CONSULTATION DRAFT - NOT FOR OFFICIAL USE

Public consultation - Information Guidelines Explanatory Note: Uncertainty Analysis for Groundwater Models

The Independent Expert Scientific Committee on Coal Seam Gas and Large Coal Mining Development (IESC) is seeking comment on the updated *Information Guidelines Explanatory Note: Uncertainty Analysis for Groundwater Models*.

The IESC welcomes feedback on the content, usability and applicability. In particular, views are sought on:

- **the content of the updated Information Guidelines Explanatory Note: Uncertainty Analysis for Groundwater Models, particularly any areas where further explanation would be useful;**
- **the relevance to your specific area of work; and**
- **potential options to increase uptake and adoption.**

These updated guidelines are targeted more towards state and Commonwealth government regulators than the previous version.

The IESC and the Information Guidelines Explanatory Note Uncertainty Analysis for Groundwater Models

The IESC is a statutory body under the *Environment Protection and Biodiversity Conservation Act 1999 (Cth).* One of the IESC's key legislative functions is to provide independent scientific advice to the Australian Government Environment Minister and relevant state ministers in relation to coal seam gas and large coal mining development proposals that are likely to have a significant impact on water resources.

For some topics, Explanatory Notes have been written to supplement the Guidelines, giving more detailed guidance to help the coal seam gas and large coal mining industries prepare environmental impact assessments. These topics are chosen based on the IESC's experience of providing over 150 pieces of advice on development proposals. Case studies and practical examples of how to collect and present relevant information are also included.

The Information Guidelines Explanatory Note Uncertainty Analysis – Guidance for groundwater modelling within a risk management framework were first published in January 2018. The Information Guidelines Explanatory Note Uncertainty Analysis for Groundwater Models is currently being reviewed and amended to update reference material, cover developments in leading practice and knowledge, take account of the IESC's recent experience and incorporate comments from users.

The IESC Information Guidelines Explanatory Note: Uncertainty Analysis for Groundwater Models outline the information about the use of uncertainty analysis for regulators making decisions, using simple and intuitive concepts which expands on the first iteration of the Uncertainty Analysis – Guidance for groundwater modelling within a risk management framework.

Independent Expert Scientific Committee on Coal Seam Gas and Large Coal Mining Development

Information Guideline Explanatory Note

Uncertainty Analysis for Groundwater Models

This report should be attributed as 'Peeters LJM and Middlemis H 2022. *Information Guidelines Explanatory Note; Uncertainty Analysis for Groundwater Models*, A report prepared for the Independent Expert Scientific Committee on Coal Seam Gas and Large Coal Mining Development through the Department of Climate Change, Energy, the Environment and Water, Commonwealth of Australia, 2022'.

This publication is funded by the Australian Government Department of Climate Change, Energy, the Environment and Water. The views and opinions expressed in this publication are those of the authors and do not necessarily reflect those of the Australian Government or the Minister for the Environment.

Contact details

For information about this report or about the work of the IESC, please contact:

IESC Secretariat

Office of Water Science
Department of Climate Change, Energy, the Environment and Water
GPO Box 3090
CANBERRA ACT 2601

The report can be accessed at http://www.iesc.environment.gov.au/

Images

Information Guidelines Explanatory Note

Uncertainty Analysis for Groundwater Models

Executive Summary

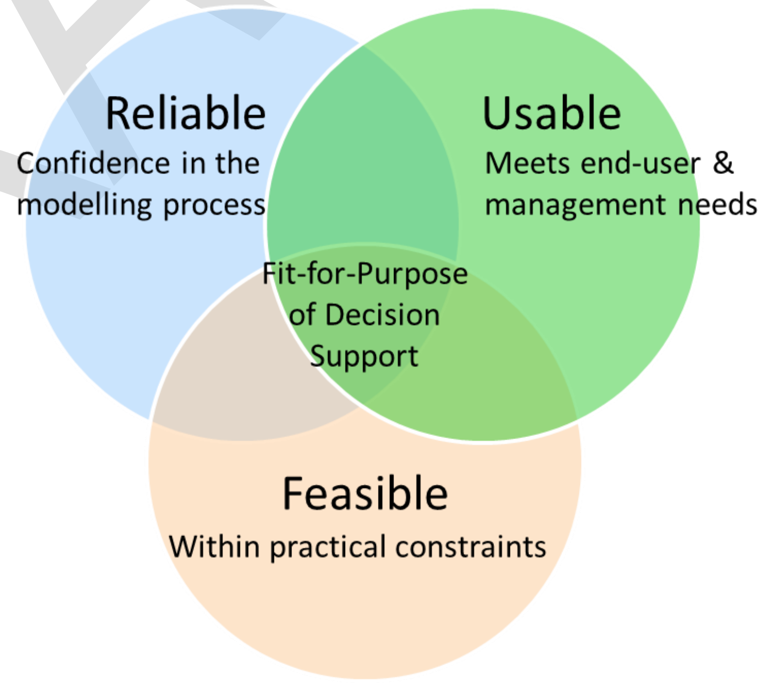
The IESC Information Guidelines are essential prior reading for this Explanatory Note on uncertainty analysis for groundwater, which aims to provide high-level guidance for non-specialists such as proponents and regulators who commission and/or review groundwater modelling studies. It focuses on key aspects to make numerical modelling and uncertainty analysis relevant to decision makers, using simple and intuitive concepts. Technical background is provided in break out boxes. It is also suitable for use by proponents/regulators on other resources projects. This IESC Explanatory Note on uncertainty analysis:

- augments the best practice Australian Groundwater Modelling Guidelines ('AGMG'; Barnett et al. 2012) that is currently being updated, and
- complements the detailed uncertainty methodologies that are championed by the Groundwater Modelling Decision Support Initiative (GMDSI.org) and are aimed more at technical specialists.

This Explanatory Note recognises that each project is unique in its hydrogeology, the hazards associated with development and the constraints on the modelling. The strategic guidance provided in this document is therefore designed not to be prescriptive. It rather aims at ensuring that model reports provide the information that readers, reviewers and decision-makers need to decide if the groundwater model and uncertainty analysis are fit-for-purpose.

Fit-for-purpose means here that the results of the model are:

- **Usable**
  relevant to the decision-making process, providing information about the uncertainty in conceptualisations and modelling simulations in a way that allows decision-makers to understand the effects of uncertainty on project objectives and the effects of potential bias.
- **Reliable**
  demonstrate that the range of model outcomes is consistent with the system knowledge and honours historical observations, providing objective evidence that uncertainties affecting decision-critical predictions of impacts on aquifer resources and dependent systems are not underestimated.
- **Feasible**
  trade-offs due to budget, time and technical constraints are reasonable and justifiable within the risk context of the project.



after Hamilton et al. (2022)

This Explanatory Note does not advocate for a particular approach of groundwater modelling and uncertainty analysis. It rather illustrates how different approaches can be justified, depending on the project and risk context, and points to advantages and drawbacks of different approaches. It also highlights particular aspects of the uncertainty analysis that require careful design, execution and review. These can be summarised as:

- What is the quantity of interest (key prediction) to the decision maker?
- What are the main sources of uncertainty to the quantity of interest?
- How do system knowledge and historical observations constrain or condition the quantity of interest (the key prediction(s) for informing decisions)?

This edition of the Explanatory Note is still underpinned by the same principles that underpinned the 2018 Explanatory Note (Middlemis and Peeters, 2018); the need to account for and report on predictive uncertainty in an open and transparent way. Where the emphasis in the 2018 Explanatory Note was on encouraging uptake of uncertainty analysis as an essential part of groundwater modelling practice, this Explanatory Note aims to make groundwater model outcomes more relevant to decision makers, provide a way to establish confidence in the results and inspire creativity and innovation in quantifying and reducing uncertainty in groundwater model results. To this end, workflow figures, classifications and checklists that can be perceived as prescriptive have been removed. Instead, the guidance is presented as open-ended questions, like those listed above, to be addressed in groundwater model reports.

Table of Contents

1 Introduction

1.1 Responsible Modelling

Saltelli et al. (2020) uses the context of the Covid-19 pandemic and epidemiological modelling to propose a 5-step manifesto of principles for responsible mathematical modelling:

- **Mind the assumptions** - uncertainty quantification and sensitivity analysis are complementary approaches to measuring the robustness and usefulness of model predictions.
- **Mind the hubris** - models are simplified representations of real systems or processes and should be developed with an optimum trade-off between complexity and error.
- **Mind the framing** - framing refers to the different lenses, worldviews, or underlying assumptions that guide how individuals or groups perceive a particular issue; transparent framing can support effective results communication and enhance trust with stakeholders.
- **Mind the consequences** – well-executed mathematical modelling helps society make wise decisions; but when not done well, models can lead to wrong or simply unjustified choices.
- **Mind the unknowns** - failure to acknowledge and communicate uncertainties can artificially limit policy options and open the door to unintended consequences.

A key point is worth highlighting:

'*Modellers must not be permitted to project more certainty than their models deserve; and politicians must not be allowed to offload accountability to models of their choosing.*'

These principles were adopted and published by UNESCO (Saltelli 2022), and this Explanatory Note on uncertainty analysis for groundwater is imbued with these principles. This is apt for a high-level Explanatory Note and many of the same ideas underpin other texts on implementing uncertainty analysis to inform decision making, such as Doherty (2022).

1.2 Context

This Explanatory Note was prepared for the Independent Expert Scientific Committee on Coal Seam Gas and Large Coal Mining Development (IESC). The IESC is a statutory body under the Environment Protection and Biodiversity Conservation Act 1999 (Cth) which provides robust scientific advice to government regulators on the potential water-related impacts of such development proposals, but it does not make regulatory decisions. The IESC Information Guidelines (2022) detail the statutory context and the information that the environmental assessment for a project proposal should include and are essential prior reading for this Explanatory Note.

Explanatory Notes supplement the IESC Information Guidelines with guidance on specialist topics, such as:

- Uncertainty analysis for groundwater models (this report)
- Assessing groundwater-dependent ecosystems (Doody et al. 2019; crucial input to groundwater model design in terms of causal pathways for impacts)

- Deriving site-specific guideline values for physico-chemical parameters and toxicants (Huynh and Hobbs, 2019)
- Ecohydrological Conceptual Models (draft in prep. 2022)
- Characterisation and modelling of geological fault zones (Murray and Power, 2021).

The linkages between the Information Guidelines, Explanatory Notes and other IESC reports, and the complementary best practice resources are shown in Figure 1.

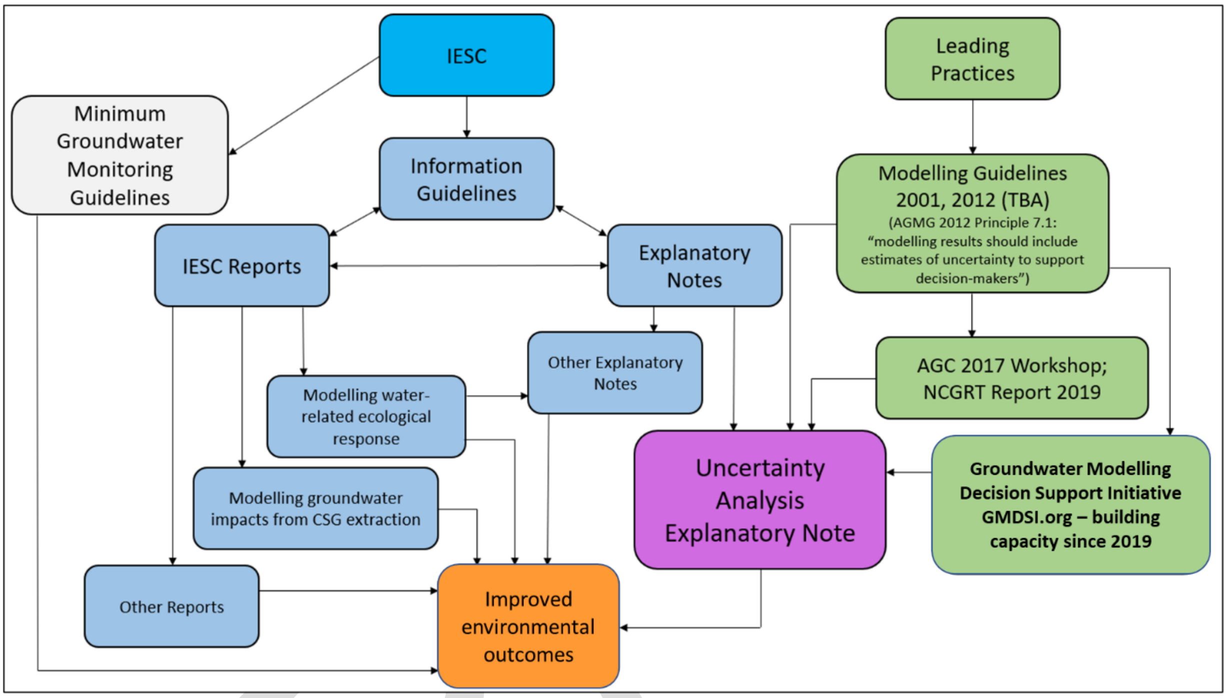


Figure 1 Contextual diagram for Explanatory Note on uncertainty analysis

As shown in Figure 1, this Explanatory Note on uncertainty analysis for groundwater models augments the best practice Australian Groundwater Modelling Guidelines ('AGMG'; Barnett et al. 2012) and complements the detailed methodologies that are championed by the Groundwater Modelling Decision Support Initiative (GMDSI.org) which is aimed more at technical specialists. GMDSI resources provide more detailed information on the practical application of uncertainty analysis methods, and the necessary focus on hypothesis testing, such as Doherty (2022), which is beyond the scope of this report. This Explanatory Note has a high-level focus and is aimed more at non-specialists such as proponents and regulators who commission and/or review modelling studies. The main focus in application for this Explanatory Note is on coal mining and coal seam gas development, but the material is also relevant in the context of other resources development projects.

1.3 Information Guidelines

The Information Guidelines are essential prior reading for this Explanatory Note. However, to help set the context for this Explanatory Note, while trying to minimize repetition, the points below summarize some key Information Guidelines requirements that are foundational to this Explanatory Note.

The development proposal should present sufficient evidence for independent verification of:

- the processes of cause and effect between the project and water resources; and

- the materiality and likelihood of the potential impacts and risks to water resources.

Early in the investigation process, preliminary assessments should consider all available data to:

- identify and characterize water resources and water-dependent ecosystems in the region
- develop the initial hydrogeological and ecohydrological conceptual models (Ecohydrological Conceptual Models Explanatory Note in prep. 2022)
- identify potential stressors and causal pathways for impacts on water-dependent ecosystems (see Doody et al. 2019 on groundwater-dependent ecosystems)
- conduct a preliminary risk assessment, with consideration of treatments and/or mitigations, and develop preliminary analytical/numerical models if warranted for the risk context
- engage with relevant regulators and stakeholders on the initial findings and next steps.

It is noteworthy that the above steps are consistent with the NSW Gateway process requirements[1].

Conceptual models are very useful at any stage of environmental risk assessment, but especially at the problem definition stage, to understand and communicate the complex interplay of stressors, sources, causal pathways for impacts and possible multiple cause-effect pathways (see Doody et al. 2019). An ecohydrological conceptual model (ECM Explanatory Note in prep. 2022) integrates the hydrological components (surface and groundwater) with the ecological components. An ECM illustrates the likely causal pathways for impacts on key aspects of water resources (e.g. flow/level regime, water quality, biota, ecological function), and the likely relative importance of these pathways during operations and post-project.

During the comprehensive environmental impact assessment process, evidence-based estimates are required of the range of potentially significant impacts to water resources and water-dependent ecosystems. Significant impact is defined in the Glossary, based on the MNES Significant Impact Guidelines (CoA 2013). The assessment should involve an iterative process using best practice source-pathway-receptor methods (Morrison-Saunders, 2018). Most large developments will likely require numerical modelling unless a low-risk context and simpler methods can be justified. Progressive results should be used to refine the models (conceptual, ecohydrological, numerical/analytical) to decrease uncertainties and to inform plans for mitigation, management and monitoring.

For each causal pathway for significant impact (Peeters et al. 2022), the intensity, duration, timing, magnitude and geographic extent of the potential impacts to the water resource should be clearly described, along with the resultant impact to any water-dependent ecosystems and the consequence or significance of the impact. Any modelling should be conducted at spatial and temporal scales suitable to represent the physical, chemical and ecological processes associated with the water resources or water-dependent ecosystems. The assessment should estimate the range of impacts at all phases (construction, operation and post-closure) and the potential cumulative impact of all past, present and reasonably foreseeable actions and significant water-affecting activities In the area.

The level of resources (data, time, budget, methods) applied to the assessment should be commensurate with the level of risk, by considering the probability and potential consequences of significant impacts/risks, and the value, condition and the vulnerability and/or sensitivity of the ecosystems affected. As the assessment progresses, effort should focus on the ecosystems at greatest risk.

The combination of conceptual, hydrogeological and ecohydrological models, coupled with risk analysis, may be sufficient to adequately estimate low-impact outcomes with low levels of uncertainty such that analytical and/or

[1] https://www.planning.nsw.gov.au/Policy-and-Legislation/Mining-and-Resources/Gateway-Assessment-and-Site-Verification

numerical models may not be required or that modelling may only be needed to test a limited number of causal pathways.

However, when analytical and/or numerical models are required, they should be designed for the purpose of supporting decision making (Doherty 2022; Doherty and Moore 2019, 2021), to:

- quantify the uncertainties of predictions of management interest (the 'quantities of interest') so that risks can be estimated for proposed management actions;
- justify the model simplifications and abstractions involved in assessing these uncertainties;
- reduce predictive uncertainties through assimilation of data; and
- ensure that parameter and structural variability is represented in the range of uncertainty analysis outputs so that the uncertainties of decision-critical model predictions are not underestimated and management risks are not understated.

In that sense, a 'model' should not be thought of as simply a noun or a deliverable, but more as a verb for a risk-based process of investigating the complex interplay of stressors, sources and causal pathways for impacts on receptors and the uncertainties involved, in order to provide information to support decision-making and environmental management. The results from any modelling should show the range and likelihood of possible outcomes, based on sensitivity and uncertainty analysis predictions that are sufficiently robust to support risk management and regulatory decision-making.

Sufficient information should be provided in the environmental assessment documentation to allow an independent reviewer (IESC or others) to evaluate the justifications for the assertions made. For example, on the basis for the conceptual or ecohydrological models, the underlying assumptions for the numerical models and why an ecosystem is or is not groundwater-dependent. An independent reader of the environmental assessment documents should be able to verify all significant assumptions, methodologies, techniques, assertions and conclusions made by the proponent. To that end, <u>assessment reports should be as self-contained as possible</u>, minimising the need for the reader to consult several volumes of technical reports to understand the methodologies and assumptions applied and the predictions of impacts on groundwater and surface water systems and dependent ecosystems (but allowing for the flow of information between disciplines conducting the assessments). For example, there are usually linkages between the groundwater and surface water assessments, such as to estimate recharge/runoff and to evaluate the site water balance (usually presented in the surface water report). These site water balance assessments typically use outputs from the groundwater model as key inputs, and produce outputs that can form inputs to post-mining groundwater model scenarios of a final void lake (Figure 2). All of these assessments should involve uncertainty assessment methods. See also section 4.1.

