline:**<br>• highlight landscape classes that are considered (and those that are not) based on the hydrological and conceptual modelling<br>• discuss connectivity between landscape classes, if appropriate |

| Section number | Title of section | Main content to include in section |
|---|---|---|
| 2.7.3 | Landscape class #1 | **Summary**<br>**2.7.3.1 Description**<br>• general description of landscape class (typically based on review and collation ahead of qualitative modelling workshop); provides opportunity to include some of that evidence base<br>**2.7.3.2 Qualitative model**<br>• both results (from the qualitative modelling workshop) and narrative<br>• highlight key assets where possible<br>• include evidence base<br>• include signed digraph and analysis for every landscape class<br>**2.7.3.3 Choice of hydrological response variables and receptor impact variables**<br>On the basis of the advice from both experts and the Assessment team:<br>• describe choice of HRVs and RIVs from the qualitative models<br>• describe how those variables are represented in the surface water and groundwater models<br>**2.7.3.4 Elicitation scenarios**<br>• for ecological assets: these are the elicitation scenarios considered and questions asked of experts – to find out relationship between HRVs and RIVs<br>• details about the scenarios (e.g. questions) should be compiled as a document (doesn't need to align with BA look and feel); this document then should be registered as a dataset and cited here<br>**2.7.3.5 Receptor impact model**<br>• outline results from elicitation [data]<br>• describe and interpret the statistical model (based on data) that describes ΔRIV given one or more hydrological response variables and the different reporting times<br>• note: this is just the model – not the application of the model; instead, the results arising from applying this model are presented in product 3-4 |
| 2.7.4 | Landscape class #2 | As per above |
| 2.7.5 | Add more landscape classes as appropriate | As per above |
| 2.7.6 | Limitations and gaps | This is the final section of this product. The numbered section may vary depending on the number of landscape class sections prior to this section. |

HRV = hydrological response variable, RIV = receptor impact variable

The receptor impact modelling methodology (this product) and its implementation was affected by design choices that have been made within BA. Some of these broader choices are described in companion submethodology M10 (as listed in Table 1) for analysing impacts and risks (Henderson et al., 2018). Table 12 summarises some of the assumptions made for the receptor impact modelling, the implications of those assumptions for the results, and how those implications are acknowledged through the BA products.

**Table 12 Summary of the receptor impact modelling assumptions, the implications of those assumptions and how the potential implications are acknowledged through bioregional assessment products**

| Assumptions of receptor impact modelling | Implications | Acknowledgement |
|---|---|---|
| Discretisation of continuous landscape surface | Provided a defined spatial scope for experts to address. Connections between landscape classes broken. Changes in one landscape class may have implications for adjacent landscape classes | Identify potential connections between landscape classes where possible in the impact and risk product |
| Data underpinning landscape classes (omissions / incorrect attribution) | Landscape class definition required data input from pre-existing data sources. Prioritisation for qualitative mathematical models and RIMs may be affected. Minimal effect on model development for RIMs | Acknowledge issues with data in the impact and risk product (also done in the conceptual modelling product); in product 3-4 acknowledge that mapped results reflect the mapped inputs |
| Areas of landscape classes are constant over modelling period | Provided a defined spatial scope for experts to address. BA is about identifying existing areas that are at risk from coal resource development as opposed to predicting the changes in areal extent or transition to different landscape classes. Some potential exists for changes in the area of the landscape class to affect its sensitivity to hydrological change but this would need to be assessed on an asset by asset basis. | Acknowledge in Methods |
| Other developments and users of water (e.g. agriculture) are constant over time | Provided a defined context for experts to consider. BA is about identifying existing areas that are at risk from coal resource development as opposed to predicting the changes due to other developments or the relative attribution | Acknowledge in Methods |
| Landscape characteristics other than hydrological variables are not represented in quantitative RIMs | Refined scope for experts to how RIMs were associated with hydrological variables that could be provided by hydrological models developed by BA. Loss of within-landscape class predictive performance from the RIMs | Identify as knowledge gap where model does not represent some dependencies that are not captured by statistical dependencies with the chosen hydrological response variables. Acknowledge importance of local (versus regional) analyses where the concern is over particular parts of a landscape class |
| Selection of experts, limited expert availability, and impact on represented domain knowledge and expertise | Experts provided domain expertise and experience that informed both model structure and provided quantifiable predictions of RIV response to novel hydrological scenarios. Expert availability affected the quality/utility of the qualitative mathematical model; identification of RIVs that reflect expertise of those in the room | Acknowledge that the receptor impact variable is an ‘indicator’ of the potential ecosystem response. Identify as knowledge gap where part of the landscape class is not represented |