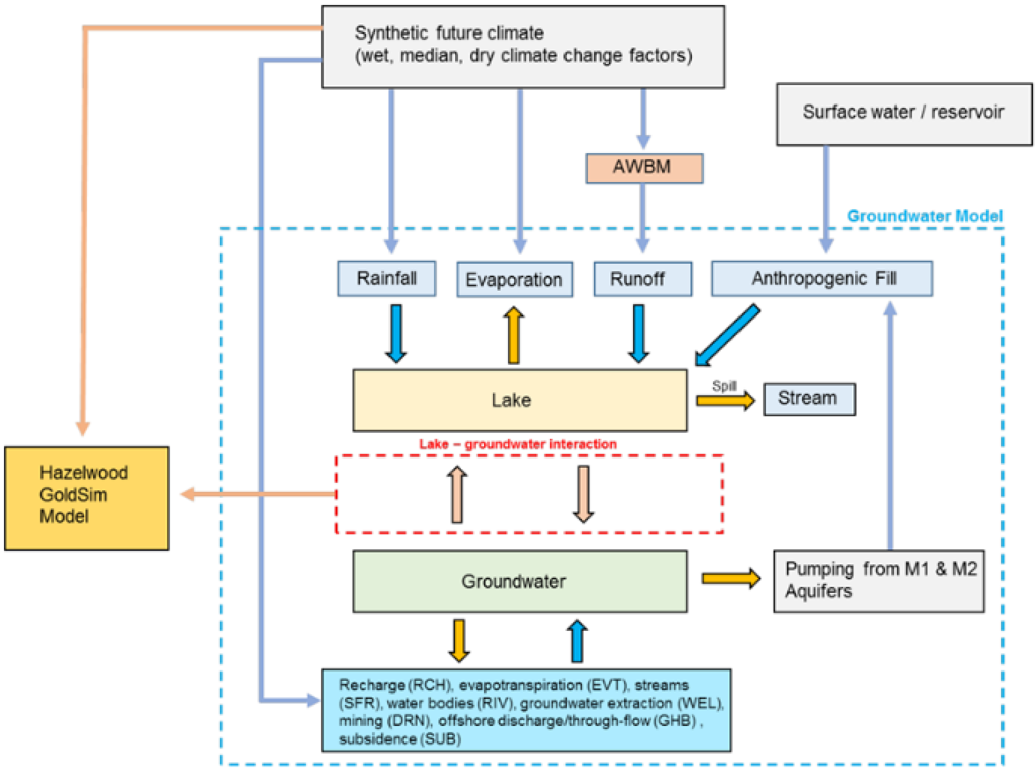


Figure 2 – Example of interactions between groundwater and surface water models (after Gresswell et al. 2019)

1.4 Purpose of this Explanatory Note

This Explanatory Note complements the Information Guidelines summarised above, with a focus on the role of uncertainty analysis in risk-based decision-making. Explanatory Notes provide guidance, rather than mandatory requirements, to assist the preparation and review of environmental impact assessments. The IESC recognizes that approaches, methods, tools and software will continue to develop, and reviews and updates the guidance accordingly. This Explanatory Note on Uncertainty Analysis for Groundwater Models is an update to the first version published in 2018 (Middlemis and Peeters, 2018) and is again designed to complement best practice initiatives (Figure 1).

The aim of this Explanatory Note is to provide a high-level introduction and strategic overview of uncertainty analysis relating to groundwater modelling for environmental impact assessment, management and decision making. Rather than retaining the prescriptive 'fatal flaws' checklist and groundwater model confidence classification, this edition introduces simple and intuitive concepts and diagrams to facilitate discussions on whether a groundwater model and uncertainty analysis is fit for purpose.

As summarized in Table 1, this document is not a formal guideline or an instruction manual, and it does not provide a step-by-step guide to conducting uncertainty analysis (see break-out box on modelling guidelines). This would not be feasible, given the many approaches and methodologies that are available and continue to be developed, for instance through the GMDSI initiative (GMDSI.org). As will be discussed in section 3, the appropriate approach will depend on the characteristics of the proposed development and its risk profile and should be designed in consultation with relevant regulators.

Table 1 Aims and exclusions of this Explanatory Note on uncertainty analysis for groundwater

Explanatory Note is intended/designed to…	Explanatory Note exclusions
Complement the IESC Information Guidelines by providing a high-level strategic overview of uncertainty analysis for groundwater modelling.	Not a textbook, instruction manual or a formal guidelines document.
Complement the Australian Groundwater Modelling Guideline, which requires uncertainty analysis (ch. 7), and the GMDSI initiative gmdsi.org	Not a step-by-step guide to uncertainty analysis (that is not workable, given the diversity of methods, projects and risks).
Integrate uncertainty analysis within a risk management framework, to a level of detail that is commensurate with the potential risks and/or consequences. See also point opposite.	Does not identify a single preferred method of uncertainty analysis, but requires: • qualitative UA as a minimum, and • quantitative UA for high risks/consequences.
Outline a workflow process of uncertainty analysis for groundwater modelling to support environmental impact assessment, management and decision making, to assist proponents and regulators commission and review studies.	Does not preclude use of any techniques, provided satisfactory justifications are provided and the effects of bias are explored in clear reports amenable to review.

Break-out box 1: Modelling guidelines

2 Risk Management and Decision Support

2.1 Risk Management and Decision Support

Freeze et al. (1990) characterize the role of models in decision support as quantifying the level of risk associated with management options. This succinct principle remains valid today. It follows that if a model is applied to support environmental decision-making that seeks to avoid unwanted outcomes (so-called 'bad things'), its simulations of the consequences of management options must quantify the related uncertainties (Doherty 2022; Doherty and Moore 2019, 2021). Quantifying uncertainties conveys the confidence in the model results.

Uncertainty is integral to risk management, as outlined in the fundamental principles of the risk management standard AS/NZS ISO 31000:2009:

- risk is defined as the effect of uncertainty on project objectives
- risk is characterized as a function of the likelihood and consequence of an outcome.

It follows that risk cannot be assessed without an understanding of uncertainty, and an uncertainty analysis should be conducted within a risk management context:

- where the model predictions of the consequences (impacts) of developments or management options should be quantified with related uncertainties (likelihoods),
- to inform decisions that seek to avoid the risk of unwanted impacts, and
- to inform assessments of risk treatment options, mitigation measures and management.

2.2 Causal Pathways and Ecohydrological Conceptual Models

A groundwater impact and uncertainty assessment that is integrated within a risk management framework begins with an identification of potential causal pathways for impacts under an initial conceptualization and investigates possible risk mitigations. A causal pathway is the logical chain of events, either planned or unplanned, that link the planned resource development activities and potential impacts on water resources and water-dependent ecosystems.

Best practice source-pathway-receptor impact assessment methods (Morrison-Saunders, 2018), such as the risk-based mining project methods of Howe et al. (2010) that were curated by the National Water Commission, are consistent with the risk-based causal impact pathway aspects of the best practice groundwater modelling guidelines (Barnett et al. 2012), the IESC Information Guidelines and this Explanatory Note. A common principle is that if an exposure pathway between mine-water affecting activities (a 'source') and a receptor is shown to not exist, the impact assessment 'chain' breaks, rendering that particular risk redundant (Howe et al. 2010). Peeters et al. (2022) developed a methodology to combine and evaluate the various causal pathways in an environmental impact assessment into a logically consistent causal network. Examples of such causal network for assessing potential impacts on water and the environment of unconventional gas resource development can be found at https://gba-explorer.bioregionalassessments.gov.au/

Identifying causal pathways is an important part of uncertainty analysis. Which causal pathways require investigation will determine the modelling approach, the sources of uncertainty to consider and, most importantly, the model outcomes required. Causal pathways should be identified by conservatively considering potential connectivities between groundwater units and/or surface water features and related water-dependent ecosystems. An Explanatory Note on Ecohydrological Conceptual Models, including causal pathways, is in preparation (2022).

All projects require at a very minimum a qualitative uncertainty analysis, discussing the potential causal impact pathways and the perceived effect of uncertainties on the impact assessment outcomes. A qualitative uncertainty analysis should use conceptual models and qualitative or semi-quantitative estimates of the likelihood of risks to and impacts on environmental receptors and related values, along with the level of confidence of scientific advice on these risks and impacts. Minimising and acknowledging bias in such investigations of causal pathways is a key element of the ecological values analysis at the problem definition stage, along with data analysis, conceptualisation, and the initial risk analysis and treatment options assessment.

Where resource development projects are classified as posing high environmental risk, a quantitative uncertainty assessment is also warranted, using groundwater models that are designed to:

- investigate the causal pathways for potential impacts on water resources and water-dependent ecosystems
- quantify the likelihood of the impact and assess the consequence (magnitude, extent, severity)
- enable the investigation of effective risk treatment and avoidance strategies or mitigation measures as an integral part of a risk management framework.

The next section, section 3, focuses on this quantitative uncertainty analysis. As the modelling and assessment workflow proceeds through its iterations, the objectives should be reviewed according to risk, and complexity may be added or refined as necessary. In the preliminary stages, there may not be any need for numerical modelling. If risks are not high at any stage, nothing more may be required, and the investigation may be cut short. However, many resource projects are likely to pose high environmental risks. This means that the proponent should conduct a quantitative uncertainty assessment to a level of detail commensurate with the potential risks and consequences of the project. During the risk and uncertainty assessment process, socially and economically acceptable and effective risk treatments or mitigations may be identified that can be part of adaptive management strategies.

Bioregional Assessments (and Geological and Bioregional Assessments) provide useful regional-scale case studies, with an emphasis on causal pathways, for environmental impact assessments of large coal mines and CSG proposals. They also illustrate how impact assessments can address principles from the Information Guidelines requiring consideration of:

- causal pathways linking depressurisation and dewatering of coal seams at depth with impacts on anthropogenic and ecological values of water-dependent receptors
- conceptual models and quantitative, semi-quantitative or qualitative analyses for estimating the likelihood of risks to and impacts on receptors and related values, along with the level of confidence of scientific advice on these risks and impacts
- potential direct, indirect and cumulative impacts on water resources
- monitoring, evaluation and review programs, and related risk assessment and treatment studies, to minimise or mitigate impacts on water resources.

However, the bioregional assessments approach should not be considered a template for an environmental impact assessment, as the objectives, scope and scale are quite different. Bioregional assessments provide advice on development stressors, causal pathways, receptors and ecosystems but they are not development-specific. Bioregional assessments do, however, inform environmental impact assessment studies by providing regional context information and, importantly, independent cumulative impacts assessment.

2.3 Adaptive Management

Adaptive management is often justifiably used to address environmental issues in the face of uncertainty. However, the lengthy time lags affecting groundwater processes can mean that it may be difficult to effectively reverse the impacts of an action once initiated (Walker 2017). By the time monitoring shows that a valued ecosystem will be affected, it may be too late to prevent impacts occurring. For example, groundwater drawdown could continue to increase due to the hydrogeological time lag effects despite groundwater extraction ceasing.

Tolerance of an unwanted outcome ('failure') is related to the cost of failure. If the cost is relatively low, then a moderate likelihood of failure may be tolerated, provided there are economically and socially acceptable risk-reduction options that can be implemented in a timely fashion. On the other hand, if the cost of failure is high (e.g. unwanted impacts on high-value ecosystems), the likelihood of failure must be low for a management option or adaptive management plan to be deemed socially and economically acceptable, and effective.

This drives the need for a conservative approach to impact assessment. Such an approach includes careful analysis of uncertainties and investigation of options for risk treatments and mitigation. It is also important to communicate the residual risk, i.e. the risk remaining after taking into account risk treatments and mitigation, and be able to adaptively manage it.

However, even the most comprehensive modelling and uncertainty analysis study cannot completely rule out the potential for unwanted outcomes. Decisions on developments are often accompanied by a set of conditions. These can be aimed at reducing uncertainty by addressing knowledge gaps, for instance through additional field investigations. They can also stipulate requirements to monitor the responses to the development of groundwater, surface water and ecology. Uncertainty analysis and sensitivity analysis, however, can greatly assist in setting such conditions. Sensitivity analysis, as a complementary part of an uncertainty analysis, identifies which parameters contribute most to predictive uncertainty and it also identifies which parameters are constrained by the currently available observations. These insights can help in identifying the most important knowledge gaps to be addressed and in designing monitoring strategies. Modelling methods can also be applied in a formal data worth analysis (i.e. expanding on the above) to identify where as-yet-ungathered data may most effectively reduce the uncertainties of decision-critical predictions, which can help objectively prioritise monitoring strategies.

3 Designing Uncertainty Analysis

3.1 Fit for purpose

The goal of an uncertainty analysis in general is to provide a range of model predictions that are consistent with the knowledge of the system and with observations relevant to the prediction. Within the context of risk-based decision-making, the purpose of uncertainty analysis becomes more specific as it needs to provide:

- objective evidence that the uncertainties affecting decision-critical predictions of impacts on aquifer resources and dependent systems are not underestimated;
- information about the uncertainty in conceptualisations and model simulation outputs in a way that allows decision-makers to understand the effects of uncertainty on project objectives (echoing the AS/NZS ISO 31000:2009 risk definition) and the effects of potential bias.

A groundwater modelling investigation and associated uncertainty analysis is fit-for-purpose if it achieves the above. It emphasizes the need to clearly define the quantity of interest, the model output or key prediction that is relevant to decision makers, and to develop the groundwater model and execute the uncertainty analysis accordingly.

Hamilton et al. (2022) discuss whether a model is fit-for-purpose in terms of a trade-off between usability, reliability and feasibility (Figure 3). Usability relates to the intended use of the model results, how the model will be used to inform decision making. Reliability is about the confidence in the model results and the trust in the modelling process, both of which will be informed by the uncertainty quantification, assessing the likelihood and consequences of potential impacts. Feasibility pertains to the pragmatic aspects of delivering the modelling project within the time and budget constraints.

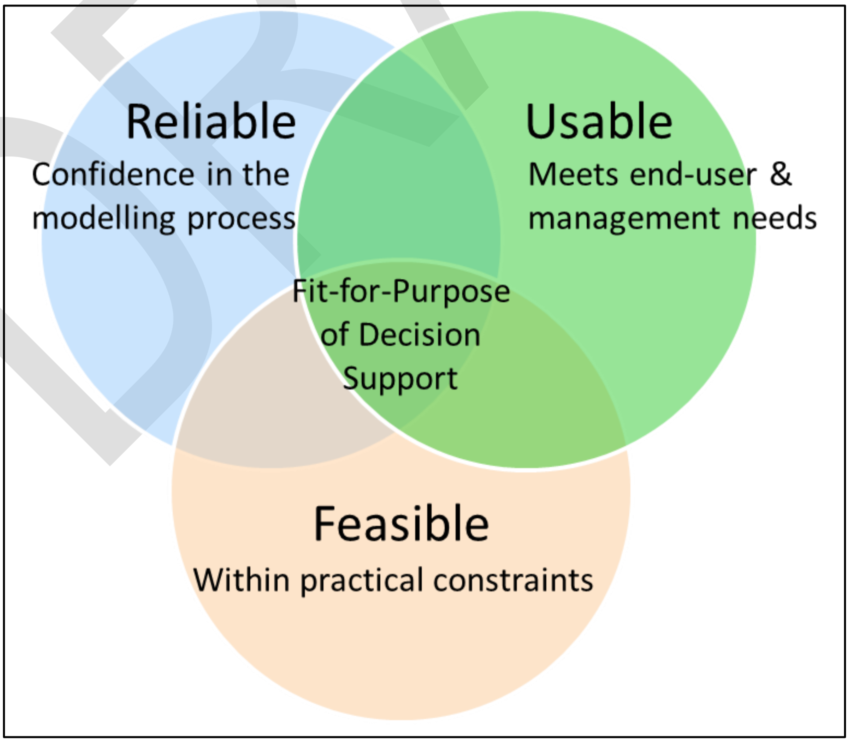


Figure 3 Fit for purpose is a trade-off between usability, reliability and feasibility (after Hamilton et al. 2022)

Uncertainty analysis primarily addresses the reliability dimension by providing a likelihood of an event or a prediction interval based on the range of predictions, where this range of predictions is constrained or conditioned by the available knowledge of the system and observations. An open and transparent representation of the range of model outputs helps in building trust and confidence in the model approach and results, especially if it complements an honest reporting of the state of knowledge on the aquifer system and how this is captured in the groundwater model. The extent to which uncertainty can be quantified, however, cannot be separated from the other dimensions of fitness for purpose. The usability refers to the context in which risk-based decision making occurs and the feasibility refers to the trade-offs required for the study resources available (data, time, budget).

A fit-for-purpose assessment of a groundwater model and uncertainty analysis will be highly context-specific and difficult to capture in formalised checklists or classifications (see break-out box 2: Australian Groundwater Modelling Guidelines Confidence Classification).

Break-out box 2: Australian Groundwater Modelling Guidelines Confidence Classification

In developing a groundwater model, from creating a conceptual model based on available data and knowledge to implementing this in a numerical model and performing an uncertainty analysis, trade-offs will be made between the different aspects of the fit-for-purpose framework (Figure 3). These assumptions and model choices will introduce uncertainty in the results that is hard to quantify, but the potential implications are essential to understand in the interpretation of the model predictions. The next sections explore further what these trade-offs can look like in groundwater modelling and uncertainty analysis. Section 4 expands on the need to transparently and openly capture these trade-offs, assumptions and model choices in the model reports.

3.2 Usability: quantity of interest (QoI)

The usability of a groundwater model is mainly focused on how the model outcomes will inform risk-based decision-making. Hazard identification identifies how the proposed development can cause an environmental impact, which requires an understanding of causal pathways. A risk assessment evaluates the likelihood and consequence of each hazard and possibly which actions can be taken to mitigate the hazard. The outcomes of a groundwater model are only usable for risk assessment if it is clear which hazards they relate to and how the outcomes will quantify likelihood and consequence. The outcomes of a groundwater model often only address a component on the hazard causal pathway. Consider the hazard of the persistence of a groundwater-dependent ecosystem being adversely affected by an open pit mine. The causal pathway for this hazard can look like this:

Open pit mine → *mine dewatering* → *drawdown at GDE* → *persistence of GDE*

A groundwater model could be designed to assess the likelihood and magnitude of mine dewatering or to assess the likelihood and magnitude of drawdown at the GDE. A groundwater model would, however, not be able to provide information on the effect of drawdown at the GDE on the health and persistence of that GDE. This requires an ecohydrological understanding of the functioning of the GDE and how it is dependent on groundwater (see Doody et al. 2022).