| Assumptions of receptor impact modelling | Implications | Acknowledgement |
|---|---|---|
| Simplification of complex systems | Provided formal approach to model identification and selection of candidate RIVs. Not all components and relationships are represented by receptor impact models | Acknowledge that one or two RIVs can under estimate complex ecosystem function; make assumptions clear; high-level interpretation of results; emphasise importance of interpreting the hydrological change |
| The common set of modelled hydrological response variables are used across each landscape class | Refined scope for experts to how RIMs were associated with hydrological variables that could be provided by hydrological models developed by BA. Enables some simplification of complex systems. Loss of local specificity in predictions of receptor impact variables | The need for local-scale information is identified (in multiple places) |
| RIV selection (assumption that RIV is good indicator of ecosystem response) | The qualitative models informed the selection of RIVs within the additional constraints imposed by expert availability given project timelines. Focus of the quantified relationships within the landscape class | The need for local-scale information is identified (in multiple places) |
| Extrapolation of predictions beyond elicitation scenarios | The ranges of hydrological scenarios to be considered at the expert elicitation sessions were informed by preliminary hydrological modelling output and hydrological expert advice within BA. However, final model results sometimes extended beyond this preliminary range due to necessary changes in underlying hydrological modelling assumptions and assimilation of data. Extrapolation beyond the range of hydrological response variables considered by the expert elicitation increases uncertainty in receptor impact variable predictions | Identify as a limitation for the appropriate landscape class in the impact and risk product where this occurs |
| For qualitative models, focus on impacts of long-term sustained hydrological changes (press perturbations) to ecosystems. Note that quantitative RIMs can and do account for pulse perturbations and associated RIV responses | Qualitative models may under-represent impacts of shorter-term hydrological changes (pulse perturbations) on ecosystems and landscape classes | Describe rationale for the focus on press perturbations in the receptor impact modelling submethodology (this product). Note that many potential pulse perturbations are caused by accidents and managed by site-based processes. Identify as a limitation / knowledge gap. Note that quantitative RIMs do account for pulse perturbations |

BA = bioregional assessment, HRV = hydrological response variable, RIM = receptor impact model, RIV = receptor impact variable

Product 2.7 (receptor impact modelling) does not cover results for the prediction of receptor impact variables. These results are addressed as part of product 3-4 (impact and risk analysis). Companion submethodology M10 (as listed in Table 1) for analysing impacts and risks (Henderson et al., 2018) describes some of the additional underlying methodology for these results, and in particular the choice of summary and aggregation of predictions to landscape classes and water-dependent assets. These results for landscape classes and assets are derived by applying the methodology described in Section 8.1 and Section 8.2.

# References

Barrett DJ, Couch CA, Metcalfe DJ, Lytton L, Adhikary DP and Schmidt RK (2013) Methodology for bioregional assessments of the impacts of coal seam gas and coal mining development on water resources. A report prepared for the Independent Expert Scientific Committee on Coal Seam Gas and Large Coal Mining Development through the Department of the Environment, Australia. Viewed 12 April 2017, http://data.bioregionalassessments.gov.au/submethodology/bioregional-assessment-methodology.

Boulton A, Brock M, Robson B, Ryder D, Chambers J and Davis J (2014) Australian freshwater ecology: processes and management. Second edition. John Wiley & Sons.

Chaloner K and Verdinelli I (1995) Bayesian experimental design: a review. Statistical Science 10(3), 273–304.

Dambacher JM and Ramos-Jiliberto R (2007) Understanding and predicting effects of modified interactions through a qualitative analysis of community structure. The Quarterly Review of Biology 82, 227–250.

Dambacher JM, Li HW, Rossignol PA (2002) Relevance of community structure in assessing indeterminacy of ecological predictions. Ecology 83, 1372–1385.

Dambacher JM, Li HW, Rossignol PA (2003a) Qualitative predictions in model ecosystems. Ecological Modelling 161, 79–93.