This simple causal pathway highlights a couple of challenges common to many groundwater modelling exercises for impact assessment. A groundwater model can be used to estimate mine dewatering rates and a groundwater model can be used to estimate drawdown caused by that dewatering. For the former, the relevant model outcome is a prediction of inflows to the mine void and/or active dewatering wells as mining progresses, while for the latter it is a difference between groundwater levels at the location of a GDE at a future time, with and without mine development. We refer to these model outcomes or key predictions as the Quantity of Interest (QoI). A quantity of interest is a model outcome from a specified model scenario, with a predefined spatial and temporal setting, that is relevant to assessing the likelihood and consequence of a causal pathway element representing a hazard.

The concept of QoI is crucial in assessing whether a groundwater model and uncertainty analysis is usable; it needs to be demonstrated that the groundwater model is able to provide reliable predictions of the QoI, within the feasibility constraints of time, budget and computational resources. All assumptions and model choices need to be evaluated within the QoI in terms of whether the choice will affect the QoI and if so, whether the potential error introduced will be acceptable.

Identifying a QoI is an activity that should involve all parties, proponent, regulator, stakeholders and modelling team. The QoI will represent the environmental, social and economic values of importance and what is considered an acceptable level of risk is a value-laden decision that needs broad support.

In the example of the causal pathway above, there are two quantities of interest; mine dewatering rate and drawdown. The model choices and assumptions that are appropriate for simulating mine dewatering are not necessarily appropriate for simulating drawdown at a distance from the mine pit. In an ideal case, a groundwater model is designed and developed for each QoI, starting from the same knowledge base. If the same groundwater model is used for multiple QoIs, it is essential that all assumptions and model choices are discussed for each QoI. Building a groundwater model for a single QoI often allows a greatly simplified conceptualisation as only the processes and structures relevant to the QoI need to be represented. In a model with multiple QoIs, it is much harder to justify simplifications that are appropriate for each QoI and such models can end up as a compromise, with sub-optimal results for all QoIs.

The GDE example also highlights another issue in specifying a QoI, in that 'drawdown at GDE' is quite vague. Research into the ecology of the GDE will be essential to figure out what is actually relevant for the persistence of the GDE. Is there an absolute groundwater level below which the vegetation can no longer access groundwater? Is the GDE most sensitive to a change in groundwater level or is it salinity? Is there a seasonal component, in which the vegetation is more sensitive to changes in groundwater level in the dry period than in the wet period? The answer to these questions will lead to subtle changes to what outcome of the groundwater model is relevant for decision making, and this would usually be facilitated by a multi-disciplinary investigation team.

The definition of the QoI includes specifying the model scenario (key prediction). The investigation focus is often on implementing the proposed water extraction for the development. Other time-varying boundary conditions, such as recharge rates or stream stages are often implicitly assumed to remain similar to their historical observed behaviour. Including multiple scenarios for time-varying boundary conditions is worthy of consideration, for instance based on climate change scenarios. This is important as some impacts may not be additive or linear. For instance, predictions of mine inflow and drawdown under a future in which a stream is variably gaining-losing but connected, can be very different to predictions of mine inflow and drawdown under a climate-change future in which that stream is losing at maximum rate (i.e. acting as a perched recharge source well above a depressed watertable). Note that such changes in future boundary conditions can be represented in a groundwater model as individual model runs (with an uncertainty analysis for each) or can be parameterised (e.g. recharge rates; dynamic stream-aquifer interactions) and included in the predictive uncertainty analysis. In that case, the parameters that represent future conditions should not be constrained by historical data (as that would limit the range to the historical range).

The next section discusses the trade-offs in establishing reliability of a groundwater model, depending on the QoI.

3.3 Reliability: range – observations – knowledge trade-off

The traditional approach of the calibrated model implicitly assumes that there is a single model that best fits the data and knowledge and that all predictions made with such a best-fit model have the same confidence. The definition of uncertainty analysis (above) on the other hand recognises that there will not be a single model or parameter combination to fit the data. More importantly, that prediction uncertainty will differ depending on the kind of prediction. Break-out box 3 (Sources of uncertainty) discusses the different sources of uncertainty relevant to groundwater models.

Break-out box 3: Sources of uncertainty

The goal of uncertainty defined in section 3.1 focuses on range, observations and knowledge as main factors that contribute to the reliability of a model. The range of predictions serves as a quantification of prediction uncertainty, while being consistent with system knowledge, and honouring observations provides confidence that the model is an appropriate and accurate simulator of the groundwater system.

3.3.1 Range of QoI

Estimating the range of the quantity of interest is often the prime objective of uncertainty analysis. Based on the spread or range of the model predictions, likelihoods can be calculated. The range of model predictions are often presented through summary statistics, such as the mean and standard deviation or the median and the range spanned between the 10^{th} and 90^{th} percentile. The mean and standard deviation are useful when the predictions are normally distributed, while median and percentiles are more useful when the predictions have a skewed distribution. As risk is a function of likelihood and consequence, it is often not enough to represent the range of predictions by summary statistics, but as likelihoods, for instance through exceedance probabilities. An exceedance probability is the probability that a value in an ensemble is exceeded (or not exceeded), such as 'the probability of more than 20 cm drawdown at this bore is 2.4 %'. Another, less quantitative, way of presenting the possible range of predictions is through a single simulation of an extreme parameter combination, often referred to as a 'worst case scenario'. While the term 'worst case scenario' is fundamentally flawed (it is impossible to demonstrate that this is indeed the worst possible outcome) and the use of the term should be avoided, the concept has some merit. A carefully selected extreme parameter combination that can be demonstrated to overestimate impact provides an upper bound to impact estimation and can be used to exclude areas of investigation with great confidence. This is explored further in the break-out box 4 on the precautionary principle and conservatism.

Break-out box 4: Precautionary principle and conservatism

In the selection of the quantity of interest, it is also important to consider how the range of this QoI is likely to be used by decision makers. Is the focus on the most likely value, with uncertainty analysis providing an estimate of the spread around that value? Or is the focus on the largest realistic impact? Or on the probability of exceeding a predefined threshold?

3.3.2 Observations

In groundwater modelling, honouring observations is traditionally referred to as calibration, the formal minimisation of the mismatch between observed and simulated state variables, such as groundwater level measurements, river fluxes or salinity. In other fields, this process is referred to a history matching, inverse modelling or parameter inference.

The mismatch between observed and simulated values is generally captured in a single value, the objective function value. Despite its name, there are a number of often subjective choices that need to be made to create the dataset of observations used in the objective function, including:

- **Which observations to include.**
  Some observations are not representative of the system under study, and it can be justifiable to not include these. An example would be groundwater levels observed in a monitoring well very close to an unmetered pumping well. Groundwater levels in this location will be affected by the pumping regime of the well, which means that history matching based on this point may bias hydraulic property estimates if the pumping rate is not included as a parameter.
- **How to weight different types of observations.**
  The importance of each observation or group of observations can be changed through adjusting their weighting. One way of weighting observations is based on the measurement uncertainty. This often proves difficult in practice, as measurement uncertainty is seldom quantified. More subjective, ad-hoc weighting schemes are therefore usually applied. These weighting schemes need to be documented and justified, especially on how they will influence the range of predictions of the QoI.
- **How to pre-process observations.**
  Pre-processing of observations can be an effective way of better exposing the information content in observations. Examples are baseflow separation of river flow measurements or temporal and spatial differencing of groundwater levels.
- **The acceptable level of mismatch between observed and simulated values.**
  In groundwater modelling, the mismatch is generally summarised in a performance metric such the normalised root mean squared error, although other metrics exist (Bennet et al. 2013).
  It is not common practice in groundwater modelling to explicitly state what is an acceptable level of mismatch between observations and simulations, apart from the often misinterpreted 'target' of 5%-10% for the normalised root mean squared error in groundwater modelling guidelines (Middlemis et al. 2001; Barnett et al. 2012). An acceptable mismatch is project- and model-specific and depends on the measurement uncertainty, the scale and resolution of the groundwater model as well as the purpose of the model.

The list above highlights the need to discuss and justify the various choices and assumptions in compiling and presenting the dataset used to constrain or condition the model.

Another important aspect of compiling the history matching dataset is to demonstrate that the observations can constrain the parameters that are relevant to the quantity of interest. Figure 4 shows an example of a global sensitivity analysis of a groundwater model (Scheidt and Caers, 2018). The analysis shows the relative importance of parameters to both the quantity of interest and to the simulated equivalents of observations. The main goal of history matching is using historical observations to constrain or condition parameters that are important to the QoI so that the uncertainty in the QoI reduces. Sensitivity analysis allows identifying whether the QoI and the simulated equivalents to observations are affected by the same parameters and that therefore the observations will be able to constrain the prediction. In the example in Figure 4, the parameters to which the QoI is sensitive are different to those to which the observations are sensitive to. Honouring observations will not greatly reduce the uncertainty in predictions.

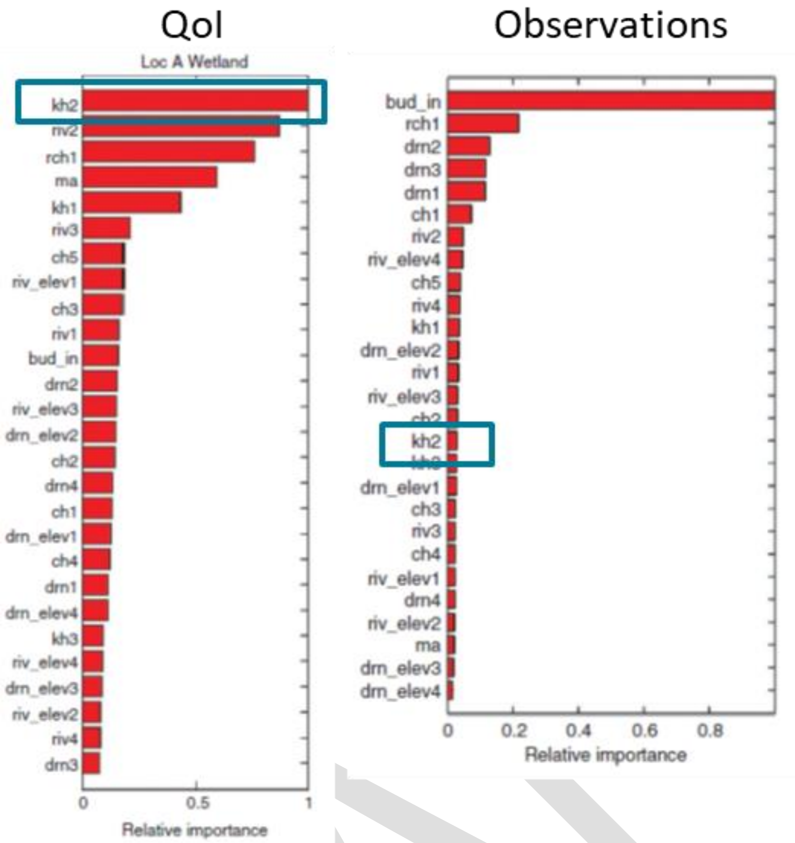


Figure 4 Sensitivity analysis (from Scheidt & Caers, 2018). The left-hand plot shows the relative importance of different parameters to the quantity of interest, drawdown under a wetland. The right-hand plot shows the relative importance of different parameters

There can, however, be an indirect effect on the QoI of constraining or conditioning parameters by observations. In the example in Figure 4, reproducing historical observations predominantly depends on the fluxes into and out of the model. Should these parameters be excluded from the parameter inference and fixed on a prior value, the inference will result in biased estimates of the hydraulic properties that do matter to the QoI.

Figure 5 shows the most typical situation in groundwater models where there is overlap between the parameter set to which the QoI is sensitive, and the parameter set to which the simulated equivalents to observations are sensitive. The figure shows that there is only part of the uncertainty in the QoI that will be reduced by honouring observations. The part of the uncertainty due to parameters to which only the QoI is sensitive will not be reduced through honouring observations. As pointed out above, parameters to which only the observations are sensitive are still relevant in parameter inference as they ensure that the parameters to which the QoI is sensitive will not be biased.

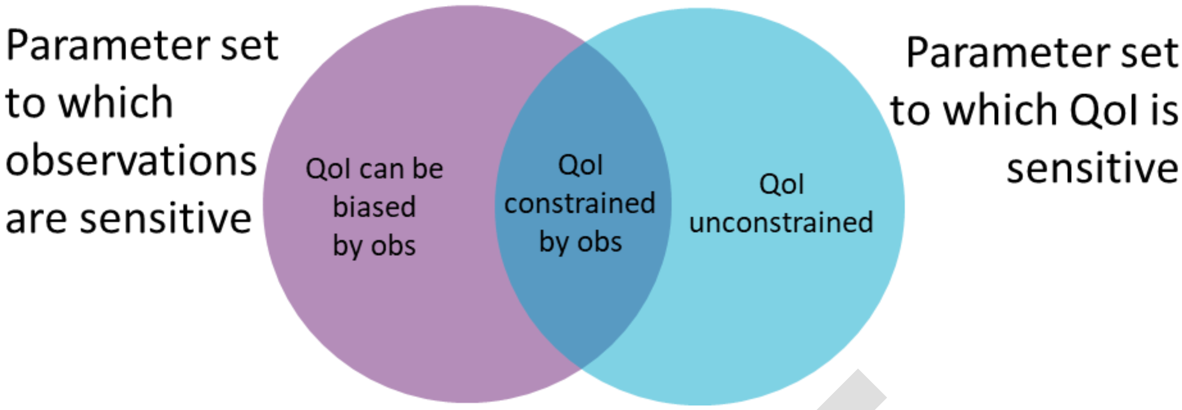


Figure 5 Only parameters to which both the observations and QoI are sensitive will reduce uncertainty in the QoI. The contribution to the range of QoI from parameters to which the observations are not sensitive will be unconstrained through history matching. Parameters to which only observations are sensitive need to be included in history matching to avoid bias in parameters to which both the QoI and observations are sensitive. Parameters to which neither the observations or QoI are sensitive should be omitted from history matching or uncertainty analysis.

For some groundwater models, there will be no overlap between both parameter sets. In that case, honouring observations will not reduce uncertainty in the QoI and will not increase confidence in the model. Parameters to which neither the observations nor the QoI are sensitive should not be included in the uncertainty analysis, except if there is a need to explicitly demonstrate that these parameters will not affect the QoI.

The section above highlights the need to present the results of a sensitivity analysis to show to which parameters the QoI is sensitive and to which parameters the observations are sensitive. It is not sufficient to only show the sensitivity analysis to observations. Break-out box 5 'Sensitivity analysis' expands on different sensitivity analysis methods.

Break-out box 5: Sensitivity analysis

3.3.3 Knowledge

The third objective of uncertainty analysis is to demonstrate that the range of outcomes is consistent with system knowledge. This recognises that observations used in history matching only capture part of the knowledge about the system. A model that matches observations does not guarantee that the model is able to adequately simulate the quantity of interest, and vice versa, a model that does not match all observations is not necessarily unsuited to simulate the quantity of interest. The reasons for this can be varied, for instance when observations are dominated by processes or local detail that is not relevant to the quantity of interest.

As the mismatch between observed and simulated values is not sufficient to evaluate whether a model is appropriate to simulate the quantity of interest, it is necessary to evaluate whether the model is consistent with the understanding of the groundwater system. A crucial aspect in this is how the conceptual model of the aquifer system is translated into a numerical model. This includes the representation of geometry and structures in the subsurface, fluxes into and out of the model and the spatial variability of parameters (Enemark et al. 2019). To assess whether the model adequately captures these aspects, the quantitative evaluation of mismatch with observed values is often complemented by a more qualitative evaluation of whether the model can reproduce temporal trends in groundwater levels, variations in space and time of recharge processes or the flow patterns around faults and discontinuities.

3.3.4 Trade-off examples

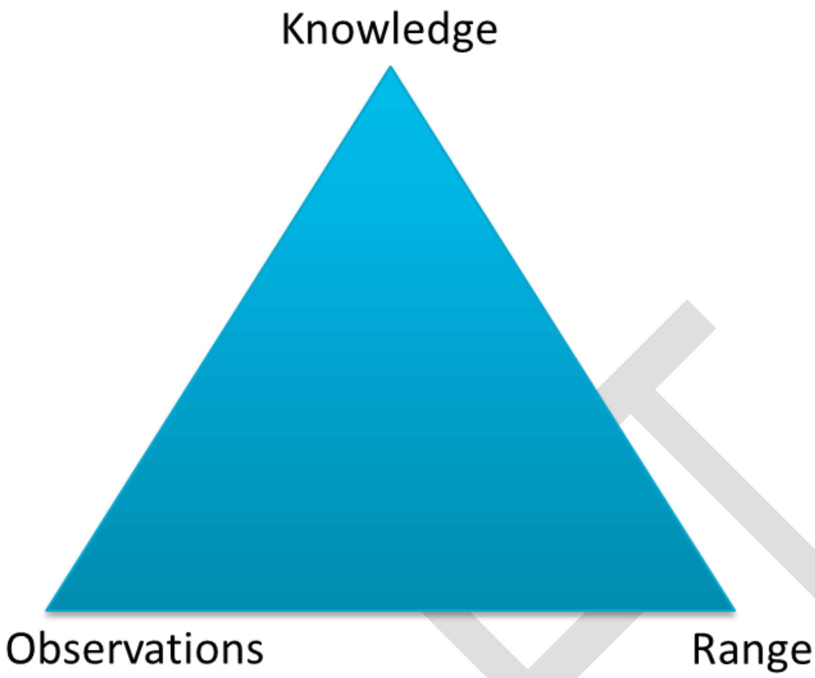


Figure 6 Trade-off between estimating the range of a prediction, being consistent with the available knowledge and honouring relevant observations

Figure 6 shows the three objectives of uncertainty analysis as the endmembers of a ternary diagram to highlight that a groundwater model with an uncertainty analysis always represents a trade-off between the three objectives. The examples below illustrate that an appropriate trade-off depends entirely on the context of the project and that different trade-offs can be justified for uncertainty analysis. While the examples are hypothetical, each includes a reference to a publicly available case-study that is illustrative for the example.