Dambacher JM, Luh H-K, Li HW, Rossignol PA (2003b) Qualitative stability and ambiguity in model ecosystems. American Naturalist 161, 876–888.

Federov VV (1972) Theory of optimal experiments. Academic Press, New York.

Ford JH, Hayes KR, Henderson BL, Lewis S, Baker PA and Schmidt RK (2016) Systematic analysis of water-related hazards associated with coal resource development. Submethodology M11 from the Bioregional Assessment Technical Programme. Department of the Environment and Energy, Bureau of Meteorology, CSIRO and Geoscience Australia, Australia. http://data.bioregionalassessments.gov.au/submethodology/M11.

Henderson BL, Barry S, Hayes KR, Hosack G, Holland K, Herron N, Mount R, Schmidt RK, Dambacher J, Ickowicz A, Lewis S and Post DA (2018) Impacts and risks. Submethodology M10 from the Bioregional Assessment Technical Programme. Department of the Environment and Energy, Bureau of Meteorology, CSIRO and Geoscience Australia, Australia. http://data.bioregionalassessments.gov.au/submethodology/M10.

Henderson BL, Hayes KR, Mount R, Schmidt RK, O'Grady A, Lewis S, Holland K, Dambacher J, Barry S and Raiber M (2016) Developing the conceptual model of causal pathways. Submethodology M05 from the Bioregional Assessment Technical Programme. Department of the Environment and Energy, Bureau of Meteorology, CSIRO and Geoscience Australia, Australia. http://data.bioregionalassessments.gov.au/submethodology/M05.

Hosack GR, Hayes KR and Dambacher JM (2008) Assessing model structure uncertainty through an analysis of system feedback and Bayesian networks. Ecological Applications 18, 1070–1082.

Hosack GR, Hayes KR and Barry SC (2016) Expert elicitation for the effectiveness of risk control options: final report to the Australian Maritime Safety Authority. CSIRO, Hobart, Australia.

Hosack GR, Hayes KR and Barry SC (2017) Prior elicitation for Bayesian generalised linear models with application to risk control option assessment. Reliability Engineering and System Safety 167, 351–361.

Hosack G, Ickowicz A, Hayes KR, Macfarlane C, Dambacher J, Marvanek SP, Zhang YQ, Viney NR, Peeters LJM, van Rooyen S, McVicar TR, O'Grady A, Barry S and Henderson BL (2018) Receptor impact modelling for the Gloucester subregion. Product 2.7 for the Gloucester subregion from the Northern Sydney Basin Bioregional Assessment. Department of the Environment and Energy, Bureau of Meteorology, CSIRO and Geoscience Australia, Australia. http://data.bioregionalassessments.gov.au/product/NSB/GLO/2.7.

Lake PS (2000) Disturbance, patchiness, and diversity in streams, Journal of the North American Benthological Society 19(4), 573–592.

Lay DC (2003) Linear algebra and its applications. Third edition. Addison-Wesley, Boston, Massachusetts, USA.

Levins R (1966) The strategy of model building in population biology. American Scientist 54, 421–431.

Levins R (1968) Evolution in changing environments: some theoretical explorations. Princeton University Press, Princeton, New Jersey, USA.

Levins R (1974) The qualitative analysis of partially specified systems. Annals of the New York Academy of Sciences 231, 123–138.

Levins R (1975) Evolution in communities near equilibrium. In: Cody M and Diamond JM (eds) Ecology and evolution of communities. Harvard University Press, Cambridge, Massachusetts, USA, 16–50.

Levins R (1998) Qualitative mathematics for understanding, prediction and intervention in complex ecosystems. In: Rapport D, Costanza R, Epstein PR, Gaudet C and Levins R (eds) Ecosystem health: principles and practices. Blackwell Science, Malden, Massachusetts, USA, 178–204.

Marcot BG, Holthausen RS, Raphael MG, Rowland MM and Wisdom MJ (2001) Using Bayesian belief networks to evaluate fish and wildlife population viability under land management alternatives from an environmental impact statement. Forest Ecology and Management 153, 29–42.

McCullagh P and Nelder JA (1989) Generalized Linear Models. Second edition. Chapman and Hall/CRC, Boca Raton, Florida.

NSW Office of Water (2012) NSW Aquifer interference policy. NSW Government policy for the licensing and assessment of aquifer interference activities. Department of Primary Industries, NSW Office of Water, September 2012.

O'Grady AP, Mount R, Holland K, Sparrow A, Crosbie R, Marston F, Dambacher J, Hayes K, Henderson B, Pollino C and Macfarlane C (2016) Assigning receptors to water-dependent assets. Submethodology M03 from the Bioregional Assessment Technical Programme. Department of the Environment and Energy, Bureau of Meteorology, CSIRO and Geoscience Australia, Australia. http://data.bioregionalassessments.gov.au/submethodology/M03.

O'Hagan A (1998) Eliciting expert beliefs in substantial practical applications. Journal of the Royal Statistical Society: Series D (The Statistician) 47, 21–35.

O'Hagan A, Buck CE, Daneshkhah A, Eiser JR, Garthwaite PH, Jenkinson DJ, Oakley JE and Rakow T (2006) Uncertain judgements: eliciting experts' probabilities. John Wiley & Sons, Chichester.

Peeters L, Pagendam D, Gao L, Hosack G, Jiang W and Henderson B (2016) Propagating uncertainty through models. Submethodology M09 from the Bioregional Assessment Technical Programme. Department of the Environment and Energy, Bureau of Meteorology, CSIRO and Geoscience Australia, Australia. http://data.bioregionalassessments.gov.au/submethodology/M09.

Puccia CJ and Levins R (1985) Qualitative modeling of complex systems: an introduction to loop analysis and time averaging. Harvard University Press, Cambridge, Massachusetts, USA.

Schmitz OJ (1998) Direct and indirect effects of predation and predation risk in old-field interaction webs. The American Naturalist 151, 327–342.

Schwarz G (1978) Estimating the dimension of a model. Annals of Statistics 6(2), 461–464.

Suter II GW (2016) Ecological risk assessment. CRC Press, Boca Raton.

Tanner MA (1996) Tools for statistical inference. Third edition. Springer-Verlag, New York.

Viney N (2016) Surface water modelling. Submethodology M06 from the Bioregional Assessment Technical Programme. Department of the Environment and Energy, Bureau of Meteorology, CSIRO and Geoscience Australia, Australia. http://data.bioregionalassessments.gov.au/submethodology/M06.

Wheeler B (2014) AlgDesign: algorithmic experimental design. R package version 1.1-7.3. Viewed 13 February 2017, https://CRAN.R-project.org/package=AlgDesign.

Wheeler RE (2004) optFederov. AlgDesign. The R project for statistical computing. Viewed 13 February 2017, http://www.r-project.org/.

# Glossary

The register of terms and definitions used in the Bioregional Assessment Programme is available online at http://environment.data.gov.au/def/ba/glossary (note that terms and definitions are respectively listed under the 'Name' and 'Description' columns in this register). This register is a list of terms, which are the preferred descriptors for concepts. Other properties are included for each term, including licence information, source of definition and date of approval. Semantic relationships (such as hierarchical relationships) are formalised for some terms, as well as linkages to other terms in related vocabularies.

additional coal resource development: all coal mines and coal seam gas (CSG) fields, including expansions of baseline operations, that are expected to begin commercial production after December 2012

annual flow: the volume of water that discharges past a specific point in a stream in a year, commonly measured in GL/year. This is typically reported as the maximum change due to additional coal resource development over the 90-year period (from 2013 to 2102).

aquifer: rock or sediment in a formation, group of formations, or part of a formation that is saturated and sufficiently permeable to transmit quantities of water to bores and springs

assessment extent: the geographic area associated with a subregion or bioregion in which the potential water-related impact of coal resource development on assets is assessed. The assessment extent is created by revising the preliminary assessment extent on the basis of information from Component 1: Contextual information and Component 2: Model-data analysis.

assessment unit: for the purposes of impact analysis, a geographic area that is used to partition the entire assessment extent into square polygons that do not overlap. The spatial resolution of the assessment units is closely related to that of the bioregional assessment groundwater modelling and is, typically, 1 x 1 km. Each assessment unit has a unique identifier. The partitioned data can be combined and recombined into any aggregation supported by the conceptual modelling, causal pathways and model data.

asset: an entity that has value to the community and, for bioregional assessment purposes, is associated with a subregion or bioregion. Technically, an asset is a store of value and may be managed and/or used to maintain and/or produce further value. Each asset will have many values associated with it and they can be measured from a range of perspectives; for example, the values of a wetland can be measured from ecological, sociocultural and economic perspectives.