- **Observations + Range of QoI > Knowledge**
  Consider a model to predict dewatering rates for an open pit mine. The geology is very complex with steeply dipping layers and faults with unknown hydraulic behaviour. There is an existing monitoring network to measure groundwater levels in the vicinity of the existing mine and several pumping tests have been carried out. With such a database of historical observations, the structure and parameter values of the model can be inferred through history matching and justified. The range of model predictions can subsequently be estimated through linear uncertainty analysis or ensemble methods. This trade-off is emphasizing the honouring of observations and estimating the range of predictions over the consistency with system knowledge. Note that this does not mean the model is inconsistent with system knowledge. Rather, it conveys that the representation of reality is simplified, for instance by having a limited number of model layers with a constant thickness rather than a model grid that closely follows the complex geology. **Hypothetical example inspired by**: Simultaneous interpretation of six pumping tests in the Pilbara (Manewell and Doherty, 2021).
- **Knowledge + Range of QoI > Observations**
  Consider a model to predict drawdown in a fractured and weathered unconfined aquifer due to coal seam gas depressurisation. The existing monitoring is scarce and mostly strongly influenced by local processes, such as surface water-groundwater interaction, that will not be able to greatly constrain or condition hydraulic properties relevant to predicting mine water inflow, such as the transmissivity and storativity of the coal measures and the aquitard. Reservoir modelling by the proponent does provide an estimate of the

expected volumes of coproduced water. In this setting, it can be justified to focus on accurately representing the geometry of the underground, the planned development and defining prior distributions for hydrogeological properties based on site characterisation and literature, and on using ensemble methods to provide an estimate of the range of drawdown consistent with this existing system knowledge, including the coproduced water volumes expected by the proponent.

Hypothetical example inspired by: Bioregional Assessments for the Gloucester subregion (Peeters et al. 2018).

- **Observations + Knowledge > Range of QoI**

  Consider a model to investigate whether a fault near a proposed mine can be considered a barrier to groundwater flow. As part of the investigation, a long-duration pumping test has been carried out, with drawdown measured in observation wells on both sides of the fault.

  In this case, it can be justifiable to focus on accurately representing the fault geometry and pumping test set up in the model and to use a highly parameterised model to capture the response to pumping in detail. The range of model predictions is then only used to estimate the likelihood of the fault being permeable, and the related implications.

  Hypothetical example inspired by: Understanding, detecting and conceptualising hydrogeological barriers (Marshall, 2021).

- **Range of QoI > Knowledge + Observations**

  Consider a screening model to prioritise areas for further research in estimating cumulative drawdown caused by multiple developments. In this case, it is advisable to be overly conservative to ensure only areas with a very low likelihood of being impacted are excluded from further research. Conservative model assumptions, i.e. assumptions that can be shown to overestimate impacts, provide confidence that the range of model prediction includes extremes of drawdown. Such assumptions are often not consistent with existing knowledge and will not be able to honour observations. An example is not representing recharge in the model. This is not consistent with available knowledge as the aquifer is known to receive recharge. Not accounting for recharge, however, will lead to an overestimate of drawdown.

  Hypothetical example inspired by: GBA Beetaloo drawdown estimates in Cambrian Limestone Aquifer. (GBA 2021).

- **Range of QoI = Knowledge = Observations**

  Consider a high profile mine development close to a wetland system with international ecological significance. In this setting, it is of utmost importance that the model results can withstand a high level of public scrutiny and that the uncertainty analysis can be shown to have estimated a wide range of model predictions, consistent with the current system understanding and honouring all relevant information.

  Hypothetical example inspired by: majority of coal mining and coal seam gas development referred to IESC for advice. These developments are typically heavily scrutinised and the opportunity for trade-offs between the range of QoI, knowledge and observations is limited. An example of finding a balance between these three components can be found in Mather et al. (2022), in the context of constraining the response of continental-scale groundwater flow for climate change.

The examples above illustrate that different trade-offs are possible. The list of examples is not exhaustive; there are more examples and combinations possible. Justifying these trade-offs typically requires being able to demonstrate which model aspects are important for the quantity of interest or that model choices are conservative and will overestimate the impact on the quantity of interest. This is discussed in more detail in the break-out boxes 4 and 5 on conservative assumptions and sensitivity analysis respectively.

As pointed out earlier, groundwater modelling should be considered more as a verb describing an iterative process of investigation, rather than simply as a noun describing a deliverable. The emphasis on the different components in the trade-offs are therefore not static and can change throughout the life of the project. In the initial stage of a project, the emphasis can be on screening models to prioritise areas for further investigation, then during characterisation the focus can be on matching field data and refining system knowledge, while later, when more data is available, the focus can be more predicting the range and honouring observations.

The examples also highlight that the trade-offs cannot be easily codified (see break-out box 2 on confidence classification;), they require an open and transparent discussion early in the modelling process. The trade-offs made at this stage will affect many of the other components of groundwater modelling and uncertainty analysis, as well as the time, budget and computational resources, which are discussed in the next section.

3.4 Feasibility: uncertainty analysis techniques

Uncertainty analysis techniques commonly used in groundwater modelling can be organized in three main groups: scenario analysis, linear error propagation and ensemble methods.

Scenario analysis describes approaches in which the range of model predictions is explored by evaluating a limited number of parameter combinations. An example is changing the recharge of a groundwater model by 10% and inspecting the changes in the results. The advantages of this method are that it is easy to implement and explain and only requires a single model run per scenario or parameter combination. The drawbacks are that this approach does not provide a quantitative estimate of likelihood and that uncertainty analysis is not comprehensive as the practical upper limit on the number of scenarios that can be evaluated and analysed is small, generally less than 10.

Linear error propagation techniques create a linear approximation of the effect of the parameters on the predictions. When the uncertainty in the parameters and observations can be characterized as normal or normal after log transform, error propagation equations can be used to propagate the uncertainty of the parameters onto the predictions. The linear approximation of the model is captured in the Jacobian matrix, which includes the change in each model outcome due to a small change in each parameter. The advantage of this approach is that it is a well-established technique that is computationally efficient. At a minimum, it requires two model runs per parameter. However, as the Jacobian matrix is an essential part of gradient-based optimization, and if the model has been calibrated or history-matched through gradient-based optimization, the Jacobian matrix for the calibrated parameter values is available from the optimization. Compared to evaluating the groundwater model, error propagation with Jacobian methods is computationally very efficient. The drawbacks of linear error propagation are that the results can be biased if the parameters, observation uncertainties or predictions are not normally distributed or the model cannot be accurately represented with a linear approximation, and that the uncertainty analysis is not comprehensive as only the parameter space in the immediate vicinity of the base parameter set is explored.

Ensemble methods rely on evaluating a large number of parameter combinations and estimating the likelihood of predictions based on the resulting ensemble of model outcomes. The advantages are that this approach does not require a linear approximation of the model or normally distributed parameters, observation uncertainties or predictions. The drawbacks are that estimating robust summary statistics, especially extreme percentiles of skewed distributions, requires large ensemble sizes (i.e. more model runs) and that finding ensembles that are consistent with observations is often extremely computationally inefficient, where only a fraction of evaluated parameter combinations are accepted in the final ensemble. Ensemble methods usually require hundreds to thousands of model evaluations.

Technological advances in uncertainty analysis techniques have mainly focused on overcoming the various drawbacks of each technique. An overview of some of the more common techniques used in groundwater modelling is provided in break-out box 6 on uncertainty analysis techniques.

Break-out box 6: Uncertainty analysis approaches

In practice, each project has limitations in time, budget and computational resources. The selection of uncertainty analysis technique will be a trade-off between the aspects of reliability discussed previously and the project limitations, which mainly focus on the time needed to develop the model and carry out the uncertainty analysis. These computational demands are affected by three main aspects:

1. **Complexity of the model**
   More complex models have finer discretization to represent spatial variations in structure, geometry and hydraulic properties and/or more complex representations of processes affecting fluxes into and out of the model. The development time and runtime tend to increase with increased complexity, while numerical stability tends to decrease.
2. **Number of parameters**
   Each aspect of the model that needs to be included in the uncertainty analysis needs to be represented by parameters. Especially when representing spatial or temporal variability, this can lead to a very large number of parameters. More parameters allow for more detail to be captured and increase the fit to observations. The number of model evaluations needed for uncertainty analysis generally increases with the number of parameters. Calculating a Jacobian matrix for a model with 100 parameters for instance would require at least 200 model runs.
3. **Number of model evaluations**
   The robustness of likelihood estimates from ensembles increases with the number of model evaluations. While the mean and standard deviation can often be robustly estimated from a small ensemble of around 100 model evaluations, estimating more extreme percentiles, such as the 90^{th} or 95^{th} percentile, often require many more model evaluations, often more than 1000, especially if the model distribution is skewed. A summary statistic is considered robust if the value has converged, i.e. that the value of the summary statistic does not change by increasing the ensemble size. This is further explored in break-out box 7 on prediction bounds and convergence.

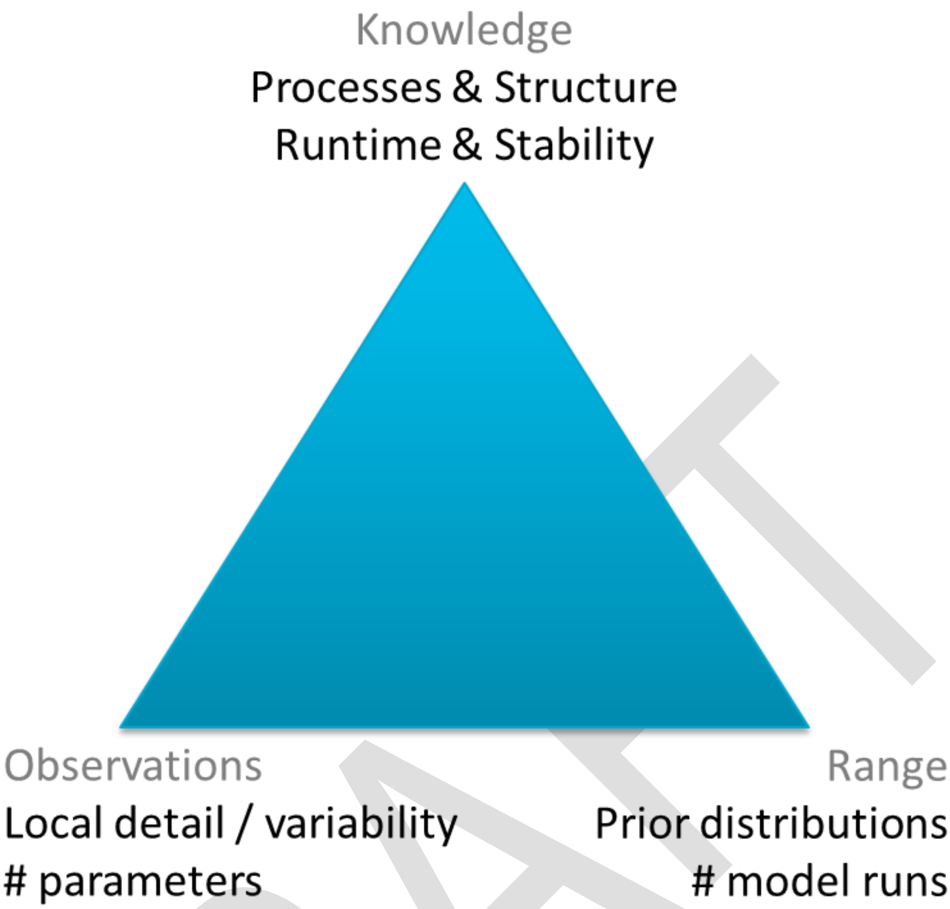


Figure 7: Trade-offs in reliability affect feasibility

Figure 7 illustrates how the aspects that affect the computational resources needed for a project are linked to the trade-offs in reliability. A focus on being consistent with system knowledge requires more complexity in the way structures and processes are represented in the model, which increases development time and runtime and decreases stability. An emphasis on honouring historical observations will require a model with many parameters to represent local detail and flexibility to adjust these parameters as a function of the observations. More parameters will increase the number of model evaluations needed for uncertainty analysis. When the range of model predictions is of importance, more time and resources need to be invested in characterising the prior distributions for each parameter, especially if these cannot be constrained or conditioned by available observations, and more model evaluations are needed if robust estimates of more extreme summary statistics are needed.

These trade-offs must be carefully justified in relation to the project resources and constraints, as discussed in the next section.

Break-out box 7: Prediction intervals and convergence

4 Reporting Uncertainty Analysis

Section 3 highlights the many assumptions, choices and trade-offs needed in a modelling and uncertainty analysis project. The main goal of the model report is to document, discuss and justify the assumptions, choices and trade-offs in an open and transparent way that is amenable to review. The report is crucial in decision making – less so the actual model dataset, scripts and results, as these are usually only accessible to the modelling team.

The key to successful communication is to present the information about uncertainty in a way that is most likely to aid decision-making. To achieve this, analysis of uncertainty information in model output needs to be: (i) adequately tailored to decision makers' needs; (ii) focused on the messages that are most likely to be relevant to their decisions; and (iii) presented in plain and clear (precise, non-technical) language (Richardson et al. 2017).

4.1 General

As outlined herein, the environmental assessment documentation should provide enough information to allow an independent reviewer (IESC or others) to evaluate the justifications for the assertions made. For example, on the basis for the conceptual or ecohydrological models, the underlying assumptions for the numerical models and why an ecosystem is or is not groundwater-dependent. An independent reader of the environmental assessment documents should be able to verify all significant assumptions, methodologies, techniques, assertions and conclusions made by the proponent. and evaluate whether the analysis effort applied is commensurate with the risk.

To that end, assessment reports should be as self-contained as possible, minimising the need for the reader to consult several volumes of technical reports to understand the methodologies and assumptions applied and the predictions of impacts on groundwater and dependent ecosystems (but allowing for the flow of information between disciplines conducting the assessments). For example, there are usually linkages between the groundwater and surface water assessments, such as to estimate recharge and/or runoff (Figure 8). Groundwater model outputs typically form inputs to evaluate the site water and salt balance (see also Figure 2) that is usually presented in the surface water report. The site water and salt balance itself produces outputs that become inputs to post-mining groundwater model scenarios, such as for the final void lake, which may also require information from geochemical assessments (Figure 9). Post-mining scenarios need information on closure from the rehabilitation report, and the groundwater model also needs information on groundwater-dependent ecosystems from the ecological report.

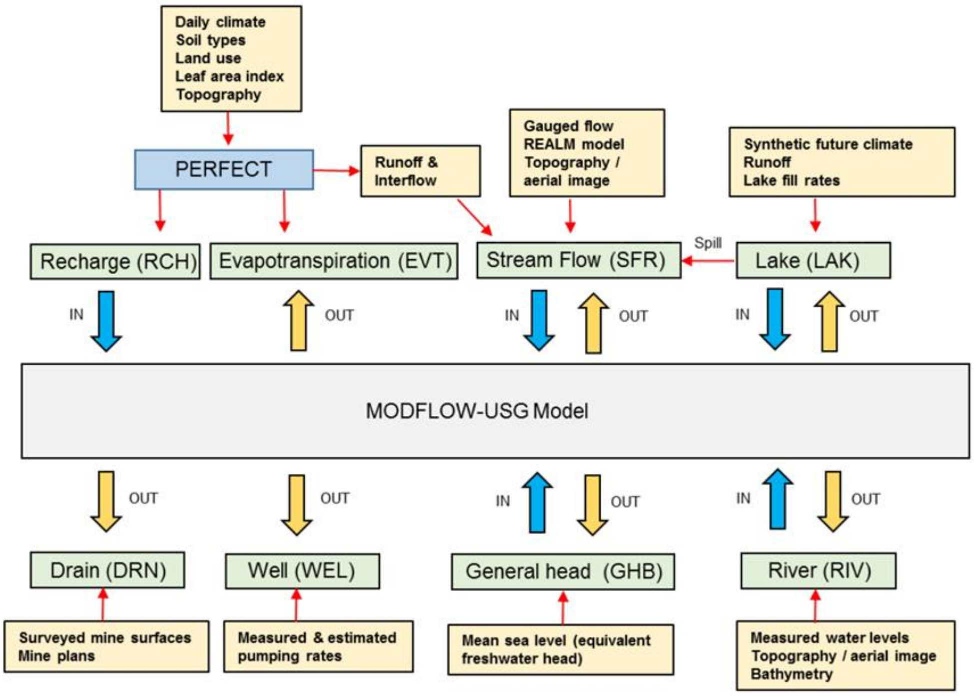


Figure 8 – example of interactions between groundwater and surface water models (after Gresswell et al. 2019). Terms in capitals refer to models or model components.

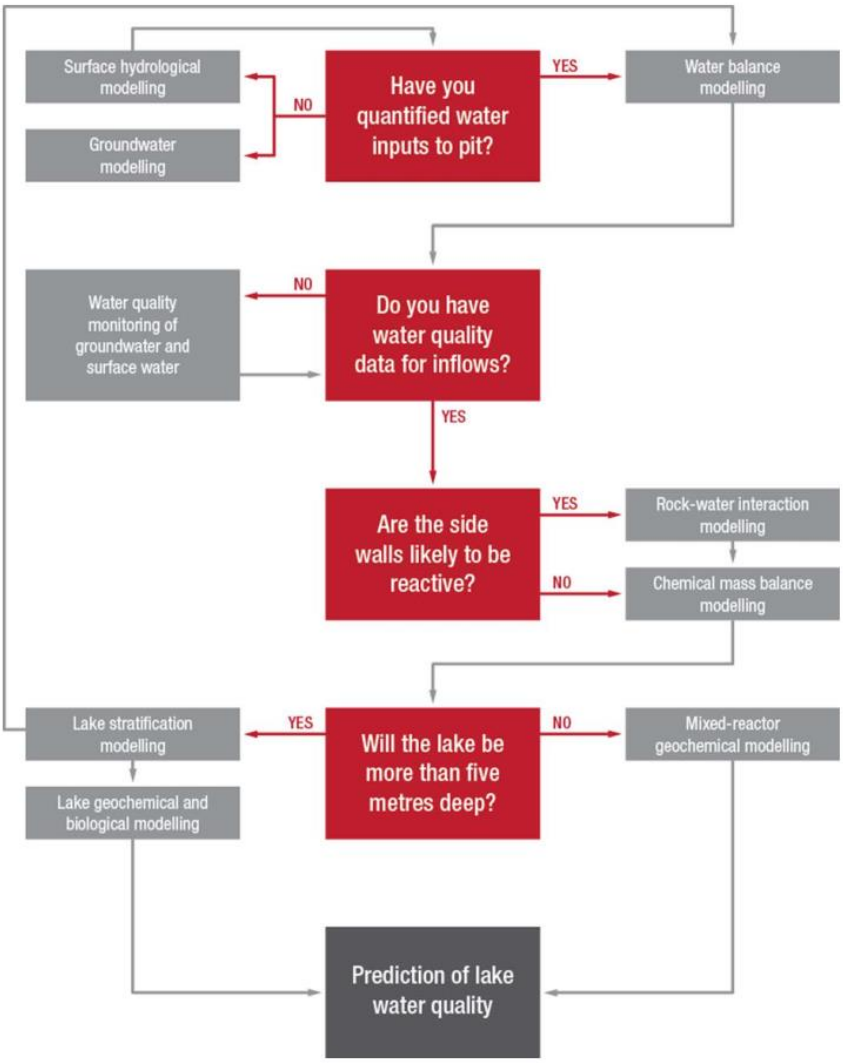


Figure 9 – example workflow for final void lake water balance and quality modelling (after WA DMIRS 2020)

4.2 Bias

The effects of bias must be acknowledged and discussed in the report.