asset element: individual spatial features – points, lines and polygons – that describe an asset spatially

baseline coal resource development: a future that includes all coal mines and coal seam gas (CSG) fields that are commercially producing as of December 2012 bioregion: a geographic land area within which coal seam gas (CSG) and/or coal mining developments are taking place, or could take place, and for which bioregional assessments (BAs) are conducted

bioregional assessment: a scientific analysis of the ecology, hydrology, geology and hydrogeology of a bioregion, with explicit assessment of the potential direct, indirect and cumulative impacts of coal seam gas and coal mining development on water resources. The central purpose of bioregional assessments is to analyse the impacts and risks associated with changes to water-dependent assets that arise in response to current and future pathways of coal seam gas and coal mining development.

bore: a narrow, artificially constructed hole or cavity used to intercept, collect or store water from an aquifer, or to passively observe or collect groundwater information. Also known as a borehole or piezometer.

causal pathway: for the purposes of bioregional assessments, the logical chain of events – either planned or unplanned – that link coal resource development and potential impacts on water resources and water-dependent assets

coal resource development pathway: a future that includes all coal mines and coal seam gas (CSG) fields that are in the baseline as well as those that are expected to begin commercial production after December 2012

component: for the purposes of Impact Modes and Effects Analysis (IMEA), a group of activities associated with a coal seam gas (CSG) operation or coal mine. For example, components during the development life-cycle stage of a coal mine include developing the mine infrastructure, the open pit, surface facilities and underground facilities. Components are grouped into life-cycle stages.

conceptual model: abstraction or simplification of reality

connectivity: a descriptive measure of the interaction between water bodies (groundwater and/or surface water)

context: the circumstances that form the setting for an event, statement or idea

cumulative impact: for the purposes of bioregional assessments, the total change in water resources and water-dependent assets resulting from coal seam gas and coal mining developments when all past, present and reasonably foreseeable actions that are likely to impact on water resources are considered

dataset: a collection of data in files, in databases or delivered by services that comprise a related set of information. Datasets may be spatial (e.g. a shape file or geodatabase or a Web Feature Service) or aspatial (e.g. an Access database, a list of people or a model configuration file).

direct impact: for the purposes of bioregional assessments, a change in water resources and water-dependent assets resulting from coal seam gas and coal mining developments without intervening agents or pathways *dmax*: maximum difference in drawdown, obtained by choosing the maximum of the time series of differences between two futures. For example, to calculate the difference in drawdown between the coal resource development pathway (CRDP) and baseline, use the equations dmax = max (dCRDP(t) – dbaseline(t)) where d is drawdown, or dmax = max (hbaseline(t) – hCRDP(t)) where h is groundwater level and t is time.

dmaxRef: maximum difference in drawdown under the baseline future or under the coal resource development pathway future relative to the reference period (1983 to 2012). This is typically reported as the maximum change due to additional coal resource development.

drawdown: a lowering of the groundwater level (caused, for example, by pumping). In the bioregional assessment (BA) context this is reported as the difference in groundwater level between two potential futures considered in BAs: baseline coal resource development (baseline) and the coal resource development pathway (CRDP). The difference in drawdown between CRDP and baseline is due to the additional coal resource development (ACRD). Drawdown under the baseline is relative to drawdown with no coal resource development; likewise, drawdown under the CRDP is relative to drawdown with no coal resource development.

ecosystem: a dynamic complex of plant, animal, and micro-organism communities and their non-living environment interacting as a functional unit. Note: ecosystems include those that are human-influenced such as rural and urban ecosystems.

effect: for the purposes of Impact Modes and Effects Analysis (IMEA), change in the quantity and/or quality of surface water or groundwater. An effect is a specific type of an impact (any change resulting from prior events).

formation: rock layers that have common physical characteristics (lithology) deposited during a specific period of geological time

fractile: the value of a distribution below which some fraction of the sample lies. For example, the 0.95-fractile is the value below which there is a probability of 0.95 occurrence (or equivalently, 95% of the values lie below the 0.95-fractile).

Gloucester subregion: The Gloucester subregion covers an area of about 348 km². The Gloucester subregion is defined by the geological Gloucester Basin. It is located just north of the Hunter Valley in NSW, approximately 85 km north-north-east of Newcastle and relative to regional centres is 60 km south-west of Taree and 55 km west of Forster.

groundwater: water occurring naturally below ground level (whether stored in or flowing through aquifers or within low-permeability aquitards), or water occurring at a place below ground that has been pumped, diverted or released to that place for storage there. This does not include water held in underground tanks, pipes or other works.

groundwater zone of potential hydrological change: outside this extent, groundwater drawdown (and hence potential impacts) is very unlikely (less than 5% chance). It is the area with a greater than 5% chance of exceeding 0.2 m of drawdown due to additional coal resource development in the relevant aquifers.

hazard: an event, or chain of events, that might result in an effect (change in the quality and/or quantity of surface water or groundwater)

hydrogeology: the study of groundwater, including flow in aquifers, groundwater resource evaluation, and the chemistry of interactions between water and rock

hydrological response variable: a hydrological characteristic of the system that potentially changes due to coal resource development (for example, drawdown or the annual flow volume)

impact: a change resulting from prior events, at any stage in a chain of events or a causal pathway. An impact might be equivalent to an effect (change in the quality and/or quantity of surface water or groundwater), or it might be a change resulting from those effects (for example, ecological changes that result from hydrological changes).

impact mode: the manner in which a hazardous chain of events (initiated by an impact cause) could result in an effect (change in the quality and/or quantity of surface water or groundwater). There might be multiple impact modes for each activity or chain of events.

Impact Modes and Effects Analysis: a systematic hazard identification and prioritisation technique based on Failure Modes and Effects Analysis

indirect impact: for the purposes of bioregional assessments, a change in water resources and water-dependent assets resulting from coal seam gas and coal mining developments with one or more intervening agents or pathways

landscape class: for bioregional assessment (BA) purposes, an ecosystem with characteristics that are expected to respond similarly to changes in groundwater and/or surface water due to coal resource development. Note that there is expected to be less heterogeneity in the response within a landscape class than between landscape classes. They are present on the landscape across the entire BA subregion or bioregion and their spatial coverage is exhaustive and non-overlapping. Conceptually, landscape classes can be considered as types of ecosystem assets.

landscape group: for the purposes of bioregional assessments (BAs), a set of landscape classes grouped together based on common ecohydrological characteristics that are relevant for analysis purposes

likelihood: probability that something might happen

low-flow days (LFD): the number of low-flow days per year. This is typically reported as the maximum change due to additional coal resource development over the 90-year period (from 2013 to 2102). The threshold for low-flow days is the 10th percentile from the simulated 90-year period.

low-flow spells (LFS): the number of low-flow spells per year. This is typically reported as the maximum change due to additional coal resource development over the 90-year period (from 2013 to 2102). A spell is defined as a period of contiguous days of flow below the 10th percentile threshold.

material: pertinent or relevant

model node: a point in the landscape where hydrological changes (and their uncertainty) are assessed. Hydrological changes at points other than model nodes are obtained by interpolation.

Monte Carlo simulation: a simulation technique involving random sampling of each probability distribution within the model to produce large number of plausible scenarios. Each probability distribution is sampled in a manner that reproduces the distribution's shape. The distribution of the values calculated for the model outcome therefore reflects the probability of the values that could occur.

percentile: a specific type of quantile where the range of a distribution or set of runs is divided into 100 contiguous intervals, each with probability 0.01. An individual percentile may be used to indicate the value below which a given percentage or proportion of observations in a group of observations fall. For example, the 95th percentile is the value below which 95% of the observations may be found.

probability distribution: the probability distribution of a random variable specifies the chance that the variable takes a value in any subset of the real numbers. It allows statements such as 'There is a probability of x that the variable is between a and b'.

quantile: a set of values of a variate that divide the range of a probability distribution into contiguous intervals with equal probabilities (e.g. 20 intervals with probability 0.05, or 100 intervals with probability 0.01). Within bioregional assessments, probability distributions are approximated using a number of runs or realisations.

receptor: a point in the landscape where water-related impacts on assets are assessed

receptor impact model: a function that translates hydrological changes into the distribution or range of potential ecosystem outcomes that may arise from those changes. Within bioregional assessments, hydrological changes are described by hydrological response variables, ecosystem outcomes are described by receptor impact variables, and a receptor impact model determines the relationship between a particular receptor impact variable and one or more hydrological response variables. Receptor impact models are relevant to specific landscape classes, and play a crucial role in quantifying potential impacts for ecological water-dependent assets that are within the landscape class. In the broader scientific literature receptor impact models are often known as ‘ecological response functions’.