Richardson et al. (2017) discuss some of the cognitive biases to which everyone is prone, such as availability bias, confirmation bias, and confidence and framing bias (see Glossary). Although a groundwater model is designed to be an objective representation of physical reality, the multitude of choices and assumptions that need to be made during modelling and uncertainty analysis make bias in predictions unavoidable.

Minimising and acknowledging bias in investigations and modelling of causal pathways is a key element of the ecological values analysis at the problem definition stage, along with data analysis, conceptualisation, and the initial risk analysis and treatment options assessment.

Conceptual models for resource development projects should consider and minimise potential biases when analysing how causal pathways can transmit direct, indirect and cumulative impacts from coal seams to water resources or water-dependent ecosystems. More than one model conceptualisation or realisation may need to be tested to understand the effect of conceptual or other sources of uncertainty and bias on model outputs. This may lead to more than one mathematical model, as outlined in the AGMG (Barnett et al. 2012). The multiple models may be of different types – e.g. conceptual, analytical or numerical – depending on the objective to be investigated.

While bias in modelling can never be eliminated, known biases need to be honestly and transparently communicated as part of the uncertainty analysis. From a management perspective, modelling is considered to have failed if there is sufficient bias for a poor decision to be made (e.g. through lack of transparency or inadequate uncertainty analysis), especially if the consequence is significant (Walker 2017).

4.3 Justifications

Most technical reports focus on what has been done and how it has been done, with often less focus on why a particular choice is made. This means that, despite often very lengthy and detailed descriptions of the model, the uncertainty analysis technique and the results, reviewers frequently ask for more information. Just as this document does not advocate for or prescribe a particular approach to uncertainty analysis, it does not recommend any particular structure or content in a report on groundwater model uncertainty analysis. A technical report should, however, provide clear, unambiguous and easy to find information to address three key questions:

1. How is this model relevant to a decision maker? What is the Quantity of Interest (or 'key prediction') and how will/can it be used in decision making?
2. What are the main sources of uncertainty affecting this QoI? How have they been identified and how are they represented in the model and uncertainty analysis?
3. What is the range of predictions? How is this range obtained and how is this range constrained or conditioned with available knowledge and observations?

The first question addresses the usability component of the fit for purpose trade-offs (section 3.2; Figure 3). Usability emphasises the importance of demonstrating why a model is needed and which of its outputs (key predictions) are relevant for decision making, including how uncertainty in the quantity of interest will be used in the decision-making process. This should be front and centre in a model report. It not only demonstrates the relevance of the model, but the justification of all assumptions, model choices and trade-offs (i.e. questions 2 and 3 above) will be evaluated in the context of this quantity of interest.

The main three approaches for such justification are (1) theoretical hydrogeological research, (2) project-specific conceptual hypothesis testing and (3) sensitivity analysis:

1. Theoretical hydrogeological research provides insights that transcend the specifics of a project. It can range from using the Glover-Balmer analytic solution for streamflow depletion to demonstrate that recharge does not affect streamflow depletion estimates (Glover and Balmer, 1954) to more elaborate work, such as Cook et al. (2016) on simulating the impacts of unconventional gas development on water resources, Peeters and Turnadge (2019) on when hydraulic properties can be constrained or conditioned with groundwater level observations, Marshall et al. (2019) on the effect of undetected barriers on groundwater drawdown and recovery or Herckenrath et al. (2015) on the potential bias due to ignoring the effect of a gas phase in simulating the effect of coal seam gas extraction.

2. Addressing more project- and site-specific assumptions or model choices can be done through project-specific hypothesis testing, using a combination of field work and modelling. Examples can include testing the hypothesis that a mapped groundwater-dependent ecosystem is relying on a regional unconfined aquifer, assessing the hydraulic behaviour of a fault through a pumping test across a fault or identifying the source aquifers of a spring complex. This is an opportunity to tie-in hydrogeological field investigations closer with the groundwater modelling and uncertainty analysis. Doherty (2022) presents an monograph on hypothesis testing and the scientific method, which can find practical application in demonstrating the low risk (or otherwise) of particular causal pathways for potential impacts.

3. The first two methods of theoretical research and project-specific hypothesis testing are ideal to justify model choices at the initial stages of developing the groundwater model and designing the uncertainty analysis. There will, however, always remain sources of uncertainty that, based on experience, would be considered of minor importance, but cannot be excluded conclusively based on theoretical grounds or field work. One approach can be to parameterise these sources of uncertainty and use a formal sensitivity analysis to quantify the effect of the source of uncertainty on the QoI. Such a check after the modelling has been concluded can provide confidence that the initial choice or assumption was justified.

A final comment on justification is that some trade-offs are pragmatically driven by constraints in budget, time and computational budget. It is recommended to report on these openly and transparently. An example would be a project assessing potential reductions in flow to a spring due to an underground mine over the life of that mine. The main source in simulating this reduction is the mine dewatering rate. This rate will not only depend on the mine design and the local hydrogeological conditions, but also on the economic conditions and mine management. Including all these economic and engineering scenarios will often be beyond the budget of the groundwater modelling project. The report should acknowledge this and that all results will be conditional to the scenario adopted, which is often the largest source of uncertainty.

4.4 Qualitative Uncertainty Analysis

Groundwater field investigations, modelling and uncertainty analysis tend to be very technical, detailed and steeped in jargon. This can be quite daunting for readers with limited background in these fields. The jargon is, however, necessary to accurately capture and justify the technical content of a report. A qualitative uncertainty analysis table (Peeters, 2017) is one of the ways to summarise that information in a more accessible, concise way without jargon.

Such a table lists the main assumptions and model choices and scores the reasoning as well as the potential impact on the QoI. Each assumption is scored on whether the assumption or model choice is driven by data availability, time and budget available for the project, or technical challenges. The scoring and justification can be facilitated by prompting questions such as:

- Would a different choice have been made if more/other data were available?
- Would a different choice have been made if more time were available?
- Would a different choice have been made if technical limitations were resolved?

The most important score, however, is the perceived effect of the assumption on the model outcomes. Summarising these scores in a table allows reviewers and stakeholders to quickly assess the importance of the various model assumptions which is particularly valuable in an environmental impact assessment.

Table 2 illustrates the concept of a qualitative uncertainty analysis with an assumption that is often made in groundwater modelling studies: representing aquifer properties as spatially uniform. The table shows both the scoring and its justification. This example highlights that it is possible to score 'medium' on the prediction attribute despite scoring 'high' on the data and resources attributes.

Table 2: Illustrative example of qualitative uncertainty analysis

Attribute	Data	Resources	Technical	Prediction
Score	High	High	Medium	Medium
Model choice: Spatially uniform aquifer properties	Large dataset of property observations needed to characterise *a priori* spatial variability. Large dataset of head or flux observations needed to infer posterior spatial variability.	Spatially heterogeneous parameterisation increases dimensionality of parameter inference and uncertainty quantification. As this increases the number of model evaluations, a larger budget and longer time frame is needed for modelling.	Most groundwater model codes allow spatially variable hydraulic properties. Efficiently implementing a spatial field generator that adheres to prior knowledge remains challenging.	If the range of equivalent properties in the uniform parameterisation capture natural variability, the range of drawdown predictions will be comparable to those of a spatially variable parameterisation. Drawdown predictions will be locally different if spatial heterogeneity is high or localised (such as through faults).

4.5 Calibrated Language

The presentation of outcomes of uncertainty analysis remains challenging. Outputs of groundwater models are often multi-dimensional (space, time, different state variables). Uncertainty analysis adds another dimension to this, making reporting and visualising results challenging. It is recommended to use calibrated language on likelihood throughout a report. Using consistent and precise language to communicate the analysis will help to prevent the subjective biases of the water manager or the project proponent affecting their decision-making. It is critical for all parties not to distort the implications of the findings presented in the assessment. To reduce the scope for distortion, the modeller should present the methods and results in a way that is not open to misinterpretation.

The IPCC (2013) devised a set of narrative descriptors of the likelihood of future climate outcomes that relate directly to probability classes (reflecting uncertainty). Those principles have been combined with risk-based visualisation methods to develop an approach (Table 3) to effective communication (after Richardson et al. 2017). This comprises:

- narrative descriptors of the likelihood of a given outcome, with careful consideration of which description best fits the impact being assessed
- quantitative ranges of probabilities from an uncertainty analysis
- qualitative visual methods (risk assessment style colour-coding).

In Table 3, the colour coding relates to the likelihood of exceedance and is designed to support the narrative descriptions. The 10th (dark green) and 90th (red) percentiles each have about a 10 per cent probability of occurring.

For the 10th percentile, there is approximately a 10 per cent probability that the outcome will be *less* than the prediction, while for the 90th percentile there is approximately a 10 per cent probability that the outcome will be *greater* than the prediction. Approximately 80 per cent of outcomes will lie within the light green to orange categories. It is important to note that an 80 per cent probability based on a set of 1000 simulations means that 200 simulations predicted outcomes outside the criteria range selected.

Table 3: Example of a combined numeric, narrative and visual approach to describing likelihood

PERCENTILE (outcomes ranked from small to large)	DESCRIPTION (in terms of likelihood of exceedance)	ALTERNATIVE DESCRIPTION OR FRAMING
<10%	It is **very likely** that the outcome is **larger** than this value	It is **very unlikely** that the outcome is **smaller** than this value
10–33%	It is **likely** that the outcome is **larger** than this value	It is **unlikely** that the outcome is **smaller** than this value
33–67%	It is as **likely as not** that the outcome is **larger** than this value	It is as **likely as not** that the outcome is **smaller** than this value
67–90%	It is **unlikely** that the outcome is **larger** than this value	It is **likely** that the outcome is **smaller** than this value
>90%	It is **very unlikely** that the outcome is **larger** than this value	It is **very likely** that the outcome is **smaller** than this value

5 Technical aspects / break-out boxes

5.1 Break-out box 1: Modelling guidelines

Hill et al. (2004) summarise some fundamental principles that are shared by all groundwater modelling guidelines, based on ideas set out in the MDBC modelling guidelines report (Middlemis et al. 2001) that was the foundation for the AGMG (Barnett et al. 2012). These principles remain valid for this Explanatory Note:

The aim of most guidelines is to reduce and reveal model uncertainty for the users of modelling studies, including resource management decision makers and the community. This is achieved by promoting transparency in modelling methodologies and encouraging innovation, consistency, and best practice. Guidance is provided to non-specialist modelers and auditors or reviewers of models by outlining the steps involved in scoping, managing, and evaluating the results of groundwater modelling studies. Guidelines serve modelling specialists by providing a baseline set of ideas and procedures from which they can innovate.

Guidelines are intended for use in raising the minimum standard of modelling practice and allowing appropriate flexibility, without limiting necessary creativity or rigidly specifying standard methods. Guidelines also should not limit the ability of modellers to use simple or advanced techniques, appropriate for the study purpose. Techniques recommended in the guidelines may be omitted, altered, or enhanced, subject to the modeller providing a satisfactory explanation for the change and negotiation with the client and/or regulator as required. Not all aspects of guidelines would necessarily be applicable to every study. It also is acknowledged that there is a need for subjective judgment during the model development process.

Guidelines should be seen as a best practice reference point for framing modelling projects, assessing model performance, and providing clients with the ability to manage contracts and understand the strengths and limitations of models across a wide range of studies (scopes, objectives, budgets) at various scales in various hydrogeological settings. The intention is not to provide a prescriptive step-by-step guidance, as the site-specific nature of each modelling study renders this impossible, but to provide overall guidance and to help make the reader aware of the complexities of models, and how they may be managed.

The guiding principles outlined above allow for flexibility, but this Explanatory Note insists on minimum standards:

- clear definition of the quantity of interest and the model outcomes sought in specific terms
- justification of the methods, assumptions and assertions
- provision of objective evidence that the uncertainties affecting decision-critical predictions of impacts on aquifer resources and dependent systems are not underestimated
- transparent documentation of the methods and results in a manner that is open to scrutiny, and consideration of the effects of potential bias.

5.2 Break-out box 2: Australian Groundwater Modelling Guidelines Confidence Classification

The 2012 Australian Groundwater Modelling Guidelines (AGMG, Barnett et al. 2012) introduced the concept of a confidence classification. It is a rubric: a checklist in which the confidence of the model as a whole is determined based on a number of criteria related to the design of the model and the observation data available. It is interesting to note that the AGMG also stated that '*the confidence level classification of the model predictions can be expressed quantitatively in a formal model uncertainty analysis*'. In the decade since the AGMG publication, the confidence level concept has become redundant, replaced by the development and uptake of efficient and effective uncertainty analysis methods.

The AGMG approach to assessing confidence in a model is inconsistent with the foundational principles of this Explanatory Note (based on Saltelli et al. 2020; see section 1.1). The confidence level classification can be very misleading and can lead to unrealistic expectations that nothing less than a 'high confidence' model is 'acceptable', when that is not warranted or justified. It is unlikely that the rubric adequately captures whether the assumptions and model choices that underpin a groundwater model and associated uncertainty analysis are appropriate to represent or understand the quantity of interest relevant for decision-making on any project.

The AGMG is currently being revised (2022) and the confidence level scheme will likely be replaced by uncertainty analysis methodologies (consistent with this Explanatory Note). The first version of this Explanatory Note (Middlemis and Peeters, 2018) recommended to no longer use the model confidence level rubric to commission, design or review groundwater models. The 2022 version of the Information Guidelines reinforced this recommendation by stating that the confidence classification should not be used as determining factor in deciding whether a groundwater model is fit-for-purpose.

Rather than relying on such a rubric, each model report should clearly identify and discuss the model choices and assumptions and to what extent these may affect the simulated quantity of interest and more importantly, the findings of the modelling investigation. Reporting of uncertainty analysis, with an emphasis on justifying model assumptions, is discussed in Section 4.

5.3 Break-out box 3: Sources of uncertainty

The subsurface environment is complex, heterogeneous and difficult to directly observe, characterise or measure. Groundwater systems are influenced by geology, topography, vegetation, climate, hydrology and human activities. Thus, uncertainty affects our ability to accurately measure or describe the existing or predicted states of these systems.

Walker et al. (2003) provide an overview of uncertainty definitions for model-based decision support which, almost two decades later, is still relevant. They recognise two types of uncertainty: epistemic and variability uncertainty. Variability or aleatory uncertainty is due to inherent variability. An example is a fair coin; we know that the probability of heads is 50%, but we do not know if the next toss will result in heads or tails. Epistemic uncertainty is uncertainty due to imperfect knowledge. The difference between both types of uncertainty is important from a theoretical point, as epistemic uncertainty can be reduced by improving our knowledge of the system, while variability uncertainty cannot be reduced. In practice, however, most sources of uncertainty in groundwater models will have both an epistemic and variability component. Uncertainty in recharge, for example, is both aleatory, as it depends on the natural variability of rainfall, and epistemic, as we often do not fully capture the recharge processes in a groundwater model.

Sources of uncertainty are often organised in categories such as structural, conceptual, input, parameter, scenario and measurement uncertainty. Such categories are helpful when identifying which sources of uncertainty to consider in an uncertainty analysis. While many of us have an intuitive understanding of these terms, clearly defining each source of uncertainty is much more challenging. Take recharge for instance. Recharge can be categorized as:

- conceptual model uncertainty, if it is not clear which recharge process is most important,
- structural uncertainty if recharge depends on which hydrostratigraphic unit is outcropping,
- input uncertainty as recharge is one of the driving forces of a model, parameter uncertainty if it is varied in calibration and uncertainty analysis,
- measurement uncertainty if the value of recharge is based on different recharge estimation techniques and
- scenario uncertainty if future projections of climate are considered.

For clarity, in this text we will use the term ''parameter in its most general sense; a parameter is any aspect of a model that is allowed to vary in an uncertainty analysis.

5.4 Break-out box 4: The precautionary principle and conservatism

The precautionary principle is a key component of many environmental policies. While there are many definitions of this principle, they often can be traced back to the formulation in Principle 15 of the 1992 Rio Declaration on Environment and Development:

In order to protect the environment, the precautionary approach shall be widely applied by States according to their capabilities. Where there are threats of serious or irreversible damage, lack of full scientific certainty shall not be used as a reason for postponing cost-effective measures to prevent environmental degradation.

The principle provides a rule for environmental decision-making that complements risk-based decision-making. The interpretation and implementation of this principle can, however, be challenging, as discussed in Steel (2014) and Stefánsson (2019). Application of the precautionary principle can lead to a reversal of the burden of proof (Peterson 2006), where a proponent needs to demonstrate that a development will **not** harm the environment with sufficient certainty. This is the context in which conservatism in modelling features, in that conservative model choices and assumptions bias a model by overestimating impacts. This provides confidence in the model results that the 'bad thing' is not going (or is very unlikely) to happen.

Underschultz et al. (2018) showed that conservatism in geological representation and in groundwater modelling led to overestimated impacts of coal seam gas development. Reported (measured) water and salt production in Queensland were about 25% of those predicted by academia and government and about 70% of those predicted by industry.

Making conservative assumptions is especially relevant for model aspects that are not usually varied in uncertainty analysis (for instance, conceptual model choices) and in defining prior distributions for parameters that cannot be constrained by the available observations. Which model choice will lead to an overestimate depends on the kind of prediction and is not always trivial, especially in complex models.

Consider a project with two quantities of interest: drawdown in an unconfined aquifer and streamflow depletion in a river connected to that unconfined aquifer. Drawdown increases with decreasing transmissivity and storativity, so underestimating transmissivity and storativity will lead to an overestimate of drawdown and is therefore conservative. Streamflow depletion, however, increases with increasing transmissivity and storativity. The conservative choice would be to overestimate streamflow depletion. It is not possible to make choices that lead to overestimates for both QoIs, so conservatism is not a valid approach to guide model development. In this case, both transmissivity and storativity should be included in the uncertainty analysis and constrained or conditioned by available observations.

Another example is to consider a situation with a single QoI, drawdown prediction in an unconfined aquifer, with transmissivity and storativity included in uncertainty analysis, constrained or conditioned by historical observations of groundwater level. The groundwater level observations are, however, influenced by historical pumping rates that have not been metered. The available computational budget does not allow inclusion of historical pumping rates in the uncertainty analysis. Should the pumping rates in the model be overestimated or underestimated? History matching with fixed pumping rates will lead to biased hydraulic properties. If pumping rates are overestimated, the parameter inference will compensate for that by increasing hydraulic properties. Vice versa, underestimated pumping rates will lead to underestimated hydraulic properties. In this case, underestimating pumping rates during history matching is justified as it will ensure that estimated hydraulic properties in the vicinity of pumping wells are underestimated and will lead to overestimated drawdown.

These examples illustrate that making conservative choices is not trivial and that, especially in high-profile projects, one should strive to include and parameterise as many known sources of uncertainty as possible.

5.5 Break-out box 5: Sensitivity analysis

Sensitivity analysis complements an uncertainty analysis, but it is not an alternative to uncertainty analysis.

In its most general sense, sensitivity analysis is the study of how the 'outputs' of a 'system' are related to, and influenced by, its 'inputs' (Razavi et al. 2021). In the context of groundwater modelling, the 'system' is the groundwater model, 'inputs' are the aspects of the model that are parameterised and 'outputs' are the simulated equivalents to observations and the quantity of interest. Where the main goal of uncertainty analysis is about quantifying uncertainty in outputs, sensitivity analysis focuses on identifying the main sources of uncertainty. The outcome of uncertainty analysis is directly relevant to decision makers (e.g. is the likelihood of an unwanted event sufficiently low for the risk to be acceptable?). The outcomes of sensitivity analysis are directly relevant to modellers as they provide additional insight into the model. This makes sensitivity analysis relevant to decision-makers as well, as it not only provides additional confidence in the model results, but it also allows development of efficient strategies to reduce uncertainty in future.

Razavi et al. (2021) provide an excellent review of the current state of the art in sensitivity analysis, including an overview of the main approaches and software implementations. They distinguish between local sensitivity analysis and global sensitivity analysis. Local sensitivity analysis is based on changing one or more parameters from an initial 'base' value and evaluating the resulting changes in outcome. This is the most commonly reported sensitivity analysis in groundwater modelling practice. Analysis of sensitivity based on a Jacobian matrix (see break-out box 6 on uncertainty analysis techniques), often complementing automated history matching and linear error propagation, should also be considered a local sensitivity analysis, as the Jacobian matrix is populated by perturbing a 'base' parameter combination. Local sensitivity analysis only explores a very small part of parameter space, especially for highly parameterised models, and the results can therefore be misleading, particularly when the model response to changes in parameters is not linear (Saltelli and Annoni, 2010).

Global sensitivity analysis addresses this issue by evaluating the sensitivity of outputs based on multiple parameter combinations. Global sensitivity analysis methods differ in the selection and number of parameter combinations. Readers are referred to Razavi et al. (2021) for a comprehensive review. Methods that are especially relevant for groundwater modelling and uncertainty are given-data sensitivity analysis that provide sensitivity measures based on existing ensembles of model evaluations (e.g. Plischke et al. 2013, Pianosi and Wagener, 2018). As these approaches can be used on an existing ensemble of model evaluations, like those generated in ensemble uncertainty analysis techniques, they come at a very small additional computational cost. It may be tempting to consider that sensitivity indices based on a Jacobian matrix obtained vicariously through the Iterative Ensemble Smoother ('IES'; White, 2018) are global in the sense that they are based on multiple parameter combinations. However, they still have many of the drawbacks of local sensitivity analysis as they are based on a linearisation of the model and limited sampling of parameter space. In such cases, there would be much greater understanding obtained from the application of IES non-linear methods and analysing the posterior ensemble (see break-out box 6 on uncertainty analysis techniques).

Data worth analysis extends sensitivity analysis by quantifying the reduction in predictive uncertainty should new observations, with a known measurement uncertainty, become available (Dausman et al. 2010, Gosses and Wohling, 2021). The closely related concept of Value of Information takes this further by evaluating if the change in uncertainty due to new information would change the decision and if the cost of acquiring new information outweighs the cost of the unwanted outcome eventuating, based on the current probability (Bratvold et al. 2009).

Data worth analysis extending a linear error propagation (see Break-out box 6), does not require additional model runs, provided the Jacobian matrix includes the sensitivity to the hypothetical observation. The main drawbacks are that it requires the same assumptions for linear error propagation to be valid and, more importantly, that the new observation would not change the calibration to the extent that the parameter sensitivities change. Data worth analysis with ensemble methods is computationally very demanding as it involves randomly drawing a hypothetical observation and evaluating a Bayesian inference with the hypothetical observation added to the observation dataset. The advantage of this approach is that it is more robust against new observations changing the estimated parameter values.

Both approaches remain however fragile to new observations challenging the current conceptualisation or parameterisation, the so-called conceptual surprises (Bredehoeft, 2005). A more robust approach would be to use the results of a sensitivity analysis to identify the main sources of uncertainty and focus field investigations on addressing any knowledge or data gaps that relate to these sources of uncertainty.

5.6 Break-out box 6: Uncertainty analysis techniques

Uncertainty analysis is important for all scientific and engineering challenges. This has also led to an explosion of jargon, mixing terms from different disciplines. This break-out box provides a high-level overview of uncertainty analysis and the context for commonly used terms in uncertainty analysis (see also the Glossary).

The first goal of uncertainty analysis is to quantify a range of model predictions that reflect the effect of uncertainty in model parameters on the model outcomes. A model **parameter** is any aspect of the model that is considered uncertain and that is allowed to vary during the analysis. Parameters can represent uncertainty in model input, properties, boundary conditions and even structure or conceptualisation if carefully designed.

To quantify the uncertainty in model outcomes, the model must be evaluated in such a way that uncertainty in parameters is propagated to the model outcomes. In most modelling disciplines, this is called the **forward model**. The uncertainty in parameters is captured in the **prior probability distribution** or **prior** for short. This term comes from Bayesian statistics and it expresses one's belief about the variability of parameters. When parameters are correlated, which is often the case for spatially varying parameters, the **joint prior probability distribution** needs to be specified. Priors of parameters that can be assumed to be normally or log-normally distributed, can be fully described by the **mean** and **variance** (the square of **standard deviation**) and correlations between parameters is captured in the **covariance matrix**. In geostatistics, a **covariance function** is used to describe how a parameter varies in space.

Scenario analysis assesses uncertainty by evaluating the forward model for a number of scenarios, where a scenario is a model run in which one or more parameters is changed compared to a reference parameter combination or baseline model. The scenarios provide information on what model outputs are possible, but generally only have a subjective/qualitative assessment of likelihood, and scenarios need to be carefully designed and justified. Maier et al. (2016) show how such scenario analysis, such as to evaluate extreme parameter combinations, can be highly relevant to decision makers. The parameter combinations chosen in a scenario should be part of the prior, but scenarios usually only change a subset of the parameters that are considered uncertain.

Linear error propagation is a more comprehensive and quantitative uncertainty analysis technique in which the effect of the parameters on the outcomes is approximated by a linear function such as the **Jacobian matrix** which is the matrix of all first order partial derivatives of model outputs to parameters. The partial derivatives are numerically approximated by the slope of changes in output by perturbing each parameter of the model by a small amount. If the variability in parameters can be fully described by standard deviation and covariance matrix and the model output is a linear function of the parameters, the uncertainty in model outcomes is obtained by combining the Jacobian and the covariance matrix. The main advantage of this method is that it is computationally efficient as it only requires (at a minimum) two model runs per parameter. The main drawback is that the estimates of uncertainty are biased (i.e.

over or underestimated) if the parameters are not normally distributed or the model outputs are not a linear function of the parameters.

Ensemble methods for uncertainty analysis overcome some of the issues of scenario analysis and linear error propagation by drawing a number of random parameter combinations of the prior parameter distributions and evaluating the model for each parameter combination. The uncertainty in model outcomes can then be empirically estimated from the resulting ensemble of model outcomes. This is referred to as **Monte Carlo** sampling, to emphasise the random component. The advantage of this approach is that it does not require assumptions on normality of parameters or linearity of the model. The drawback is the computational cost of evaluating the model for each randomly generated parameter combination. Break-out box 7 on uncertainty bounds and convergence discusses the effect of ensemble size on the accuracy of the uncertainty analysis.

The second and third goal of uncertainty analysis is for the range of predictions to be consistent with system knowledge and to honour historical observations. The latter part, on how to constrain or condition model predictions with observations, has been a focus of research for many decades and is referred to as **calibration, history matching, data assimilation** and **inverse modelling**. A generic term is **parameter inference**, as the goal is to infer parameters from observations. The term **Bayesian inference** is used when Bayes law is formally used to estimate a **posterior probability distribution** or **posterior**. Bayes law describes how the probability of a model can be estimated from prior belief in the model (the prior of parameters) and the evidence (the dataset of historical observations). For a model M and dataset D, Bayes law can be written as:

$$P(M|D) = \frac{P(D|M)P(M)}{P(D)}$$

in which $P(M|D)$, the probability of a model, can be estimated from $P(M)$, the prior of the model, $P(D)$, the probability of the dataset and $P(D|M)$, the likelihood of the dataset given the model. This last term is usually computed based on the mismatch between observed and simulated values, accounting for observation uncertainty.

Many techniques for parameter inference are designed as optimisation algorithms with the goal of minimising the difference between observed and simulated values. Quantifying uncertainty bounds of inferred parameters is often only a secondary objective. In some disciplines, like geophysics, the parameters are the QoI (e.g. electric conductivity of the subsurface inferred from an airborne electromagnetic survey). In groundwater modelling however, the parameters are seldom the QoI and the uncertainty in parameters captured in the posterior still needs to be propagated to the QoI.

A common choice in **scenario analysis** is that the base case or reference scenario is the parameter combination that best fits the data. This parameter combination is often obtained through manual calibration, i.e. sequentially adjusting parameter values based on the modeller's experience until a predefined calibration target is achieved or the available computational budget is exhausted.

Many iterative least squares curve fitting methods, like the Levenberg-Marquardt algorithm implemented in PEST, use the Jacobian matrix to find the local minimum of the objective function. Linear error propagation can therefore be readily implemented as the Jacobian matrix is available (Moore and Doherty, 2005). These algorithms are computationally efficient but prone to get stuck in a local minimum rather than finding the global minimum of the objective function. While computationally efficient, the number of model runs still scales with the number of parameters, which means that these methods are still not feasible for highly parameterised models, i.e. hundreds of parameters. The **Iterative Ensemble Smoother (IES)** method (Chen and Oliver, 2013; White 2018) approximates the Jacobian by estimating it from an ensemble. This allows assimilation of data very efficiently, within around 1,000 model evaluations, regardless of the number of parameters. Evensen (2018) does point out however that the IES does not precisely sample the posterior probability distribution. This means that using the IES does not guarantee that prediction intervals from IES ensembles are accurate or that extremes are properly characterised.

Ensemble methods aim to estimate the posterior probability distribution by randomly sampling from the prior using a likelihood function, where parameter combinations that result in a small mismatch between observed and simulated values are more likely to be accepted in the posterior distribution. Many proposed parameter combinations from random sampling of the prior can have a low likelihood and will therefore not be accepted in the posterior. **Markov chain Monte Carlo (McMC) sampling** is a more efficient sampling strategy in which a new parameter combination is chosen in such a way that it is more likely that the new parameter combination has a greater likelihood than the previous. McMC methods mainly differ in how they select the next parameter combination in a sequence. A good overview of these methods, including **DiffeRential Evolution Adaptive Metropolis (DREAM)** sampling, can be found in Vrugt and Massoud (2018).

The likelihood function that underpins McMC sampling often assumes that the model error is normally distributed, which cannot always be justified (Schoups and Vrugt, 2010). **Generalised Likelihood Uncertainty Estimation (GLUE)** and **Approximate Bayesian Computation (ABC)** (Vrugt and Beven, 2018) instead use an approach in which samples are accepted in the posterior if a summary metric of the mismatch between simulated and observed values is less than a predefined threshold. The commonly used approach in groundwater modelling practice of evaluating a random ensemble of parameter values from the posterior and only accepting those for which the root mean squared error is below a predefined threshold, can be considered as ABC.

Common to all parameter inference methods is that they require many model evaluations, often proportional to the number of parameters. One way of overcoming this problem is making the inference methods more efficient, i.e. require fewer model evaluations. Another way is by reducing the runtime of models. **Model emulators** or surrogate models are black-box models that reproduce the dynamics of a model, i.e. how model outputs vary as a function of the parameter values but run much quicker. An overview of such approaches is presented in Asher et al. (2015). Recently, machine learning algorithms are increasingly considered as model emulators (Razavi 2021).

Another common feature to all parameter inference methods is that they assume that (1) the model is an adequate, unbiased simulator of reality and that (2) the prior parameter combination contains the posterior, i.e. that the parameter combination with the lowest error is within the prior. The first assumption can often not be strictly justified as each model will have a structural error component. This can be due to factors such as the resolution of the model, inaccuracies in historical boundary conditions such as pumping rates or lack of spatial variability. Structural error can lead to biased parameters, as parameter inference attempts to compensate for these structural errors (White et al. 2014). The effect of structural errors can be mitigated by increasing complexity of models or increasing parameterisation, but not completely eliminated. The second assumption is seldom tested and if it is not true it will also lead to a biased posterior as the area of highest likelihood is not included. The effect of both structural error and incorrect priors often manifests itself in posterior parameter distributions that are close to the bounds of the prior distribution. Scheidt and Caers (2018) recommend before starting a parameter inference, to evaluate a small, initial ensemble of the prior and test if the resulting ensemble of simulated values encompasses the observed values. Such a small initial ensemble also serves as a 'stress test' of the model to test if the model is stable and converging over the entire parameter range of the prior.

5.7 Break-out box 7: Uncertainty bounds and convergence

Ensemble methods rely on Monte Carlo or random sampling of a distribution to estimate summary statistics, such as the mean and prediction intervals. The 95th percentile prediction interval is the range spanned between the 2.5th and 97.5th percentile of a distribution. The accuracy of the estimates of mean and prediction interval increases with the size of the ensemble. Creating large ensemble sizes, with more than 1,000 or 10,000 model realizations is often computationally very expensive and not feasible within the constraints of a project. A much-used approach to assess whether the size of the ensemble is sufficient is to create a convergence plot. In a convergence plot, the evolution of the summary metric is plotted as more evaluations are added to the ensemble. The summary statistic is said to have converged if its value does not change with adding more values. Such plots can be quite misleading. Figure 10 shows the convergence plot for the 5th, 50th and 95th percentile of a skewed distribution. The left plot uses the full ensemble size of 1,000 evaluations and the right plot only the first 100 evaluations. Looking at the plot of 100 evaluations in isolation would indicate that the 95th percentile is stable and the ensemble size is adequate. The plot of 1,000 evaluations shows however that the 95th percentile only really stabilizes after 300 model evaluations.

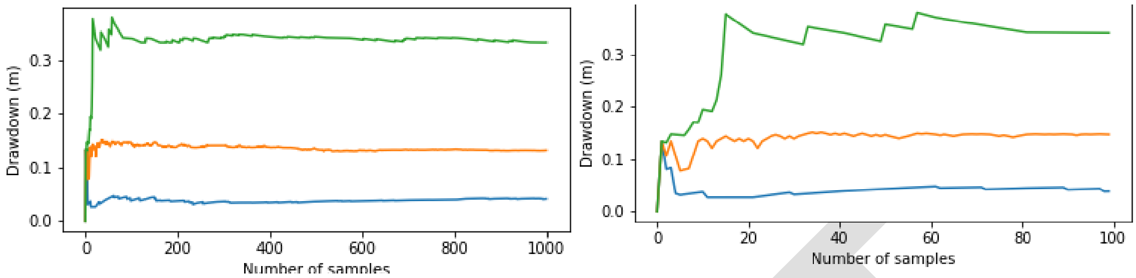


Figure 10: Convergence plot of 5th (blue), 50th (orange) and 95th (green) percentile for 1,000 model evaluations (left) and 100 model evaluations (right)

The accuracy of prediction intervals from Monte Carlo sampling is explored in more detail by Roy and Gupta (2021) (Figure 11). They show that the uncertainty in estimates of extreme percentiles can be considerable for small ensemble sizes (100), especially for skewed distributions.

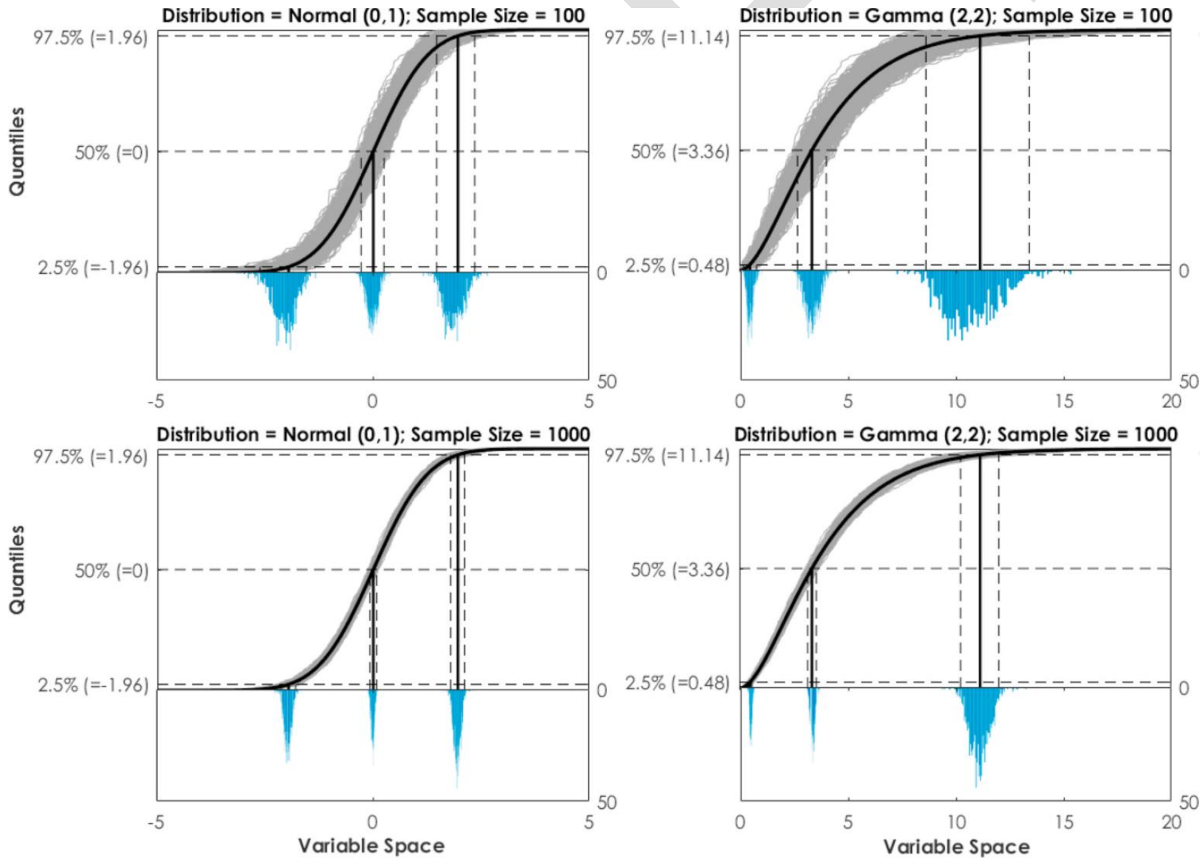


Figure 11: Illustration of sampling a normal (left) and skewed (right) distribution 1,000 times with a sample size of 100 (top) and sample size of 1,000 (bottom). The blue histograms show the distribution of estimated 2.5th, 50th and 97.5th percentiles. (after Roy and Gupta, 2021)

Roy and Gupta (2021) developed a method to adjust for this uncertainty by inflating prediction intervals through selecting a more extreme percentile as a function of the ensemble size (Table 4). The adjustment does not depend on the shape of the distribution that is sampled. Note that for small ensemble sizes, less than 100, the adjustment is considerable and that it is unlikely that 95th percentile prediction intervals can be reliably estimated.