receptor impact variable: a characteristic of the system that, according to the conceptual modelling, potentially changes due to changes in hydrological response variables (for example, condition of the breeding habitat for a given species, or biomass of river red gums)

riparian: An area or zone within or along the banks of a stream or adjacent to a watercourse or wetland; relating to a riverbank and its environment, particularly to the vegetation.

risk: the effect of uncertainty on objectives

source dataset: a pre-existing dataset sourced from outside the Bioregional Assessment Programme (including from Programme partner organisations) or a dataset created by the Programme based on analyses conducted by the Programme for use in the bioregional assessments (BAs) stressor: chemical or biological agent, environmental condition or external stimulus that might contribute to an impact mode

subregion: an identified area wholly contained within a bioregion that enables convenient presentation of outputs of a bioregional assessment (BA)

surface water: water that flows over land and in watercourses or artificial channels and can be captured, stored and supplemented from dams and reservoirs

surface water zone of potential hydrological change: outside this extent, changes in surface water hydrological response variables due to additional coal resource development (and hence potential impacts) are very unlikely (less than 5% chance). The area contains those river reaches where a change in any one of nine surface water hydrological response variables exceeds the specified thresholds. For the four flux-based hydrological response variables (annual flow (AF), daily flow rate at the 99th percentile (P99), interquartile range (IQR) and daily flow rate at the 1st percentile (P01)), the threshold is a 5% chance of a 1% change in the variable. That is, if 5% or more of model runs show a maximum change in results under coal resource development pathway (CRDP) of 1% relative to baseline. For four of the frequency-based hydrological response variables (high-flow days (FD), low-flow days (LFD), length of longest low-flow spell (LLFS) and zero-flow days (ZFD)), the threshold is a 5% chance of a change of 3 days per year. For the final frequency-based hydrological response variable (low-flow spells (LFS)), the threshold is a 5% chance of a change of 2 spells per year.

*tmax*: year of maximum change

tmaxRef: the year that the maximum difference in drawdown relative to the reference period (1983 to 2012) (dmaxRef) occurs

uncertainty: the state, even partial, of deficiency of information related to understanding or knowledge of an event, its consequence, or likelihood. For the purposes of bioregional assessments, uncertainty includes: the variation caused by natural fluctuations or heterogeneity; the incomplete knowledge or understanding of the system under consideration; and the simplification or abstraction of the system in the conceptual and numerical models.

*very unlikely*: less than 5% chance

water-dependent asset: an asset potentially impacted, either positively or negatively, by changes in the groundwater and/or surface water regime due to coal resource development

water use: the volume of water diverted from a stream, extracted from groundwater, or transferred to another area for use. It is not representative of 'on-farm' or 'town' use; rather it represents the volume taken from the environment.

well: typically a narrow diameter hole drilled into the earth for the purposes of exploring, evaluating or recovering various natural resources, such as hydrocarbons (oil and gas) or water. As part of the drilling and construction process the well can be encased by materials such as steel and cement, or it may be uncased. Wells are sometimes known as a ‘wellbore’.

zero-flow days (ZFD): the number of zero-flow days per year. This is typically reported as the maximum change due to additional coal resource development over the 90-year period (from 2013 to 2102).

zero-flow days (averaged over 30 years) (ZQD): the number of zero-flow days per year, averaged over a 30-year period. This is typically reported as the maximum change due to additional coal resource development.

zone of potential hydrological change: outside this extent, hydrological changes (and hence potential impacts) are very unlikely (less than 5% chance). Each bioregional assessment defines the zone of potential hydrological change using probabilities of exceeding thresholds for relevant hydrological response variables. The zone of potential hydrological change is the union of the groundwater zone of potential hydrological change (the area with a greater than 5% chance of exceeding 0.2 m of drawdown due to additional coal resource development in the relevant aquifers) and the surface water zone of potential hydrological change (the area with a greater than 5% chance of exceeding changes in relevant surface water hydrological response variables due to additional coal resource development).

www.bioregionalassessments.gov.au

Australian Government
Department of the Environment and Energy
Bureau of Meteorology
Geoscience Australia

CSIRO