Table 4: Prediction interval adjustment to account for limited ensemble size. The theoretical 90% prediction interval (the range spanned by the 5th and 95th percentile) for an ensemble with size 100 can be approximated by the prediction interval spanned by the 2nd and 98th percentile estimated from the ensemble. Values are calculated using equation 3 in Roy and Gupta (2021).

Theoretical	N=25	N=50	N=100	N=500	N=1,000	N=10,000
97.5	99.9	99.6	99.3	98.6	98.3	97.8
95	99.3	98.6	98.0	96.6	96.2	95.4
5	0.7	1.3	2.0	3.4	3.8	4.6
2.5	0.1	0.4	0.7	1.4	1.7	2.2

6 References, Glossary and Abbreviations

6.1 References

AS/NZS ISO 31000:2009. *Risk management—Principles and guidelines*. Standards Australia, Sydney.

Asher, M.J., Croke, B.F.W., Jakeman, A.J., Peeters, L.J.M., 2015. *A review of surrogate models and their application to groundwater modeling*. Water Resources Research 51, 5957–5973. https://doi.org/10.1002/2015wr016967

Barnett B, Townley LR, Post V, Evans RE, Hunt RJ, Peeters L, Richardson S, Werner AD, Knapton A and Boronkay A 2012 (update in prep. 2022). *Australian groundwater modelling guidelines*. Waterlines report. National Water Commission, Canberra. Available [online]: https://www.groundwater.com.au/media/W1siZiIsIjIwMTIvMTAvMTcvMjFfNDFfMzZfOTYwX0F1c3RyYWxpYW5fZ3JvdW5kd2F0ZXJfbW9kZWxsaW5nX2d1aWRlbGluZXMucGRmIl1d/Australian-groundwater-modelling-guidelines.pdf accessed June 2022.

Bennett, N.D., Croke, B.F.W., Guariso, G., Guillaume, J.H.A., Hamilton, S.H., Jakeman, A.J., Marsili-Libelli, S., Newham, L.T.H., Norton, J.P., Perrin, C., Pierce, S.A., Robson, B., Seppelt, R., Voinov, A.A., Fath, B.D., Andreassian, V., 2013. *Characterising performance of environmental models*. Environmental Modelling and Software 40, 1–20. https://doi.org/10.1016/j.envsoft.2012.09.011

Bratvold, R. B., Bickel, J. E., & Lohne, H. P. (2009). Value of Information in the Oil and Gas Industry: Past, Present, and Future. *SPE Reservoir Evaluation & Engineering*, *12*(04), 630–638. https://doi.org/10.2118/110378-pa

Bredehoeft, J. (2005). The conceptualization model problem--surprise. *Hydrogeology Journal*, *13*(1), 37–46. https://doi.org/10.1007/s10040-004-0430-5

Chen, Y., Oliver, D.S., 2013. *Levenberg-Marquardt forms of the iterative ensemble smoother for efficient history matching and uncertainty quantification*. Computational Geosciences 17, 689–703. https://doi.org/10.1007/s10596-013-9351-5

Commonwealth of Australia (CoA) 2013. Significant Impact Guidelines 1.3: Coal seam gas and large coal mining developments—Impacts on water resources. Available [online]: https://www.dcceew.gov.au/environment/epbc/publications/significant-impact-guidelines-11-matters-national-environmental-significance. Accessed August 2022.

Cook, P.G., Miller, A., Shanafield, M., Simmons, C.T., 2016. *Predicting Water Resource Impacts of Unconventional Gas Using Simple Analytical Equations*. Groundwater. https://doi.org/10.1111/gwat.12489

Dausman, A. M., Doherty, J., Langevin, C. D., & Sukop, M. C. (2010). Quantifying data worth toward reducing predictive uncertainty. *Ground Water*, *48*(5), 729–740. https://doi.org/10.1111/j.1745-6584.2010.00679.x

Doherty J 2022. *Scientifically-Based, Decision-Support Groundwater Modelling: A Conceptual Framework, Draft of Part 1*. Available [online] https://gmdsi.org/blog/scientifically-based-decision-support-groundwater-modelling/ Accessed June 2022

Doherty J and Moore C 2019. *Decision Support Modeling: Data Assimilation, Uncertainty Quantification and Strategic Abstraction*. Groundwater, 58(3) 327-337 Available [online]: https://doi.org/10.1111/gwat.12969

Doherty J and Moore C 2021. *Decision Support Modelling Viewed through the Lens of Model Complexity*. A GMDSI Monograph. Available [online]: https://gmdsi.org/blog/monograph-decision-support-modelling-viewed-through-the-lens-of-model-complexity/ accessed June 2022.

Doody TM, Hancock PJ, Pritchard JL 2019. *Information Guidelines explanatory note; Assessing groundwater-dependent ecosystems.* A report prepared for the Independent Expert Scientific Committee on Coal Seam Gas and Large Coal Mining Development through the Department of the Environment and Energy, Commonwealth of Australia. Available [online]: https://iesc.environment.gov.au/publications/information-guidelines-explanatory-note-assessing-groundwater-dependent-ecosystems accessed June 2022.

Enemark, T., Peeters, L.J.M., Mallants, D., Batelaan, O., 2019. *Hydrogeological conceptual model building and testing: A review.* Journal of Hydrology 569, 310–329. https://doi.org/10.1016/j.jhydrol.2018.12.007

Evensen, G., 2018. *Analysis of iterative ensemble smoothers for solving inverse problems*. Computers and Geosciences 22, 885–908. https://doi.org/10.1007/s10596-018-9731-y

Freeze RA, Massmann J, Smith L, Sperling T and James B 1990. *Hydrogeological decision analysis:1. A framework*. Ground Water*,* 28(5): 738–766. https://doi.org/10.1111/j.1745-6584.1990.tb01989.x

Geological and Bioregional Assessment Program (2021) Fact sheet 18: Modelling of potential hydraulic impact pathways in the Beetaloo GBA region [online document]. Fact sheet for the Geological and Bioregional Assessment Program. Available [online]: https://www.bioregionalassessments.gov.au/sites/default/files/fact_sheet_18_modelling_of_potential_hydraulic_impact_pathways_in_the_beetaloo_gba_region_0.pdf accessed June 2022

Glover, R.E., Balmer, G.G., 1954. *River depletion resulting from pumping a well near a river.* Trans. Am. Geophys. Union 35, 468. https://doi.org/10.1029/tr035i003p00468

Gresswell R, Foley G and Faithful J. 2019. Hydrogeological modelling to inform closure planning for Hazelwood Mine. Mine Closure 2019, Fourie and Tibbett (eds), 2019 Australian Centre for Geomechanics, Perth. www.ghd.com/en/about-us/resources/106_Gresswell.pdf

Gosses, M., & Wöhling, T. (2021). Robust Data Worth Analysis with Surrogate Models. *Groundwater*, *59*(5), 728–744. https://doi.org/10.1111/gwat.13098

Hamilton, S.H., Pollino, C.A., Stratford, D.S., Fu, B., Jakeman, A.J., 2022. *Fit-for-purpose environmental modeling: Targeting the intersection of usability, reliability and feasibility*. Environmental Modelling and Software. 148, 105278. https://doi.org/10.1016/j.envsoft.2021.105278

Herckenrath, D., Doherty, J., Panday, S., 2015. *Incorporating the effect of gas in modelling the impact of CBM extraction on regional groundwater systems*. J. Hydrol. 523, 587–601. https://doi.org/10.1016/j.jhydrol.2015.02.012

Hill MC, Middlemis H, Hulme P, Poeter E, Riegger J, Neumann S, Williams H and Anderson M. 2004. *Brief overview of selected groundwater modelling guidelines*. Proceedings of FEM-Modflow: International conference on finite element models, Modflow and more. Karlovy Vary, Czech Republic, 13–16 September 2004. Available [online]: https://www.researchgate.net/profile/Hugh-Middlemis/publication/342302546_Hill_Middlemis_etal_2004_GWmodel_guidelines_overview/data/5eec2de9458515814a6ab53e/Hill-etal-2004-guidelines-overview.pdf accessed June 2022

Howe P, Moran C and Vink S. (2010). *Framework for assessing cumulative effects of mining operations on groundwater systems*. Water in Mining 2010, 2nd International Congress on Water Management in the Mining Industry, Chile, June 2010.

Huynh, T. and Hobbs, D. (2019). *Deriving site-specific guideline values for physico-chemical parameters and toxicants*. https://iesc.environment.gov.au/system/files/resources/249ff82e-f853-499b-ac06-d90726f8a394/files/information-guidelines-explanatory-note-site-specific-guidelines-values.pdf Accessed August 2022

IESC 2022. *Information Guidelines for proponents preparing coal seam gas and large coal mining development proposals*. Available [online]: http://iesc.environment.gov.au/publications/information-guidelines-independent-expert-scientific-committee-advice-coal-seam-gas accessed June 2022. Update in prep. 2022.

Maier, H.R., Guillaume, J.H.A., van Delden, H., Riddell, G.A., Haasnoot, M., Kwakkel, J.H., 2016. *An uncertain future, deep uncertainty, scenarios, robustness and adaptation: How do they fit together?* Environmental Modelling and Software 81, 154–164. https://doi.org/10.1016/j.envsoft.2016.03.014

Manewell, N., Doherty, J., 2021. *Rapid Data Assimilation Using an Appropriately Complex Model*. A GMDSI worked example report. Available [online]: https://gmdsi.org/blog/worked-example-simultaneous-interpretation-of-six-pumping-tests/ accessed June 2022

Marshall, S.K., 2021. *Understanding , detecting and conceptualising hydrogeologic barriers in groundwater modelling studies* PhD thesis, College of Science and Engineering, Flinders University. Available [online]: https://theses.flinders.edu.au/view/1643b2a4-73f6-4636-b6e5-fecf8f12192a/1 accessed June 2022.

Marshall, S.K., Cook, P.G., Miller, A.D., Simmons, C.T., Dogramaci, S., 2019. *The Effect of Undetected Barriers on Groundwater Drawdown and Recovery*. Groundwater 57, 718–726. https://doi.org/10.1111/gwat.12856

Mather, B., Müller, R.D., O'Neill, C., Beall, A., Vervoort, R.W., Moresi, L., 2022. *Constraining the response of continental-scale groundwater flow to climate change*. Scientific Reports 12, 1–16. https://doi.org/10.1038/s41598-022-08384-w

Middlemis H, Merrick N, Ross J, and Rozlapa K 2001. *Groundwater flow modelling guideline*. Report prepared for Murray–Darling Basin Commission by Aquaterra Consulting Pty Ltd. Available [online]: www.mdba.gov.au/sites/default/files/archived/mdbc-GW-reports/2175_GW_flow_modelling_guideline.pdf accessed June 2022

Middlemis, H., Peeters, L.J.M., 2018. *Uncertainty analysis—Guidance for groundwater modelling within a risk management framework*. A report prepared for the Independent Expert Scientific Committee on Coal Seam Gas and Large Coal Mining Development through the Department of the Environment and Energy, Commonwealth of Australia. Available [online]: https://iesc.environment.gov.au/publications/information-guidelines-explanatory-note-uncertainty-analysis accessed June 2022

Moore, C., Doherty, J., 2005. *Role of the calibration process in reducing model predictive error.* Water Resources Research 41, 1–14. https://doi.org/10.1029/2004WR003501

Morrison-Saunders, A. (2018). *Advanced introduction to environmental impact assessment*. Edward Elgar Publishing. 174p.

Murray TA and Power WL 2021. *Information Guidelines Explanatory Note: Characterisation and modelling of geological fault zones*. Report prepared for the Independent Expert Scientific Committee on Coal Seam Gas and Large Coal Mining Development through the Department of Agriculture, Water and the Environment, Commonwealth of Australia. Available [online] https://iesc.environment.gov.au/publications/information-guidelines-explanatory-note-characterisation-modelling-geological-fault-zones accessed June 2022

Peeters, L. J. M. 2017 *Assumption Hunting in Groundwater Modeling: Find Assumptions Before They Find You*, Groundwater, 55(5), pp. 665–669. https://10.1111/gwat.12565.

Peeters, L.J.M., Holland, K.L., Huddlestone-Holmes, C., Boulton, A.J., 2022. *A spatial causal network approach for multi-stressor risk analysis and mapping for environmental impact assessments.* Science of the Total Environment. 802, 149845. https://doi.org/10.1016/j.scitotenv.2021.149845

Peeters, L.J.M., Pagendam, D.E., Crosbie, R.S., Rachakonda, P.K., Dawes, W.R., Gao, L., Marvanek, S.P., Zhang, Y.Q., McVicar, T.R., 2018. *Determining the initial spatial extent of an environmental impact assessment with a probabilistic screening methodology*. Environmental Modelling and Software 109, 353–367. https://doi.org/10.1016/j.envsoft.2018.08.020

Peeters, L.J.M., Turnadge, C., 2019. *When to Account for Boundary Conditions in Estimating Hydraulic Properties from Head Observations?* Groundwater 57, 351–355. https://doi.org/10.1111/gwat.12880

Peterson, D.C., 2006. *Precaution: Principles and practice in Australian environmental and natural resource management.* Australian Journal of Agricultural and Resource Economics. 50, 469–489. https://doi.org/10.1111/j.1467-8489.2006.00372.x

Pianosi, F., Wagener, T., 2018. *Distribution-based sensitivity analysis from a generic input-output sample.* Environmental Modelling and Software 108, 197–207. https://doi.org/10.1016/j.envsoft.2018.07.019

Plischke, E., Borgonovo, E., Smith, C.L., 2013. *Global sensitivity measures from given data*. European Journal of Operational Research 226, 536–550. https://doi.org/10.1016/j.ejor.2012.11.047

Razavi, S., 2021. *Deep learning, explained: Fundamentals, explainability, and bridgeability to process-based modelling.* Environmental Modelling and Software 144, 105159. https://doi.org/10.1016/j.envsoft.2021.105159

Razavi, S., Jakeman, A., Saltelli, A., Prieur, C., Iooss, B., Borgonovo, E., Plischke, E., Lo Piano, S., Iwanaga, T., Becker, W., Tarantola, S., Guillaume, J.H.A., Jakeman, J., Gupta, H., Melillo, N., Rabitti, G., Chabridon, V., Duan, Q., Sun, X., Smith, S., Sheikholeslami, R., Hosseini, N., Asadzadeh, M., Puy, A., Kucherenko, S., Maier, H.R., 2021. *The Future of Sensitivity Analysis: An essential discipline for systems modeling and policy support.* Environmental Modelling and Software 137, 104954. https://doi.org/10.1016/j.envsoft.2020.104954

Richardson S, Irvine E, Froend R, Boon P, Barber S and Bonneville B 2011. *Australian groundwater-dependent ecosystem toolbox part 1: Assessment framework.* Waterlines report. National Water Commission, Canberra. Available

[online]:
http://webarchive.nla.gov.au/gov/20160615084626/http://archive.nwc.gov.au/library/waterlines/69-70 accessed August 2018.

Roy, T., Gupta, H., 2021. *How certain are our uncertainty bounds? Accounting for sample variability in Monte Carlo-based uncertainty estimates*. Environmental Modelling and Software 136, 104931. https://doi.org/10.1016/j.envsoft.2020.104931

Saltelli A 2022. Reckoning with uncertainty, social sciences lessons for mathematical modelling. In Dhersin J-S, Kaper H, Ndifon W, Roberts F, Rousseau C and Ziegler GM (Eds), *Mathematics for action: supporting science-based decision-making*, UNESCO [62236]. Available [online]: https://unesdoc.unesco.org/ark:/48223/pf0000380883.locale=en. Accessed June 2022

Saltelli, A., Annoni, P., 2010. *How to avoid a perfunctory sensitivity analysis*. Environmental Modelling and Software 25, 1508–1517. https://doi.org/10.1016/j.envsoft.2010.04.012

Saltelli, A., Bammer, G., Bruno, I., Charters, E., Di Fiore, M., Didier, E., Nelson Espeland, W., Kay, J., Lo Piano, S., Mayo, D., Pielke, R., Portaluri, T., Porter, T.M., Puy, A., Rafols, I., Ravetz, J.R., Reinert, E., Sarewitz, D., Stark, P.B., Stirling, A., van der Sluijs, J., Vineis, P., 2020. *Five ways to ensure that models serve society: a manifesto.* Nature 582, 482–484. Available [online]: https://doi.org/10.1038/d41586-020-01812-9

Scheidt, C., Caers, J., 2018. *Quantifying Uncertainty in Subsurface Systems*, Geophysical Monograph Series. John Wiley & Sons, Inc., Hoboken, NJ, USA. https://doi.org/10.1002/9781119325888.ch8

Schoups, G., Vrugt, J.A., 2010. *A formal likelihood function for parameter and predictive inference of hydrologic models with correlated, heteroscedastic, and non-Gaussian errors.* Water Resources Research 46, W10531. https://doi.org/10.1029/2009WR008933

Steel, D., 2014. *Philosophy and the precautionary principle: Science, evidence, and environmental policy.* Cambridge University Press. https://doi.org/10.1017/CBO9781139939652

Stefánsson, H.O., 2019. *On the Limits of the Precautionary Principle.* Risk Analysis 39, 1204–1222. https://doi.org/10.1111/risa.13265

Underschultz, J.R., Vink, S., Garnett, A., 2018. *Coal seam gas associated water production in Queensland: Actual vs predicted.* Journal Natural Gas Science and Engineering 52, 410–422. https://doi.org/10.1016/j.jngse.2018.02.010

Vrugt, J.A., Beven, K.J., 2018. *Embracing equifinality with efficiency: Limits of Acceptability sampling using the DREAM(LOA) algorithm*. Journal of Hydrology 559, 954–971. https://doi.org/10.1016/j.jhydrol.2018.02.026

Vrugt, J.A., Massoud, E.C., 2018. *Uncertainty Quantification of Complex System Models: Bayesian Analysis*, in: Duan, Q., Pappenberger, F., Wood, A., Cloke, H.L., Schaake, J.C. (Eds.), Handbook of Hydrometeorological Ensemble Forecasting. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 1–74. https://doi.org/10.1007/978-3-642-40457-3_27-1

WA Department of Mines, Industry, Regulation and Safety. 2020. Mine Closure Plan Technical Guidance. www.dmp.wa.gov.au/Environment/Mine-Closure-Plan-6034.aspx Accessed 6 July 2022.

Walker G. 2017. *Predictive uncertainty in groundwater modelling: how is it best used?* Discussion paper no.1 in Middlemis et al. 2019 (eds). *NCGRT national groundwater modelling uncertainty workshop*, Sydney, 10 July 2017. National Centre for Groundwater Research and Training, Australia. Available [online] https://doi.org/10.25957/5ca5641defe56 Accessed June 2022

Walker, W.E., Harremoes, P., Rotmans, J., van der Sluijs, J.P., van Asselt, M.B.A., Janssen, P., von Krauss, P.M., 2003. *Defining Uncertainty: A Conceptual Basis for Uncertainty Management in Model-Based Decision Support*. Integrated Assessment 4, 5–17. https://doi.org/10.1076/iaij.4.1.5.16466

White, J.T., 2018. *A model-independent iterative ensemble smoother for efficient history-matching and uncertainty quantification in very high dimensions*. Environmental Modelling Software 109, 191–201. https://doi.org/10.1016/j.envsoft.2018.06.009

White, J.T., Doherty, J.E., Hughes, J.D., 2014. *Quantifying the predictive consequences of model error with linear subspace analysis*. Water Resources Research 50, 1152–1173. https://doi.org/10.1002/2013WR014767

6.2 Glossary

The glossary provides short, succinct definitions of key terms used in the report. Some terms can have different meanings, depending on the context and the discipline, which means that definitions provided here may differ from definitions found elsewhere. This glossary is not intended to provide 'definitive' definitions, it is intended to clarify the use of terms within this Explanatory Note for uncertainty analysis.

Term	Description
Adaptive management	An approach to natural resource management that involves implementing management action, monitoring and evaluating outcomes and systematically adapting those actions according to what is learned (Morrison-Saunders, 2018)
Aleatory	Aleatory uncertainty is due to inherent variability or randomness.
Bias—availability	People tend to judge events that are easily recalled as more risky or more likely to occur than events that are not readily available to memory. An event may have more availability if it occurred recently, if it was a high-profile event, or if it has some other significance for an individual or group.
Bias—confidence	People typically have too much confidence in their own judgements. This appears to affect almost all professions, as well as the lay public. The few exceptions are people who receive constant feedback on the accuracy of their predictions, such as weather forecasters. The psychological basis for this unwarranted certainty seems to be insensitivity to the weaknesses in assumptions on which judgements are based.
Bias—confirmation	Confirmation bias is the filtering of new information to fit previously formed views. In particular, it is the tendency to accept as reliable new information that supports existing views, but to see as unreliable or erroneous and filter out new information that is contrary to current views. People may ignore or dismiss uncertainty information if it contradicts their current beliefs.
Bias—framing	How probabilistic information is framed can influence how that information is understood as well as the confidence that people have in the information. 'Priming' the brain with a particular stimulus can affect how it responds to a later stimulus. Using expressions that take advantage of this priming (i.e. the direction and expression are consistent) can reduce cognitive strain, which makes it easier for stakeholders to understand the idea presented without requiring further analysis. For example, the phrase

Term	Description
	'there is a 5 per cent chance the drawdown in the groundwater level will be greater than 0.2 m' may leave a different impression than the phrase 'there is a 95 per cent chance the water drawdown level will be less than 0.2 m'; even though the two phrases contain the same information. The latter requires less mental workload because your brain is already 'primed' to think about being 'down' when it hears 'less than'. This is particularly effective when paired with explicit advice about whether precautionary action is advised.
Bioregional assessments	A series of scientific analyses of the ecology, hydrology, geology and hydrogeology of a bioregion, with explicit assessment of the potential direct, indirect and cumulative impacts of CSG and coal mining development on water resources. The central purpose of bioregional assessments is to inform the understanding of impacts on and risks to water-dependent ecosystems that arise in response to current and future pathways of CSG and large coal mining development. See www.bioregionalassessments.gov.au
Causal pathway	The logical chain of events, either planned or unplanned, that link resource development and potential impacts on water resources and water-dependent ecosystems.
Consequence	Severity of an impact, often measured as the magnitude or extent.
Controls or Countermeasures	The methods or actions currently planned, or in place, to detect hazards when they occur or to reduce the likelihood and/or consequences of these hazards should they occur.
Cumulative impact	The total impact from multiple stressors and activities of a resource development on water resources when all past, present and reasonably foreseeable actions that are likely to impact on water resources are considered.
Data worth	The ability (or otherwise) of as-yet-ungathered data to reduce the uncertainties of decision-critical predictions using modelling methods (eg. PEST). See also https://gmdsi.org/blog/data-worth-analysis/
Decision maker	Authority that makes the decision whether or not to approve a development proposal. Results of a groundwater model as part of an environmental impact assessment can be part of the information considered by a decision-maker.
Ecohydrological Conceptual Model (ECM)	A type of conceptual model that integrates the hydrological components (surface and groundwater) and processes with the ecological components (e.g. specific taxa, communities and ecosystems) to show the likely pathways by which a proposed project might impact on key aspects of water resources (e.g. water quality, flow regime, biota, ecological function).
Epistemic	Epistemic uncertainty is uncertainty due to imperfect knowledge.
Failure	Tolerance of an unwanted outcome. See also Model Failure.

Term	Description
Fit for purpose	Hamilton et al. (2022) describe a model as fit for purpose when it is useful, reliable and feasible for its given context (see Figure 3 above). A groundwater model can be considered fit for purpose in the context of impact assessment for coal mines and coal seam gas developments when it provides: • objective evidence that the uncertainties affecting decision-critical predictions of impacts on aquifer resources and dependent systems are not underestimated, and; • information about uncertainties in the conceptualisations and model simulation outputs in a way that allows decision makers to understand the effects of uncertainty on project objectives (echoing the AS/NZS ISO 31000:2009 risk definition) and the effects of potential bias
Groundwater-dependent ecosystem (GDE)	Ecosystems that require access to groundwater on a permanent or intermittent basis to meet all or some of their water requirements so as to maintain their communities of plants and animals, ecological processes and ecosystem services. GDEs include terrestrial vegetation, wetlands (swamps, lakes and rivers) and ecosystems in aquifers and caves. The types and characteristics of GDEs are discussed further in the Explanatory Note for GDEs (Doody et al. 2019).
Hazard	An event, or chain of events, that might result in an effect (change in the quality or quantity of surface water or groundwater).
Hydrogeological conceptual model (HCM)	A simplified and idealised representation (usually graphical) of the physical hydrogeologic setting and our hydrogeological understanding of the essential flow processes of the system. This includes the identification and description of the geologic and hydrologic framework, media type, hydraulic properties, sources and sinks, and important aquifer flow and surface-groundwater interaction processes. See also Ecohydrological conceptual model.
Hypothesis	In the environmental risk management context where groundwater modelling is applied, the hypothesis to be tested typically comprises the conjecture of an unwanted outcome or consequence associated with a particular development and/or management strategy. In practice, the hypothesis should be clearly stated in terms of threshold impacts (preferably regulatory-based) and/or resource condition indicators, and should be closely linked with the specified modelling objectives. The hypothesis of an unwanted outcome can never be completely rejected.
Impact	A change resulting from 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 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).
Key Prediction	See Quantity of Interest.
Knowledge	The cumulative, often qualitative, understanding of the functioning of a groundwater system. Epistemic knowledge is based on interpretation of available geological and hydrogeological observations and information.

Term	Description
Likelihood	In risk management terminology, likelihood is the chance of something happening, whether defined, measured or determined objectively or subjectively, qualitatively or quantitatively, and described using general terms or mathematically (AS/NZS ISO 31000:2009)) In statistical literature, likelihood refers to how well a sample provides support for particular values of a parameter in a model. In Bayesian inference, the equivalent to the objective function is therefore called the likelihood function. See also probability.
Material	Pertinent or relevant.
Model—analytical	A model that provides an exact mathematical solution of a given problem by making simplifying assumptions (for example, that properties of the aquifer are considered to be constant in space and time).
Model—conceptual	A descriptive and/or schematic hydrological, hydrogeological and ecological representation of a site, environment or process showing the stores, flows and uses of water, which illustrates the geological formations, water resources and water-dependent ecosystems. It provides a basis for developing water and salt balances and inferring water-related ecological responses to changes in hydrology, hydrogeology and water quality.
Model—empirical	A model that uses algorithms or mathematical relationships that are based on observations/evidence (empiricism) but do not necessarily have a physical basis (e.g. regressions that do not necessarily establish a causal relationship).
Model—lumped	A model where hydrological processes are lumped to the catchment scale (no spatial variability within the catchment/domain).
Model—numerical	Computer codes that enable simulation of physical systems and processes such as groundwater and/or surface water flow and can be applied to assess the potential impacts of a project or management plan. Similar to analytical models in that they make simplifying assumptions; however, features of the governing equations and boundary conditions in numerical models (e.g. aquifer geometry, hydrogeological properties, pumping rates or sources of solute) can be specified as varying over space and time. This enables more complex representations of groundwater or surface water systems than could be achieved with an analytical model.
Model—physically based	A model with algorithms designed to realistically represent physical processes (e.g. depth-dependent ET).
Model complexity (Middlemis et al. 2001)	The degree to which a model application resembles, or is designed to resemble, the physical hydrogeological system (adapted from the model fidelity definition given in Ritchey and Rumbaugh 1996, cited in Middlemis et al. 2001). There are three main complexities (in order of increasing complexity): basic, impact assessment and aquifer simulator. Higher complexity models have a capability to provide for more complex simulations of hydrogeological processes and/or address resource management issues more comprehensively.

Term	Description
Model simplicity (effective)	The simplicity (or parsimony) principle implies that a conceptual model has been simplified yet retains enough complexity to adequately represent the physical system and its behaviour for the specified purpose of the model. The term 'effective model simplicity' was discussed by Voss (2011a, 2011b). Model simplification involves testing and removing all redundant elements of the model to which prediction is insensitive.
Non-uniqueness	The principle that many different possible sets of model inputs can produce nearly identical computed outputs for any given model.
Observation	A quantitative measurement of a groundwater system, such as groundwater pressure, streamflow or environmental tracer concentration.
Pathway	See Causal Pathway, and Source-Pathway-Receptor.
Probability	In statistical literature, probability refers to the chance that a particular outcome occurs based on the values of parameters in a model. In risk management terminology, probability is often used interchangeably with likelihood (AS/NZS ISO 31000:2009)
Probability density function / probability distribution function	The probability distribution of a random variable specifies the chance that the variable takes a value in any subset of the real numbers. For example: 'there is a probability of p that the variable is between x and y'.
Proponent (actor in engagement process)	The person or organisation that owns the project or development (e.g. a mine). The project proponent is asking the water manager to make a decision related to impacts on a water resource. The project proponent commissions studies by outside professionals such as consulting hydrogeologists.
Quantity of Interest	A quantity of interest is a model outcome from a specified model scenario, with a predefined spatial and temporal setting, that is relevant to assessing the likelihood and consequence of a causal pathway element representing a hazard. An alternative term is 'key prediction'
Range-of QoI	The predicted range for a quantity of interest produced by a quantitative uncertainty analysis.
Receptor	A point in the landscape where water-related impacts on assets are assessed.
Reviewer (actor in engagement process)	A person conducting an external review of a modelling study. The review may be more or less comprehensive depending on the requirements of the particular case. The reviewer is typically appointed by the water manager to support them to match the modelling capability of the modeller.
Risk (calculation)	Combination of consequence and likelihood (AS/NZS ISO 31000:2009).
Risk (definition)	Risk is the effect of uncertainty on management objectives. The level of risk is the magnitude of a risk, expressed in terms of the combination of consequences and their likelihood (AS/NZS ISO 31000:2009). The effects of uncertainty can be a positive or a negative deviation from the expected.

Term	Description
Risk (residual)	The potential occurrence of an adverse event that remains after controls or countermeasures have been applied to eliminate or treat identified risks.
Sensitivity analysis	The study of how uncertainty in the output of a model (numerical or otherwise) can be apportioned to different sources of uncertainty in the model input.
Significant impact	Defined by the Significant impact guidelines (CoA 2013) as an impact which is important, notable or of consequence, having regard to its context or intensity. Whether or not an action is likely to have a significant impact depends on the sensitivity, value and quality of the potentially threatened water resource, and on the intensity, duration, timing, magnitude and geographic extent of the impacts.
Source-Pathway-Receptor	A conceptual model for a hazard assessment of the water-affecting activities of a development (eg. drawdown due to pumping) via exposure pathways by which these effects may impact upon receptors (see also 'hazard', 'causal pathway' and 'receptor'). Central to the S-P-R approach is that an exposure or causal pathway must exist between a water-affecting activity (direct effect via groundwater and/or surface water response mechanisms) and a receptor (a user of water resources – indirect effect), otherwise the logical assessment process breaks down, rendering a particular risk redundant (Howe et al. 2010).
Stressor	A change in environmental or hydrological conditions that places stress on the health and functioning of an aquifer system or dependent ecosystem. Stressors can be natural or anthropogenic in origin, and direct or indirect in their effects.
Uncertainty—measurement error	Combination of uncertainties associated with the measurement of complex aquifer system states (heads, discharges), parameters and variability (3D spatial and temporal) with those induced by upscaling or downscaling (site-specific data, climate data).
Uncertainty—parameterisation	Uncertainties associated with hydrogeological property values and assumptions applied to represent complex reality in space and time (any system aspect that can be changed in an automated way in a model via parameterisation).
Uncertainty— predictive	The quantification of uncertainty in predictions. The bias and spread associated with model predictions that are made via a model that is consistent with the conceptual understanding of the system and associated measurements.
Uncertainty— scenario	Uncertainties associated with guessing future stresses, dynamics and boundary condition changes (e.g. mining, climate variability, land and water use change).
Uncertainty—structural/conceptual	Uncertainties associated with geological structure and hydrogeological conceptualisation assumptions applied to derive a simplified view of a complex hydrogeological reality (any aspect of a system that cannot be

Term	Description
	changed in an automated way in a model). See also Barnett et al. (2012, section 3.4).
Uncertainty (definition)	Uncertainty is the state, even partial, of deficiency of information related to the understanding or knowledge of an event, its consequence, or its likelihood (AS/NZS ISO 31000:2009).
Uncertainty (source/type)	Any deficiency in information relating to understanding or knowledge in four main classes/sources of uncertainty: 1. structural/conceptual uncertainty 2. parameter/input uncertainty 3. measurement error 4. scenario uncertainties.
Uncertainty analysis (qualitative)	A formal and structured discussion of all model choices and assumptions and their effect on predictions.
Uncertainty analysis (quantitative)	Quantitative uncertainty analysis seeks to identify a range of model predictions that are consistent with available knowledge and are constrained by historical observations.
Water balance	A mathematical expression of water flows and exchanges, described as inputs, outputs and changes in storage. Surface water, groundwater and atmospheric components should be included.
Water-dependent ecosystems	Water-dependent ecosystems are defined by the Water Act 2007 (Cth) as surface water ecosystems or groundwater ecosystems, and their natural components and processes, that depend on periodic or sustained inundation, waterlogging or significant inputs of water for their ecological integrity and includes ecosystems associated with a wetland, stream, lake or waterbody, salt marsh, estuary, karst system or groundwater system. A reference to a water-dependent ecosystem includes the biodiversity of the ecosystem. The Information Guidelines states that the IESC recognises that water-dependent ecosystems are captured under the definition of water resource (see below) but sees the value of retaining the term 'water-dependent ecosystem' because everyday usage of the term 'water resource' typically refers only to surface water or groundwater that is or can be exploited for human uses. Refer to the Information Guidelines for further detail.
Water manager (actor in engagement process)	The person or organisation responsible for the management or protection of the water resources, and thus of the modelling study and the outcome (the problem owner).
Water resource	Defined by the Water Act 2007 (Cth) (CoA 2007) as '*surface water or groundwater or a watercourse, lake, wetland or aquifer (whether or not it currently has water in it); and includes all aspects of the water resource, including water, organisms and other components and ecosystems that contribute to the physical state and environmental value of the water resource.*'

6.3 Abbreviations and acronyms

Short Form	Meaning
AGMG	Australian Groundwater Modelling Guidelines (Barnett et al. 2012)
BoM	Bureau of Meteorology
CoA	Commonwealth of Australia
CSG	Coal seam gas
EIS	Environmental Impact Statement
EPBC	Environment Protection and Biodiversity Conservation
ET or EVT	Evapotranspiration
GDE	Groundwater-dependent ecosystem
FEFLOW	Commercial software for simulation of saturated and unsaturated flow, mass transport (multiple solutes) and heat, using the finite element method. Can be coupled to MIKE 11 to simulate flow in river and stream networks. https://www.mikepoweredbydhi.com/products/feflow
GMDSI	Groundwater Modelling Decision Support Initiative (GMDSI.org)
HCM	Hydrogeological Conceptual Model
HydroGeoSphere	Commercial software for simulation of saturated and unsaturated flow, transport of mass and heat. Includes solution of 2D overland flow and 1D flow in river and stream networks. Also includes discrete fracture networks. https://www.aquanty.com/hydrogeosphere/
IES	Iterative Ensemble Smoother software for efficient history-matching and uncertainty quantification (open source, model-independent and scalable)
IESC	Independent Expert Scientific Committee on Coal Seam Gas and Large Coal Mining Development
Kh	Aquifer horizontal hydraulic conductivity [L/T]
Kv or Kz	Aquifer vertical hydraulic conductivity [L/T]
m	Metres
Mike-SHE	Commercial software for integrated modelling of surface water flow (2D overland flow and 1D stream flow networks) and groundwater flow (3D,

	unsaturated-saturated). https://www.mikepoweredbydhi.com/products/mike-she
ML	Megalitres
ML/d	Megalitres per day
MODFLOW	Modular groundwater flow modelling software (open source) developed by the US Geological Survey, regarded as industry standard (https://water.usgs.gov/ogw/modflow/). Refer to AGMG for more information on this and other groundwater modelling software packages.
MODHMS	Commercial software that is coupled with MODFLOW for integrated surface water and groundwater flow and solute transport simulations. Refer to AGMG for more information.
NCGRT	National Centre for Groundwater Research and Training
OGIA	Office of Groundwater Impact Assessment (Queensland)
PEST and PEST++	Parameter ESTimation software (open-source) for application to any model and often used in uncertainty analysis methods. http://pesthomepage.org/
S	Aquifer storativity [-]
Ss	Aquifer specific storage [L^{-1}]
Sy	Unconfined aquifer specific yield [-]
T	Transmissivity [L^2/T]

Australian Government

This initiative is funded by the Australian Government

www.iesc.environment.gov.au